universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation / Title of the Doctoral Thesis

Exploring quantum information with single photons

verfasst von / submitted by

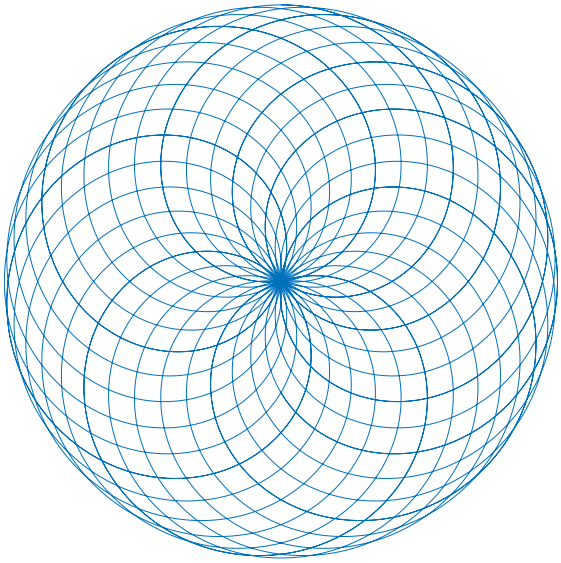
Karl Teodor Alexander Trygve Strömberg, MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, 2023 / Vienna 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 796 605 411
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Physik
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Ing. Dr. Philip Walther



Trajectory of a polarization state on the Bloch sphere.

Abstract

The field of experimental quantum information can be traced back to the pioneering tests of Bell's theorem carried out by Clauser and Aspect in the 1970s. These experiments explored the foundations of quantum theory and ruled out local hidden-variable models of quantum mechanics. Decades later Bell tests would find a practical use in establishing the security of quantum key distribution, remarkably enabling fully device-independent security. Today, as quantum information technologies stand poised to transform society, it is worth noting that many of the concepts core to their workings originated out of studies driven by fundamental curiosity.

This thesis presents several experimental works within the field of quantum information motivated by a similar curiosity and fascination by the unique and counter-intuitive aspects of quantum theory. Like the first experimental violations of Bell's inequality mentioned above, the experiments all made use of quantum states encoded in single photons. The first of the experiment is an experimental realisation of what is known as counterfactual communication, in which one party can communicate information to another party without directly transmitting any information-carrying particles. In particular, we present the first realisation of such a task in which the so-called weak trace of the information carrying particles vanishes in the laboratory of one party.

The second experiment deals with the question of how to reverse the time evolution of a quantum system. In classical systems this can be done by simply recording the evolution and undoing it step by step, but in the quantum domain the destructive nature of measurements forbids this. Time-reversal of quantum systems is nevertheless possible, and we present a realisation of a time-rewinding protocol for two-level quantum systems that works even under very weak assumptions on the nature of the time evolution and the ways in which one can interact with the system in question.

In the third experiment we study the utility of time-reversed processes in quantum information tasks. Taking inspiration from previous work on superpositions of temporal orders of quantum operations, we implement a superposition of temporal directions of quantum processes and show, in a semi-device independent way, that this leads to a quantifiable advantage in a tailored quantum information processing task. This advantage is relative to the circuit model of quantum computation, and suggests that this model does not fully capture the scope of information processing allowed by quantum theory.

The last experiment concerns the realisation of a process known as the quantum SWITCH, which is the most well-known example of a process with an indefinite causal order. Like the superposition of temporal directions discussed above, the quantum SWITCH can also provide an advantage over the quantum circuit model in performing certain tasks. Here we present a realisation of a quantum SWITCH that solves the issue of phase instability that plagued most previous implementations. The novel design of the quantum SWITCH was facilitated by development of a family of polarization rotators with unique properties under counterpropagation. We believe that these devices may also prove useful in other photonic quantum information contexts.

Zusammenfassung

Der Bereich der experimentellen Quanteninformation lässt sich bis zu den bahnbrechenden Tests des Bell'schen Theorems durch Clauser und Aspect in den 1970er Jahren zurückverfolgen. Diese Experimente untersuchten die Grundlagen der Quantentheorie und schlossen lokale Modelle der Quantenmechanik mit verborgenen Variablen aus. Jahrzehnte später fanden die Bell-Tests praktische Anwendung bei der Etablierung der Sicherheit von Quantenschlüsseln und ermöglichten auf bemerkenswerte Weise eine völlig geräteunabhängige Sicherheit. Heute, da die Quanteninformationstechnologien im Begriff sind, die Gesellschaft zu verändern, ist es erwähnenswert, dass viele der Konzepte, die ihrer Funktionsweise zugrunde liegen, aus Studien hervorgegangen sind, die von grundlegender Neugierde getrieben waren.

In dieser Arbeit werden mehrere experimentelle Arbeiten auf dem Gebiet der Quanteninformation vorgestellt, die durch eine ähnliche Neugier und Faszination für die einzigartigen und kontraintuitiven Aspekte der Quantentheorie motiviert waren. Wie die oben erwähnten ersten experimentellen Verletzungen der Bell'schen Ungleichung wurden bei allen Experimenten Quantenzustände verwendet, die in einzelnen Photonen kodiert sind. Das erste der Experimente ist eine experimentelle Umsetzung der so genannten kontrafaktischen Kommunikation, bei der eine Partei einer anderen Partei Informationen mitteilen kann, ohne direkt informationstragende Teilchen zu übertragen. Insbesondere stellen wir die erste Realisierung einer solchen Aufgabe vor, bei der die sogenannte schwache Spur der informationstragenden Teilchen im Labor einer Partei verschwindet.

Das zweite Experiment beschäftigt sich mit der Frage, wie man die zeitliche Entwicklung eines Quantensystems umkehren kann. In klassischen Systemen kann man dies tun, indem man die Entwicklung einfach aufzeichnet und Schritt für Schritt rückgängig macht, aber im Quantenbereich verbietet die zerstörerische Natur von Messungen dies. Die Zeitumkehr von Quantensystemen ist dennoch möglich, und wir stellen eine Realisierung eines Zeitumkehrprotokolls für Zweizustandssytem vor, das sogar unter sehr schwachen Annahmen über die Art der Zeitentwicklung und die Möglichkeiten der Interaktion mit dem betreffenden System funktioniert.

Im dritten Experiment untersuchen wir den Nutzen von zeitumgekehrten Prozessen bei Quanteninformationsaufgaben. In Anlehnung an frühere Arbeiten über Überlagerungen zeitlicher Ordnungen von Quantenoperationen implementieren wir eine Überlagerung zeitlicher Richtungen von Quantenprozessen und zeigen auf halb-geräteunabhängige Weise, dass dies zu einem quantifizierbaren Vorteil bei einer maßgeschneiderten Quanteninformationsverarbeitungsaufgabe führt. Dieser Vorteil ist relativ zum Quantenschaltungsmodell und deutet darauf hin, dass dieses Modell den Umfang der von der Quantentheorie erlaubten Informationsverarbeitungsweisen nicht vollständig erfasst.

Das letzte Experiment betrifft die Verwirklichung eines Prozesses, der als Quanten-SWITCH bekannt ist und das bekannteste Beispiel für einen Prozess mit unbestimmter kausaler Ordnung darstellt. Wie die oben beschriebene Überlagerung zeitlicher Richtungen kann auch der Quanten-SWITCH bei der Erfüllung bestimmter Aufgaben einen Vorteil gegenüber dem Quantenschaltungsmodell bieten. Hier stellen wir eine Realisierung eines Quanten-SWITCHs vor, die das Problem der Phaseninstabilität löst, das

die meisten früheren Implementierungen plagte. Das neuartige Design des Quanten-SWITCHs wurde durch die Entwicklung einer Familie von Polarisationsrotatoren mit einzigartigen Eigenschaften bei Gegenausbreitung erleichtert. Wir glauben, dass sich diese Geräte auch in anderen Bereichen der photonischen Quanteninformation als nützlich erweisen könnten.

Contents

List of publications

The following is a list of scientific articles produced throughout the duration of the doctoral studies. Articles contributing to this thesis are indicated in boldface.

- **Alonso Calafell, I.[†], Strömberg, T.[†], Arvidsson-Shukur, D.R.M., Rozema, L.A., Saggio, V., Greganti, C., Harris, N.C., Prabhu, M., Carolan, J., Hochberg, M., Baehr-Jones, T., Englund, D., Barnes, C. H. W. and Walther, P., Trace-free counterfactual communication with a nanophotonic processor. npj Quantum Information, 5(1), p.61. (2019).** https://doi.org/10.1038/s41534-019-0179-2

- Saggio, V., Asenbeck, B.E., Hamann, A., Strömberg, T., Schiansky, P., Dunjko, V., Friis, N., Harris, N.C., Hochberg, M., Englund, D., Wölk, S., Briegel, H. J. and Walther, P., Experimental quantum speed-up in reinforcement learning agents. Nature, 591(7849), pp.229-233. (2021). https://doi.org/10.1038/s41586-021-03242-7

- **Schiansky, P.[†], Strömberg, T.[†], Trillo, D., Saggio, V., Dive, B., Navascués, M. and Walther, P., Demonstration of universal time-reversal for qubit processes. Optica, 10(2), pp.200-205. (2023).** https://doi.org/10.1364/OPTICA.469109

- **Strömberg, T., Schiansky, P., Peterson, R.W., Quintino, M.T. and Walther, P., Demonstration of a quantum switch in a Sagnac configuration. Physical Review Letters, 131(6), p.060803. (2023).** https://doi.org/10.1103/PhysRevLett.131.060803

- **Strömberg, T., Schiansky, P., Quintino, M.T., Antesberger, M., Rozema, L., Agresti, I., Brukner, Č. and Walther, P., Experimental superposition of a quantum evolution with its time reverse. arXiv preprint arXiv:2211.01283. (2023).** https://doi.org/10.48550/arXiv.2211.01283

- Silvestri, R., Yu, H., Strömberg, T., Hilweg, C., Peterson, R.W., Walther, P., Experimental Observation of Earth's Rotation with Quantum Entanglement. arXiv preprint arXiv:2310.16903. (2023). https://doi.org/10.48550/arXiv.2310.16903

[†]These authors contributed equally to this work

List of Figures

Introduction

The development of quantum theory in the early 20th century was motivated by a need to explain seeming contradictions, such as the wave-particle duality implied by the photoelectric effect, and the ability for electrons to stay bound to a positively charged atomic nucleus. While quantum mechanics was successful in providing a framework for how to understand these effects, it did not demystify them. In fact, quantum mechanics soon provided new puzzles, such as the concept of entanglement, first identified by Einstein, Podolsky and Rosen in their famous thought experiment [1].

In the decades that followed there was little interest within the field to seriously explore these concepts and related open questions within the foundations of quantum mechanics such as the measurement problem. This was in part due to how disconnected these topics were from the unanswered questions in experimental physics at the time. However, beginning with the discovery of Bell's theorem in 1964 [2], and the experimental tests thereof performed in the decades that followed [3], the possibility of testing and even exploiting more of the unique features of quantum mechanics began to emerge.

Motivated by theoretical discoveries and realisations about the potential utility of quantum systems in information processing tasks, such as Feynman's proposal for quantum simulators [4], Bennett and Brassard's quantum key distribution [5], and Shor's algorithm for integer prime factorization [6], the field grew. In span of a few decades the experimental level of control over isolated quantum systems progressed rapidly, with many landmark demonstrations along the way, including the first experimental teleportation of a quantum state [7] and the demonstration of quantum supremacy [8].

The development of experimental methods for creating and controlling quantum states has at the same time enabled studies of concepts of a more fundamental nature in quantum theory, such as the quantum zeno effect in which repeated measurement of a system freezes it in time [9], the nonlocality of single quantum particles [10], and more recently the causal structure of quantum processes [11]. While such studies may have, and in some cases have had, technological applicability, this is not what motivated them. These types of studies were instead motivated by a fundamental interest in exploring the full riches of quantum theory.

In this thesis I present four works on quantum information that continue in this spirit, and that focus on the experimental realisation of various tasks that do not have a classical analogue.

The thesis begins with a chapter giving an introduction to quantum information, outlining the description of quantum states and their transformations, as well as some of the theoretical tools used to describe and analyse the presented experiments. Common to all these experiments is that they were realised using single photon states as carriers of quantum information. The second chapter therefore provides an overview of the quantum theory of light, and methods for both the generation and manipulation of single-photon states, with particular attention being devoted to the latter. All but one of the experiments made partial use of the polarization degree of freedom of single photons

to encode quantum information, and in the third chapter a detailed examination of the properties of polarization rotators, used to manipulate the polarization states of single photons, is presented. A set of hitherto unknown polarization rotators are introduced, their properties are studied mathematically, and proofs of universality are written down where applicable. The chapter concludes with a discussion of some of the potential applications of these devices.

The fourth chapter presents background detail on an experiment on counterfactual communication, which is included in Chapter 8. This experiment was implemented using a programmable nanophotonic waveguide, and the chapter discusses the experimental characterisation of the device, the mapping of the communication protocol onto the waveguide, as well as the relevant aspects of the protocol itself.

In the fifth chapter the second publication, presented in Chapter 9, is discussed. This article concerns the realisation of a time-rewinding protocol that works for two-level quantum systems. The chapter gives a description of the protocol from a theoretical perspective, and places it in the context of other similar protocols. A detailed description of the experimental setup and implementation of the protocol is then presented, concluding with an extended analysis of the data taking into account experimental imperfections.

The sixth chapter discusses the experimental realisation of a superposition of time directions for two quantum channels, which was the subject of the article included in Chapter 10. The chapter begins with a discussion of non-quantum-circuit-model processes, and introduces a specific example of such a process: the quantum time flip. It then presents how the quantum time flip can outperform a broad class of quantum processes in a certain task, before moving on to discuss how the concepts introduced in Chapter 3 can be used to experimentally realise this advantage. Finally, the details of the implementation are presented, together with an analysis and interpretation of the data.

In the seventh chapter the final publication, included in Chapter 11, is discussed. This article focused on demonstrating the applicability of the devices presented in Chapter 3 to the photonic realisation of a specific quantum process: the quantum SWITCH. The chapter opens with an introduction to the quantum switch, and presents an overview of its previous photonic realisations. It is then shown how the devices from Chapter 3 can simplify and improve such realisations, and the specific device employed in the experiment is re-examined. The experiment deign and device characterisation is presented, and main results of the article are discussed. Finally, the potential applicability of the novel polarization rotators to quantum switch experiments with more than two parties is explored.

1 Quantum Information

This chapter will give a largely self-contained introduction to most of the concepts in quantum information needed to understand the experiments presented in this thesis. It begins with a discussion on qubits and how to manipulate them, before moving on to some of the theoretical tools used to describe and characterise quantum states and processes. Finally, the chapter concludes with a discussion on counterfactual communication and weak measurements relevant to Publication 1.

1.1 Introduction

The field of quantum information centres on studying how the laws of quantum mechanics allow for information to be encoded, manipulated and distributed. The connection between information and physics is not an entirely obvious one. It traces its roots back to Claude Shannon, who in the 1940s began to formulate a theory of information, and in a seminal paper defined what is now known as Shannon entropy [12]. As the name suggest, this concept is closely related to the concept of entropy in statistical mechanics and in 1961 Rolf Landauer showed that, in analogy to how irreversible processes necessarily increase the entropy of a system, there is an energy cost associated with the irreversible manipulation of information [13]. The deep connection between information and physics fundamentally rests on the fact that all information is ultimately encoded in the degrees of freedom of some physical system. From this perspective, the idea of quantum information emerges naturally when one considers ever smaller information carriers, and go from quantities that can be described macroscopically, such as the voltage across a transistor, to states that require a quantum mechanical description, for instance the spin of an electron. Since the states of the latter behave in a fundamentally different way from classical states, so will the information encoded in those states.

1.1.1 The qubit

The fundamental unit of information of quantum information is the eponymous quantum bit, or qubit [14]. In analogy with so-called classical information, which is made up of binary states $b \in \{0, 1\}$ called bits, the qubit is a two level quantum mechanical system:

$$|\Psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,, \tag{1.1-1}$$

where α and β are the complex probability amplitudes of the two states $|0\rangle$ and $|1\rangle$, and obey the normalization condition

$$|\alpha|^2 + |\beta|^2 = 1. \tag{1.1-2}$$

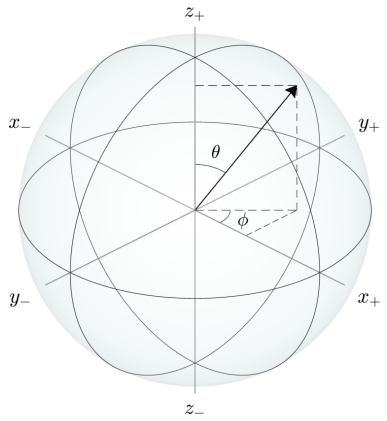


Figure 1.1: **Bloch sphere.** Pure single-qubit states can be thought of as unit vectors in 3D with polar angle θ and azimuth ϕ. The corresponding unit sphere is known as the Bloch sphere. This picture omits the global phase γ, and the reason for this is that this phase does not have any physical meaning when only considering a single qubit; it does not affect any measurement outcomes and is therefore not measurable. Mixed states correspond to vectors with a length less than unity, and the maximally mixed state is the zero vector. The x-, y-, and z-axes are chosen as the positive eigenvectors of the corresponding Pauli matrices (c.f. (1.1-15)).

Since a qubit is an element of $\mathbb{C}^2$ the kets $|0\rangle$ and $|1\rangle$ can equivalently be written as vectors:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{1.1-3}$$

From (1.1-1) it is already evident that the qubit differs significantly from a classical bit, since it can take on values between 0 and 1. The probabilities for a measurement of the qubit to reveal the outcomes 0 and 1 are given by the square of their respective probability amplitudes. The phenomenon of quantum systems inhabiting combinations of discriminable states is called superposition, and is a key resource in quantum information processing. Combining the normalization condition in (1.1-2) with (1.1-1) we see that the qubit state can also be written the following way:

$$|\Psi\rangle = e^{i\gamma}\left[\cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle\right]. \tag{1.1-4}$$

This description suggests that a qubit state can be represented as a unit vector in 3D space with polar angle θ and azimuth ϕ, while the factor $e^{i\gamma}$, which does not have a geometric interpretation, is needed to encode the global phase of the state. The corresponding unit sphere, illustrated in Fig. 1.1, is referred to as the Bloch sphere and is a powerful tool for thinking about qubit states. The vector $\vec{v} \in \mathbb{R}^3$:

$$|\Psi\rangle = \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) \\ e^{i\phi}\sin\left(\frac{\theta}{2}\right) \end{bmatrix} \leftrightarrow \vec{v} = \begin{bmatrix} \cos(\phi)\sin(\theta) \\ \sin(\phi)\sin(\theta) \\ \cos(\theta) \end{bmatrix} \tag{1.1-5}$$

is known as the Bloch vector.

1.1.2 The density operator

For a more general description of qubit states it is necessary to introduce the density matrix, or density operator. States on the form of (1.1-1), which are unit vectors on the Bloch sphere, have a density matrix given by

$$\rho = |\Psi\rangle\langle\Psi| \tag{1.1-6}$$

where $\langle\Psi| = |\Psi\rangle^\dagger$, while a general density matrix can be written

$$\rho = \sum_i p_i \, |\Psi_i\rangle\langle\Psi_i| = \sum_{ij} \rho_{ij} \, |i\rangle\langle j| \, . \tag{1.1-7}$$

The prefactors p_i correspond to the probability of observing the quantum system in the state $|\Psi_i\rangle$. States that obey

$$\mathrm{Tr}\left[\rho^2\right] = 1 \tag{1.1-8}$$

are called pure states, and correspond to states which can be written as a state vector. States for which the purity $P = \mathrm{Tr}[\rho^2]$ is less than 1 are called mixed states, and can only be described by a density matrix. Such states have a Bloch vector with a length less than one; more specifically

$$\|\vec{v}\|^2 = 2\mathrm{Tr}\left[\rho^2\right] - 1. \tag{1.1-9}$$

The probabilities p_i in (1.1-7) can be thought of as representing ignorance of which state the system is actually in, as opposed to a fundamental uncertainty in the state itself, as in the case of a quantum superposition state. A density matrix describing a physical state has to obey

$$\mathrm{Tr}[\rho] = 1 \tag{1.1-10}$$

$$\rho^\dagger = \rho \tag{1.1-11}$$

$$\rho \geq 0 \tag{1.1-12}$$

where the condition (1.1-10) is equivalent to the probabilities p_i summing to 1, (1.1-11) ensures that the probabilities are real, and (1.1-12) guarantees that the probabilities are non-negative.

1.1.3 Pauli matrices

For the description of qubits it is helpful to introduce the so-called Pauli matrices [15], which are a set of three Hermitian and unitary 2×2 matrices named after Wolfgang Pauli. The matrices are

$$\sigma_x = X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma_y = Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad \sigma_z = Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{1.1-13}$$

The Pauli matrices obey

$$\sigma_i \sigma_j = \delta_{ij} I + i \sum_k \epsilon_{ijk} \sigma_k, \tag{1.1-14}$$

where I is the identity operator and ϵ_{ijk} is the Levi-Civita symbol.[1] Using these matrices it is possible to describe the manipulation and measurement of qubits. The eigenvectors and eigenvalues of the Pauli matrices are

$$X: \quad \begin{aligned} |x_+\rangle &= \frac{|0\rangle + |1\rangle}{\sqrt{2}} = |+\rangle\,, \qquad \lambda_+ = 1 \\ |x_-\rangle &= \frac{|0\rangle - |1\rangle}{\sqrt{2}} = |-\rangle\,, \qquad \lambda_- = -1 \end{aligned} \tag{1.1-15}$$

$$Y: \quad \begin{aligned} |y_+\rangle &= \frac{|0\rangle + i\,|1\rangle}{\sqrt{2}} = |L\rangle\,, \qquad \lambda_+ = 1 \\ |y_-\rangle &= \frac{|0\rangle - i\,|1\rangle}{\sqrt{2}} = |R\rangle\,, \qquad \lambda_- = -1 \end{aligned} \tag{1.1-16}$$

$$Z: \quad \begin{aligned} |z_+\rangle &= |0\rangle = |H\rangle\,, \qquad \lambda_+ = 1 \\ |z_-\rangle &= |1\rangle = |V\rangle\,, \qquad \lambda_- = -1. \end{aligned} \tag{1.1-17}$$

The right-hand labels $\{+, -, L, R, H, V\}$ are the ones commonly used for qubits encoded in photon polarization. The eigenvectors with positive eigenvalues correspond to the x-, y- and z-axes of the Bloch sphere, respectively. The density matrix of a general qubit state can be decomposed in terms of the Pauli matrices as

$$\rho = \frac{1}{2}\left(I + n_x X + n_y Y + n_z Z\right) = \frac{I + \vec{n}\cdot\vec{\sigma}}{2}, \tag{1.1-18}$$

where $\vec{n} = \begin{bmatrix} n_x & n_y & n_z \end{bmatrix}$ and $\vec{\sigma} = \begin{bmatrix} X & Y & Z \end{bmatrix}$ is called the Pauli vector.

1.1.4 Measuring qubits

Measurements in quantum mechanics are described by operators called observables. The eigenvalues of an observable determine the possible measurement outcomes, and in order to ensure real-valued measurement outcomes every operator associated with an observable is required to be Hermitian. The most straightforward type of measurement is called a projective measurement, and can be written as

$$M = \sum_i \lambda_i P_i = \sum_i \lambda_i\, |\psi_i\rangle\langle\psi_i| \tag{1.1-19}$$

where M is the observable, λ_i are its eigenvalues and P_i are orthogonal projectors that sum to identity, meaning they obey

$$P_i^2 = P_i \tag{1.1-20}$$

$$P_i P_j = \delta_{ij} P_i \tag{1.1-21}$$

$$\sum_i P_i = I. \tag{1.1-22}$$

A measurement of the observable M on the state $|\Psi\rangle$ projects it onto the state $|\psi_i\rangle$ with a probability given by

$$p(\lambda_i) = \langle\Psi|P_i|\Psi\rangle = |\langle\Psi|\psi_i\rangle|^2, \tag{1.1-23}$$

[1] The notation I and $\mathbb{1}$ will be used interchangeably.

and the property (1.1-22) is necessary to guarantee that the probabilities of the different possible measurement outcomes add up to unity. If the state is described by a density operator the measurement probabilities are given by

$$p(\lambda_i) = \mathrm{Tr}\left[P_i^\dagger P_i \rho\right]. \tag{1.1-24}$$

This is known as the Born rule, after Max Born. Immediately after the measurement, the post-measurement state of the system is no longer $|\Psi\rangle$, but rather the state $|\psi_i\rangle$ that was projected onto. The act of measurement itself therefore changes the state of the system. An important consequence of this is that it is impossible to extract all the information about a quantum state, since even the simple qubit state in (1.1-4) has two parameters, while the measurement only reveals a single one. More generally, given two observables M and N describing two properties of a quantum state, it is only possible to measure these properties simultaneously if the observables commute: $[M, N] = 0$, since this condition guarantees the existence of a simultaneous eigenbasis for the two operators. If the observables do not commute ($[M, N] \neq 0$) it means that the order in which the measurements are performed affects the measurement outcome.

In addition to the probabilities of individual measurement outcomes, it is often useful to speak about the expected, or average value of a measurement with respect to a state. The expression from this follows immediately from (1.1-19) and (1.1-23):

$$\langle M \rangle_\Psi = \sum_i \lambda_i p(\lambda_i) = \langle \Psi | M | \Psi \rangle. \tag{1.1-25}$$

For a state described by a density matrix the corresponding equation is

$$\langle M \rangle_\rho = \mathrm{Tr}[M\rho]. \tag{1.1-26}$$

The expectation value of an observable does not reflect the average value obtained by repeated measurement on a single quantum state, since the projective nature of the measurements ensures that repeated ones always yield the same result. Instead it should be thought of as the average measurement outcome of identical measurements on a set of identically prepared quantum states.

The Pauli matrices from (1.1-13) are Hermitian, and are therefore valid observables. Measurements of these observables correspond to projecting a qubit state onto the x-, y- or z-axes of the Bloch sphere and the expectation value of a given operators can be understood as the geometric projection of the Bloch vector onto the axis defined by the eigenstates of the observable. For this reason we can also write the Bloch vector of a general qubit state ρ as

$$\vec{v} = \begin{bmatrix} \mathrm{Tr}[X\rho] \\ \mathrm{Tr}[Y\rho] \\ \mathrm{Tr}[Z\rho] \end{bmatrix}. \tag{1.1-27}$$

This is exactly the vector $\vec{n}$ in the density matrix composition (1.1-18).

The measurement bases defined by the Pauli operators occur frequently in quantum information, and in particular the one corresponding to the Z operator is often referred

to as the computational basis, since it projects any qubit state onto $\{|0\rangle, |1\rangle\}$ and therefore reveals either a logical 0 or 1.

Many physical properties of quantum states are hard to measure directly. Take for example the spin of an electron; because of its small magnitude it has an exceptionally small influence on the macroscopic world and is therefore very challenging to detect. A way around this is to couple the spin, or some other quantity to be measured, to another degree of freedom, commonly the path or position of the particle. A canonical example of this is the famous Stern-Gerlach experiment, in which a magnetic field is used to deflect an electron in different directions depending on its spin magnetic moment [16]. A measurement of the electron's position then reveals its spin, since the two quantities have become correlated. This type of measurement is referred to as a Von-Neumann measurement [17].

1.1.5 Single qubit gates

The time evolution of quantum states is described by unitary operators, something which is a consequence of the Hermiticity of the Hamiltonian H in the Schrödinger equation governing the evolution of quantum states:

$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = H|\Psi(t)\rangle. \tag{1.1-28}$$

Let $U(t, t_0)$ be a state-independent time-evolution operator[2] from time $\tau = t_0$ to $\tau = t$ under the Hamiltonian H, then

$$U(t,0)|\Psi(0)\rangle = |\Psi(t)\rangle, \tag{1.1-29}$$

$$i\hbar\frac{d}{dt}U(t,0)|\Psi(0)\rangle = HU(t,0)|\Psi(0)\rangle. \tag{1.1-30}$$

Since (1.1-30) holds for any state $|\Psi\rangle$ it follows that the time-evolution operator also obeys the Schrödinger equation:

$$i\hbar\frac{d}{dt}U(t,0) = HU(t,0). \tag{1.1-31}$$

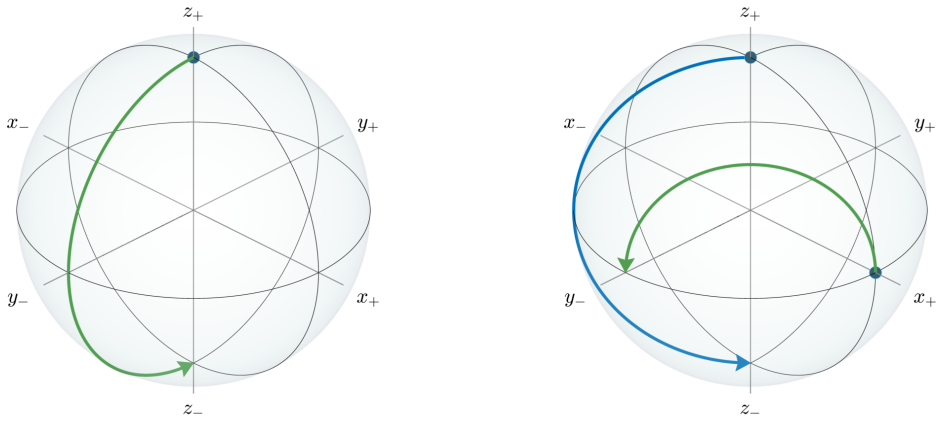
Taking the adjoint of this equation we find

$$-i\hbar\frac{d}{dt}U(t,0)^\dagger = U(t,0)^\dagger H^\dagger. \tag{1.1-32}$$

We now use these two equations to show that $U(t,0)^\dagger U(t,0)$ is time independent:

$$\begin{aligned} i\hbar\frac{d}{dt}U(t,0)^\dagger U(t,0) &= i\hbar\big(-U(t,0)^\dagger H^\dagger U(t,0) + U(t,0)^\dagger H U(t,0)\big) \\ &= i\hbar U(t,0)^\dagger (H - H^\dagger) U(t,0) \\ &= 0. \end{aligned} \tag{1.1-33}$$

[2]The Schrödinger equation for the time-evolution operator can also be derived without explicitly assuming the state independence of the operator, by instead taking the Hamiltonian as the generator of time translations [18].



(a) X-gate (b) $\frac{X-Y}{\sqrt{2}}$

Figure 1.2: **Visualization of qubit unitaries.** Qubit unitaries can be thought of as rotations on the Bloch sphere, where the axis of rotation is given by the eigenvectors of the unitary and the rotation angle is the difference between the complex arguments of the eigenvalues. In (a) the simple case of an X-gate is shown, which rotates any state about the x_+ axis by 180° in a counter-clockwise direction. Similarly, in (b) a rotation about the axis between x_+ and y_- is shown. The blue dots indicate the initial states, and the green and blue traces illustrate the trajectory of the states as the rotation angle is continuously varied.

Since $U(0,0) = I$ and therefore $U(0,0)^\dagger U(0,0) = I$, the time independence of this quantity then imples that $U(t,0)^\dagger U(t,0) = I$ for every time t, showing the unitarity.

If the Hamiltonian H is time independent then the time evolution of a system can be written down directly, and is given by the operator

$$U = e^{-\frac{i}{\hbar}Ht}. \tag{1.1-34}$$

Unitary operations can be thought of as rotations in rotations in the space of state vectors, and for qubits in particular they can be thought of as rotations on the Bloch sphere. For state vectors, these operators act directly:

$$U\left|\Psi\right\rangle = \left|\Psi'\right\rangle, \tag{1.1-35}$$

while for density matrices they 'sandwich' the state:

$$\sum_i U\left|\Psi_i\right\rangle\!\left\langle\Psi_i\right| U^\dagger = U\rho U^\dagger = \rho'. \tag{1.1-36}$$

Since the Pauli matrices are unitary in addition to being Hermitian, they also represent valid time evolutions of quantum states. In the context of information processing and computing these time evolutions are often referred to as gates, in analogy to operations acting on classical bits. The X gate, as an example, effects a bit flip in the computational basis:

$$X\left(\alpha\left|0\right\rangle + \beta\left|1\right\rangle\right) = \alpha\left|1\right\rangle + \beta\left|0\right\rangle. \tag{1.1-37}$$

On the Bloch sphere this represents a rotation by 180° around the x-axis (see Fig. 1.2a). The other Pauli operators similarly correspond to 180° rotations about their respective axes. A rotation about one of the coordinate axes by an arbitrary angle can be achieved by exponentiating the Pauli operators. For instance, a rotation around Z by an angle θ can be written as

$$R_z(\theta) = e^{-i\frac{\theta}{2}Z} = \cos\left(\frac{\theta}{2}\right) I - i \sin\left(\frac{\theta}{2}\right) Z. \tag{1.1-38}$$

More generally, a rotation around an axis $\vec{n}$ on the Bloch sphere (see Fig. 1.2b) is given by

$$R_{\vec{n}}(\theta) = e^{-i\frac{\theta}{2}\vec{n}\cdot\vec{\sigma}} = \cos\left(\frac{\theta}{2}\right) I - i \sin\left(\frac{\theta}{2}\right) \vec{n}\cdot\vec{\sigma}. \tag{1.1-39}$$

Such rotations belong to the special unitary group of 2×2 matrices,[3] abbreviated as SU(2), and the operators $\{i\sigma_x, i\sigma_y, i\sigma_z\}$ are said to be the generators of the group. Since they correspond[4] to rotations in 3D space, any rotation about an arbitrary axis can be decomposed in terms of three rotations about a set of fixed axes, either using the Euler angle decomposition

$$R_{\vec{n}} = R_i(\theta) R_j(\phi) R_i(\gamma), \tag{1.1-40}$$

or the Tait-Bryan decomposition

$$R_{\vec{n}} = R_i(\theta) R_j(\phi) R_k(\gamma), \tag{1.1-41}$$

where i, j, k are three orthogonal axes. In addition to the rotation in (1.1-39) an arbitrary qubit unitary can also contain a global phase factor:

$$U = e^{i\phi} R_{\vec{n}}(\theta). \tag{1.1-42}$$

1.1.6 Multi-qubit states

While single-qubit states are elements of $\mathbb{C}^2$, multi-qubit states live in spaces that can be decomposed as tensors products of single-qubit spaces. A trivial example of a two-qubit state is

$$|0\rangle \otimes |0\rangle = |0\rangle\,|0\rangle = |00\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2. \tag{1.1-43}$$

While the space of a two-qubit state can be written as a tensor product of single-qubit spaces, the same is not always true for the states themselves. For example, the state

$$|\Phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \tag{1.1-44}$$

cannot be factored as $|\Psi_1\rangle \otimes |\Psi_2\rangle$, and states of this type are called entangled, whereas states that can be factored are called separable. The state in (1.1-44) is one of the four

[3] This group consists of all 2×2 unitary matrices with determinant 1.

[4] The group of rotations in 3D is called SO(3), and is technically not isomorphic to SU(2), but instead there exists a two-to-one group homomorphism from SU(2) to SO(3). Loosely speaking, this is due to the fact that for example Z and $-Z$ are distinct elements in SU(2), since a qubit can have a global phase, but in SO(3) they describe the same rotation. In other words, for qubits rotating 180° clockwise and counter-clockwise is not the same.

so-called Bell states, the other three being

$$|\Phi^-\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}} \tag{1.1-45}$$

$$|\Psi^+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}} \tag{1.1-46}$$

$$|\Psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{1.1-47}$$

The corresponding density matrices are

$$|\Phi^+\rangle\langle\Phi^+| = \frac{1}{4}\left[I + X\otimes X - Y\otimes Y + Z\otimes Z\right] \tag{1.1-48}$$

$$|\Phi^-\rangle\langle\Phi^-| = \frac{1}{4}\left[I - X\otimes X + Y\otimes Y + Z\otimes Z\right] \tag{1.1-49}$$

$$|\Psi^+\rangle\langle\Psi^+| = \frac{1}{4}\left[I + X\otimes X + Y\otimes Y - Z\otimes Z\right] \tag{1.1-50}$$

$$|\Psi^-\rangle\langle\Psi^-| = \frac{1}{4}\left[I - X\otimes X - Y\otimes Y - Z\otimes Z\right]. \tag{1.1-51}$$

These states form a basis for $\mathbb{C}^2\otimes\mathbb{C}^2$ and are maximally entangled, which quantitatively means

$$\mathrm{Tr}_A[\rho_{A,B}] = \frac{1}{2}I_B. \tag{1.1-52}$$

Here, A and B label the two constituent subsystems, and Tr_A refers to the partial trace over the system A:

$$\mathrm{Tr}_A[\rho_{A,B}] = \sum_j \langle j|_A\, \rho_{A,B}\, |j\rangle_A\,. \tag{1.1-53}$$

This can be understood as the local description of the system B, without any information about the system A. What (1.1-52) says then is that the local description of one half of a maximally entangled state is that of the maximally mixed state, for which

$$\mathrm{Tr}[MI] = 0, \quad M = \vec{n}\cdot\vec{\sigma}, \tag{1.1-54}$$

meaning that the expectation value of Pauli any observable is zero, and any measurement on the maximally mixed state gives a random result.[5] The inseparability of a maximally entangled state therefore means that a description of the component subsystems is not possible, in the sense that one cannot make predictions about the outcomes of measurements on the local subsystem. On the other hand, as (1.1-48)–(1.1-51) suggest, it is possible to make predictions about the correlations between the measurement outcomes of the local subsystems. In the case of the state $|\Psi^-\rangle$, for instance, every measurement of the form: $M = \vec{n}\cdot\vec{\sigma}\otimes\vec{n}\cdot\vec{\sigma}$, i.e. measurements in the same basis for the two subsystems, will give anti-correlated measurement outcomes for the two subsystems.

[5] A qubit observable can also contain an I term, which shifts the expectation value away from zero. In this case the expectation value of the observable on a maximally mixed state is the average of the eigenvalues, but the measurement outcomes are still random.

This property persists over arbitrarily large distances, and is one of the key features of quantum systems that allow them to outperform classical ones in certain information distribution and processing tasks [19, 20]. Interestingly, while the state $|\Psi^-\rangle$ is anti-correlated in every basis, there is no two-qubit state that is correlated in every basis, since the corresponding density matrix

$$\rho_c = \frac{1}{4}\left[I + X \otimes X + Y \otimes Y + Z \otimes Z\right], \tag{1.1-55}$$

has a negative eigenvalue. One reason for this is that the gate which takes $|\Psi^-\rangle$ to the maximally correlated state is $I \otimes U_{\text{NOT}}$, where U_{NOT} is the universal-NOT gate for a single qubit. This gate has the effect of transforming a state into its orthogonal complement, which for a density matrix can be expressed as

$$\rho \mapsto I - \rho. \tag{1.1-56}$$

On the Bloch sphere this corresponds to negating the Bloch vector: $\vec{v} \mapsto -\vec{v}$. It's easy to see that given three vectors on the Bloch sphere, for example ones aligned with the coordinate axes, negating all of them does not preserve their handedness, and the transformation can therefore not be described by a rotation. Consequently it is not unitary, and is hence not allowed [21].[6]

1.1.7 Two-qubit gates

The simplest way to manipulate a two-qubit system is to act locally on the components subsystems:

$$U_a \otimes V_b(|\Psi\rangle_a \otimes |\Phi\rangle_b) = U_a\,|\Psi\rangle_a \otimes V_b\,|\Phi\rangle_b\,, \tag{1.1-57}$$

however such operations cannot take a separable state to an entangled one, or vice versa. An example of a gate that can generate an entangled state is the controlled-NOT gate

$$U_{\text{CNOT}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{1.1-58}$$

here written in the basis $\{|i\rangle \otimes |j\rangle\}$:

$$|0\rangle \otimes |0\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad |0\rangle \otimes |1\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad |1\rangle \otimes |0\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad |1\rangle \otimes |1\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \tag{1.1-59}$$

This gate has the property of flipping the second qubit, called the target, conditional on the first qubit, known as the control, being in the state $|1\rangle$:

$$U_{\text{CNOT}}\,|0\rangle\,|0\rangle = |0\rangle\,|0\rangle \tag{1.1-60}$$

$$U_{\text{CNOT}}\,|1\rangle\,|0\rangle = |1\rangle\,|1\rangle\,, \tag{1.1-61}$$

[6] A second reason is that if one considers the qubits to be represented by a physical system, for example spin-$\frac{1}{2}$ particles, and measures the spin operator S^2, one finds: $\text{Tr}[S^2\rho_c] = 3\hbar^2$ which is unphysical since S^2 has eigenvalues $s(s+1)\hbar^2$.

By initializing the control qubit in a superposition of $|0\rangle$ and $|1\rangle$ an entangled state can be produced:

$$U_{\text{CNOT}} \frac{|0\rangle + |1\rangle}{\sqrt{2}} |1\rangle = \frac{|0\rangle |0\rangle + |1\rangle |1\rangle}{\sqrt{2}} = |\Phi^+\rangle . \tag{1.1-62}$$

The C-NOT gate cannot be written as a tensor product between two local operations, instead it must be expressed through an interaction Hamiltonian

$$U_{\text{CNOT}} = e^{i\pi H_{\text{CNOT}}}, \tag{1.1-63}$$

where

$$H_{\text{CNOT}} = \frac{1}{4}(I + Z \otimes X - I \otimes X - Z \otimes I). \tag{1.1-64}$$

The term $Z \otimes X$ describes the interaction between the two subsystems, and is what enables the two qubits to become correlated, while the other three terms correspond to local operations that are needed to generate the right entangling gate. Since they commute with the interaction term they can be factored out:

$$U_{\text{CNOT}} = e^{i\pi/4} \left[e^{-i\frac{\pi}{4}Z} \otimes e^{-i\frac{\pi}{4}X} \right] e^{i\frac{\pi}{4}Z \otimes X}. \tag{1.1-65}$$

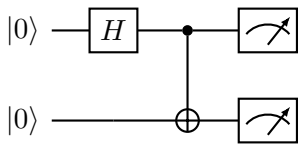
1.2 Methods in quantum information

Although the basic constructs of quantum information, such as encoding information in complex valued vectors, are relatively straightforward, they have nevertheless yielded a rich field of research with a number of surprising and striking results. The last few decades have also seen the development of a large set of both theoretical and experimental methods for studying and manipulating quantum information. A few of these results and methods will be summarised here.

1.2.1 The quantum circuit model

While the laws of quantum mechanics implicitly impose limitations on what kind of information processing tasks can be carried out in nature, they do not define a structure for how these tasks must be formulated or performed. Consequently, a host of different models for quantum information processing have been developed, often being inspired by or tailored to specific physical realisations of quantum states. Two widely used paradigms are Adiabatic Quantum Computation [22], in which the solution to a problem is gradually encoded in the ground state of a Hamiltonian, and One-Way Quantum Computation [23], also known as Measurement-Based Quantum Computation (MBQC), in which a computation is realised through measurements on a large entangled resource state [24]. The most widely used model, however, is the quantum circuit model, the basic constructs of which are the ones outlined in the previous section, namely qubits, qubit gates and measurements of qubit states [15]. Additionally, the quantum circuit model allows for operations that are conditioned on classical functions of measurement outcomes, and assumes the ability to efficiently initialize any computational basis state $\bigotimes_{k=1}^{n} |j_k\rangle$.

Computations in the circuit model are often drawn using circuit diagrams, in which horizontal lines represent qubits, and vertical lines denote gates that are conditioned on the state of a qubit (controlled operations). As an example, the preparation of a maximally entangled state using the C-NOT gate in (1.1-58), followed by measurements in the computational basis is drawn as



where H is the so-called Hadamard gate

$$H = \frac{X + Z}{\sqrt{2}}. \tag{1.2-1}$$

An important concept within the circuit model is the idea of universality. A given set of gates is said to be universal if they can be used to approximate, to an arbitrarily high precision, any n-qubit unitary. Quite remarkably, this can be achieved using as primitives only a single two-qubit gate, the C-NOT, as well as arbitrary single-qubit gates [25].

It is important to note, however, that this notion of universality does not imply that any limits within the circuit model are necessarily the same as the ones imposed by quantum mechanics. For instance, the circuit model always deals with finite-dimensional Hilbert spaces, whereas infinite-dimensional ones appear naturally in quantum mechanics. These can also be used for information processing, and the related field is known continuous-variable quantum information [26]. Even in finite dimensions there are examples of tasks not possible within the circuit model that nevertheless admit physical realisations, and one such task is the focus of Publication 3.

1.2.2 The no-cloning theorem

A key property of classical information is the fact that it can be copied an arbitrary number of times. This is important for being able to provide redundancy in information storage, but also leads to problems for example in the context of data rights management. Quantum information, on the other hand, has the property that it cannot be copied, or cloned, as was proven in 1982 [27, 28]. The result is a consequence of the linearity of quantum mechanics, and the proof is the following: consider a quantum state cloning device copying a state from system C to the system T:

$$\begin{aligned}(\alpha\,|0\rangle_C + \beta\,|1\rangle_C)\,|\Psi\rangle_T &= \alpha\,|0\rangle_C\,|\Psi\rangle_T + \beta\,|1\rangle_C\,|\Psi\rangle_T \\ &\mapsto \alpha\,|0\rangle_C\,|0\rangle_T + \beta\,|1\rangle_C\,|1_T\rangle\,.\end{aligned} \tag{1.2-2}$$

Now, suppose that $|0\rangle_C + \beta\,|1\rangle_C = |\Phi\rangle_C$. Then

$$\begin{aligned}|\Phi\rangle_C\,|\Psi\rangle_T &\mapsto |\Phi\rangle_C\,|\Phi\rangle_T \\ &= \alpha^2\,|0\rangle_C\,|0\rangle_T + \alpha\beta\,|0\rangle_C\,|1\rangle_T + \alpha\beta\,|1\rangle_C\,|0\rangle_T + \beta^2\,|1\rangle_C\,|1\rangle_T\,.\end{aligned} \tag{1.2-3}$$

(1.2-2) and (1.2-3) are in contradiction with each other even though they describe the same process. We can therefore conclude that copying an arbitrary quantum state is not possible. Note that this is only true for unknown quantum states, since if one knows how a given state was prepared, then performing an identical preparation on a second quantum state naturally produces another copy of the first state.

1.2.3 Quantum teleportation

Even though it is not possible to copy a quantum state, it is nevertheless possible to perfectly transfer a quantum state from one physical system to another [7, 29, 30]. This process has, somewhat misleadingly, been dubbed quantum teleportation, because the two systems between which the quantum state is transferred do not need to be in physical proximity to each other, and can in fact be separated by arbitrarily large distances. The idea behind quantum teleportation is to exploit the fact that entangled states have correlated measurement outcomes. If one wishes to teleport a state $|\psi\rangle_A$ one first entangles it with a second system B, which is in turn entangled with a third system C. A measurement on system A will now not reveal the initial state $|\psi\rangle_A$, but will instead yield a random outcome that is correlated with the system BC. Similarly, a measurement of system B will be random but correlated with AC. Consequently, the measurement outcome for system C will be correlated with the initial state in system A. This hand-waving argument can be formalised. Consider the initial quantum state

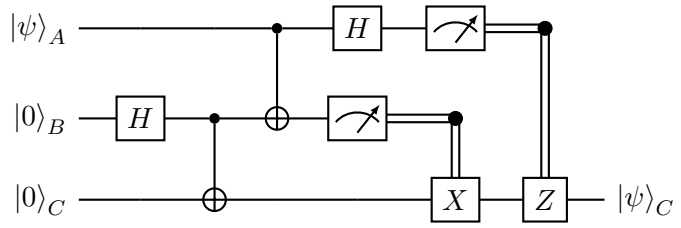
$$|\psi\rangle_A |\Phi^+\rangle_{BC} . \tag{1.2-4}$$

This state can equivalently be expressed in the Bell basis for the system AB:

$$\frac{1}{2}(|\Phi^+\rangle_{AB} |\psi\rangle_C + |\Phi^-\rangle_{AB} Z |\psi\rangle_C + |\Psi^+\rangle_{AB} X |\psi\rangle_C + |\Psi^-\rangle_{AB} XZ |\psi\rangle_C). \tag{1.2-5}$$

Measuring the state (1.2-4) in the AB Bell basis therefore projects the system C onto the state $|\psi\rangle$, modified by a local operation that depends on the measurement outcome. As long as the measurement outcome is known this local operation can be undone and the initial state recovered, however since each of the four measurement outcomes are equally likely the local description of the system C prior to receiving the measurement outcome is that of the maximally mixed state. This is what prevents quantum teleportation from enabling superluminal signalling.

The process of quantum teleportation of the state $|\Psi\rangle$ can be illustrated by the following circuit:

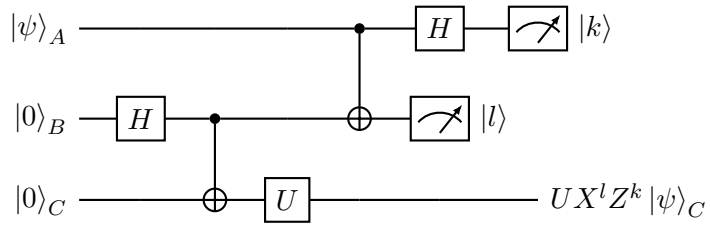


The first two gates prepare the maximally entangled state $|\Phi^+\rangle_{BC}$, while the two gates before the measurements have the effect of changing the measurement basis from the

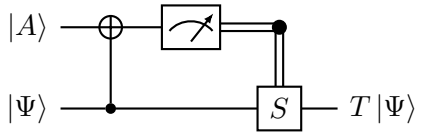
computational one to the Bell basis. This follows from the fact that the CNOT and Hadamard gate transform any Bell state into a computational basis state:

$$\begin{aligned} U^{A,B}_{\text{CNOT}}(H \otimes I)\,|\Phi^+\rangle_{AB} &= |0\rangle_A\,|0\rangle_B \\ U^{A,B}_{\text{CNOT}}(H \otimes I)\,|\Phi^-\rangle_{AB} &= |1\rangle_A\,|0\rangle_B \\ U^{A,B}_{\text{CNOT}}(H \otimes I)\,|\Psi^+\rangle_{AB} &= |0\rangle_A\,|1\rangle_B \\ U^{A,B}_{\text{CNOT}}(H \otimes I)\,|\Psi^-\rangle_{AB} &= |1\rangle_A\,|1\rangle_B\,. \end{aligned} \tag{1.2-6}$$

Finally, the thick wires represent conditional feed-forward operations that undo the local unitaries in (1.2-5). Quantum teleportation is a key building block for a multitude of different quantum protocols, including superdense coding [19] and entanglement swapping [31]. Interestingly, it is also possible to teleport a quantum state through a gate [32]:



In other words, the teleported state $|\psi\rangle$ has the gate U applied to it (up to a local unitary), even though the gate was applied on a different state before the teleportation. When the outcome of the Bell measurement is the $|\Phi^+\rangle$ state the exact gate U is teleported. For certain gates the other measurement outcomes can also be used by exchanging the order of the teleported gate and the local unitary, while modifying the latter.[7] An example of such a gate is the so-called T-gate, defined as an $R_z(\pi/4)$ rotation, which plays an important role in quantum error correction [33, 34, 35]. This gate can actually be teleported using a simpler circuit [36, 37]:



where $|A\rangle = TH\,|0\rangle$ and $S = R_z(\pi/2)$. The state $|A\rangle$ can be seen as a resource that contains the information about the gate T, allowing it to be applied without access to the gate itself. This is widely exploited in the context of quantum error correcting codes, which do not support the direct application of certain gates [38], but nevertheless allow for the approximate preparation of resource states encoding those gates [35], commonly referred to as magic states. More generally, the different forms of teleportation of both states and gates discussed above play an important role in proposals for photonic quantum computation [23, 24, 39].

[7] This is possible for gates in the set $\{U | UC_1U^\dagger \in C_2\}$, where C_1 is the Pauli group and C_2 is the Clifford group.

1.2.4 Fidelity

It is often useful to define a distance measure that quantifies how similar two quantum states $|\Phi\rangle$ and $|\Psi\rangle$ are, and a natural choice for such measure is the overlap of the respective state vectors

$$\mathcal{F} = |\langle\Psi|\Phi\rangle|^2. \tag{1.2-7}$$

This quantity is referred to as the fidelity of the state with respect to another state. In case one of the states is a mixed state represented by a density matrix the corresponding expression is

$$\mathcal{F} = \langle\Psi|\rho|\Psi\rangle, \tag{1.2-8}$$

and finally the fidelity between two density matrices ρ and σ is

$$\mathcal{F} = \mathrm{Tr}\left[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right]. \tag{1.2-9}$$

The fidelity is exactly 1 when the two states are equal, and 0 when the two states are orthogonal. For quantum gates a common measure of the fidelity is [40]

$$\mathcal{F} = \left\langle \mathrm{Tr}\left[\sqrt{\sqrt{\mathcal{E}(\rho)}\,U(\rho)\sqrt{\mathcal{E}(\rho)}}\,\right]\right\rangle_\rho. \tag{1.2-10}$$

where $\mathcal{E}$ and U are the experimental and ideal gates, respectively and the average is taken over all states in order to remove the state-dependent nature of the quantity. In practice it is typically averaged over a large number of different states randomly chosen using for example the Haar measure [41].

1.2.5 Quantum state tomography

An important task in experimental quantum information is the characterisation of quantum states. This can involve the estimation of certain properties of the state, for example whether or not the state is entangled [42], or direct estimation of the quantum state itself [43] as in the case of quantum state tomography. The destructive nature of measurements in quantum mechanics, combined with the impossibility of copying quantum information that follows from the no-cloning theorem, means that it is not possible to reconstruct a quantum state from a single copy, or even to infer certain properties of the state. In practice, measurements are instead performed on sets of independently and identically prepared states. In quantum state tomography one tries to find the density matrix of an unknown quantum state by performing a set of measurements whose outcomes completely characterise the state. In the case of a qubit state ρ_1 this amounts to estimating the expectation values

$$\mathrm{Tr}\big[X\rho_1\big], \quad \mathrm{Tr}\big[Y\rho_1\big], \quad \mathrm{Tr}\big[Z\rho_1\big] \tag{1.2-11}$$

and therefore three measurement settings are sufficient to characterise the state. For multi-qubit states the number of measurement settings needed scales with the dimension of the state, which grows exponentially in the number of qubits, and quantum state tomography is therefore an inefficient method for characterising large quatum states. This

has spurred investigation into more efficient methods of quantum state tomography [44], and state verification in general [45]. Quantum state tomography nevertheless remains a useful experimental tool for smaller quantum systems.

The expectation values in (1.2-11) can be estimated by performing measurements on a finite number of copies of the quantum state:

$$\mathrm{Tr}[\sigma_i \rho] \approx \frac{N_+ \lambda_+ + N_- \lambda_-}{2}. \tag{1.2-12}$$

A consequence of the finite number of measured copies is that this estimate will invariably be noisy, with the number of events $N_\pm$ of the positive and negative eigenvalues of the observable $\lambda_\pm$ following a Poisson distribution:

$$P(N_\pm) = \frac{m^{N_\pm} e^{-m}}{N_\pm!}, \tag{1.2-13}$$

where m is mean of the distribution. The noise in the expectation values means that an explicitly reconstructed density matrix such as

$$\rho_{exp} = \frac{1}{2}(I + X\langle X\rangle_{exp} + Y\langle Y\rangle_{exp} + Z\langle Z\rangle_{exp}) \tag{1.2-14}$$

will often be unphysical, with $\mathrm{Tr}[\rho_{exp}^2] > 1$ and $\rho_{exp} \not\geq 0$. To get a good estimate of the density matrix one therefore generally needs to fit the acquired data to a physically valid parameterisation of the density matrix [46]. One such parameterisation is the following:

$$\rho = \frac{T^\dagger T}{\mathrm{Tr}[T^\dagger T]}, \tag{1.2-15}$$

where T is a matrix of the same dimension as ρ. This parameterisation yields a matrix of with trace 1 that is clearly Hermitian, and also positive semidefinite since

$$\langle \Psi | T^\dagger T | \Psi \rangle = \langle \Phi | \Phi \rangle \geq 0. \tag{1.2-16}$$

Through the Cholesky decomposition any density matrix can be written on the form (1.2-15) with T being a lower triangular matrix with real and positive diagonal entries. In the case of a qubit state

$$T(\vec{t}) = \begin{bmatrix} t_1 & 0 \\ t_2 + it_4 & t_3 \end{bmatrix}, \qquad \vec{t} \in \mathbb{R}^4. \tag{1.2-17}$$

This is convenient in practice because it allows (1.2-15) to be inverted, and an initial guess for a density matrix can be mapped to the vector $\vec{t}$ and fed to a numerical optimizer.

A common way of fitting a density matrix to the acquired data is using the maximum likelihood technique [47, 48], which is a method for finding the density matrix that has the highest probability of having produced the experimental data. To this end, one needs to define a likelihood function expressing this probability. If the number of events $N_{i,s}$ of a certain outcome i for the observable σ_s is sufficiently high, then the Poisson distribution of these events can be approximated by a Gaussian distribution thanks to

the central limit theorem. The probability of getting N events with a certain outcome for a given observable therefore obeys the following proportionality:

$$P(N_{i,s}, \vec{t}) \propto \exp\left[\frac{-(\overline{N}_{i,s}(\vec{t}) - N_{i,s})^2}{2\sigma_{i,s}^2}\right], \tag{1.2-18}$$

where $\overline{N}_{i,s}(\vec{t})$ is the number of expected events for the density matrix $\rho(\vec{t})$, $N_{i,s}$ is the number of recorded events, and $\sigma_{i,s}$ is the standard deviation of the number of events, which due to the underlying Poisson statistics can be approximated by $\sqrt{\overline{N}_{i,s}(\vec{t})}$. The total probability that a given density matrix produced the measured data can then be expressed, up to a normalization constant A, as

$$P(\vec{t}) = \frac{1}{A}\prod_{i,j}\exp\left[\frac{-(\overline{N}_{i,s}(\vec{t}) - N_{i,s})^2}{2\overline{N}_{i,s}(\vec{t})}\right]. \tag{1.2-19}$$

This function needs be maximized numerically, however in practice it is more convenient to instead maximize the log of the function, or equivalently its negation

$$L(\vec{t}) = \sum_{i,s}\frac{(\overline{N}_{i,s}(\vec{t}) - N_{i,s})^2}{2\overline{N}_{i,s}(\vec{t})}, \tag{1.2-20}$$

where the normalization constant was disregarded since it does not affect the location of the minimum.

To estimate the effect of the counting statistics, as well as other experimental imperfections, on some quantity that is a function of the reconstructed state, it is common to perform a so-called Monte-Carlo simulation of the measurement. This involves defining a probability distribution for the various experimental parameters, repeatedly sampling from them, constructing a new density matrix using the maximum likelihood method described above at each step, and then calculating the relevant function, for example the fidelity $\mathcal{F}$. This is repeated until a convergence condition is met. One way to define this is through a confidence, for example one may want to keep running the Monte-Carlo simulation until one is 95 % confident that the uncertainty in mean value of the fidelity is below a certain threshold.

According to the central limit theorem, if n samples are drawn from a distribution with mean and variance μ and σ, respectively, and the mean of the samples is $\overline{\mathcal{F}}_n$, then the quantity

$$\mathcal{Z} = \frac{\overline{\mathcal{F}}_n - \mu}{\sigma/\sqrt{n}} \tag{1.2-21}$$

follows a standard normal distribution for sufficiently large n, i.e. a normal distribution with zero mean and unity standard deviation. The variable $\mathcal{Z}$ is called the standardised version of $\overline{\mathcal{F}}_n$ and can be used to express the confidence in the estimate of μ. For example

$$P(|\mathcal{Z}| < z) = p \tag{1.2-22}$$

means that the probability of the error in the estimate of the mean, $|\overline{\mathcal{F}}_n - \mu|$, being more than z in units of $\sigma/\sqrt{n}$ is p. The standard deviation σ can be estimated from the samples by

$$S_n = \sqrt{\frac{1}{n-1}\sum_k^n \left(F_k - \overline{\mathcal{F}}_n\right)^2}. \tag{1.2-23}$$

Substituting in the estimate of the standard deviation, the 95 % confidence condition is

$$P\left(\left|\frac{\overline{\mathcal{F}}_n - \mu}{S_n/\sqrt{n}}\right| < 1.96\right) = P\left(\left|\overline{\mathcal{F}}_n - \mu\right| < 1.96\frac{S_n}{\sqrt{n}}\right) = 0.95. \tag{1.2-24}$$

where the value 1.96 comes from the fact that for a standard normal distribution

$$\frac{1}{\sqrt{2\pi}}\int_{-1.96}^{1.96} e^{-x^2/2}dx \approx 0.95. \tag{1.2-25}$$

The condition for being 95 % confident that $\left|\overline{\mathcal{F}}_n - \mu\right| < \epsilon$ can now be expressed as

$$1.96\frac{S_n}{\sqrt{n}} < \epsilon. \tag{1.2-26}$$

1.2.6 Generalized measurements

The projective measurements described in Section 1.1.4 are not the only kind of measurements allowed in quantum mechanics. An example of a more general type of measurement is a positive operator-valued measurement, or POVM for short. A POVM is defined by a set of Hermitian positive semidefinite operators $\{E_m\}$, called the POVM elements, that sum to identity:

$$\sum_m E_m = I. \tag{1.2-27}$$

The number of POVM elements, and by extension possible measurement outcomes, can be arbitrarily high. Projective measurements are a special case of POVMs, for which the POVM elements are projectors, i.e. $E_m^2 = E_m$. In analogy with projective measurements, the probability of a given outcome m of a POVM on a state ρ is straightforwardly given by

$$P_m = \mathrm{Tr}[E_m\rho]. \tag{1.2-28}$$

Unlike in the case of projective measurements, the post measurement state of a POVM is not well defined, but instead depends on its decomposition into measurement operators M_m:

$$E_m = M_m^\dagger M_m, \tag{1.2-29}$$

which is not unique since UM_m, where U is any unitary, is also a measurement operator describing the same POVM element. Given a fixed choice of measurement operators though, the post measurement state can be expressed as

$$\frac{M_m\rho M_m^\dagger}{\mathrm{Tr}[M_m^\dagger M_m\rho]}. \tag{1.2-30}$$

In practice the operators M_m are a function of the physical realisation of the POVM. In the event that the POVM elements are rank-1 they are proportional to pure quantum states, and the measurement outcome m can be understood as revealing that the state before the measurement was not one orthogonal to E_m, since those are the only states which would never yield the outcome m.

1.2.7 Quantum channels

While closed quantum systems always follow a unitary time evolution the same is not true for open quantum systems, i.e. systems that interact with an external environment. Consider, as an example, the state

$$|\Psi\rangle_{\text{env},a} = |+\rangle_{\text{env}} |0\rangle_a \,, \tag{1.2-31}$$

evolving under the Hamiltonian

$$H = \frac{\pi}{4} Z \otimes X, \tag{1.2-32}$$

for a time $0 \leq t \leq 1$. The action of this Hamiltonian is to entangle the system a with the 'environment' system env. The time-evolved state is

$$e^{-it\frac{\pi}{4}Z\otimes X} |\Psi\rangle_{\text{env},a} = \frac{1}{\sqrt{2}} \big(e^{-it\frac{\pi}{4}X_a} |0\rangle_{\text{env}} |0\rangle_a + e^{it\frac{\pi}{4}X_a} |1\rangle_{\text{env}} |0\rangle_a \big). \tag{1.2-33}$$

Using the shorthand notation $\cos(\frac{\pi}{4}t) = \mathrm{c}_t$ and $\sin(\frac{\pi}{4}t) = \mathrm{s}_t$ (1.2-33) can be expanded as

$$|\Psi(t)\rangle_{\text{env},a} = \frac{1}{\sqrt{2}} \big(\mathrm{c}_t |0\rangle_{\text{env}} |0\rangle_a - i\mathrm{s}_t |0\rangle_{\text{env}} |1\rangle_a + \mathrm{c}_t |1\rangle_{\text{env}} |0\rangle_a + i\mathrm{s}_t |1\rangle_{\text{env}} |1\rangle_a \big). \tag{1.2-34}$$

The local description of the system a, also referred to as the reduced dynamics, is given by tracing out the environment:

$$\rho_a(t) = \mathrm{Tr}_{\text{env}} \big[|\Psi(t)\rangle\langle\Psi(t)|_{\text{env},a} \big] = c_t^2 |0\rangle\langle 0| + s_t^2 |1\rangle\langle 1| \,. \tag{1.2-35}$$

Substituting $\lambda = 2s_t^2$ this can be rewritten as

$$\rho_a = (1-\lambda) |0\rangle\langle 0| + \lambda \frac{I}{2}. \tag{1.2-36}$$

The time evolution therefore has the effect of interpolating the system b between its initial state and the maximally mixed state. This process is called depolarization, and the corresponding map is referred to as a depolarizing map.[8] The fact that these sorts of non-unitary transformations emerge naturally when a system interacts with an environment motivates the study of more general maps, referred to as quantum channels. The generic form of such a map $\mathcal{E}$ is

$$\mathcal{E}(\rho_a) = \mathrm{Tr}_{\text{env}} \Big[U(\rho_{\text{env}} \otimes \rho_a) U^\dagger \Big] \,. \tag{1.2-37}$$

[8] Technically the Hamiltonian used here does not generate a depolarizing map, because it only has this effect for the computational basis states $\{|0\rangle, |1\rangle\}$. However, by adding another qubit to the environment a depolarizing map for any input state can be achieved: $e^{-it\frac{\pi}{4}Z\otimes I\otimes X} e^{-it\frac{\pi}{4}I\otimes X\otimes Z} |+\rangle_b |0\rangle_c |\Psi\rangle_a$ has the effect of depolarizing the system a when taking the trace over b, c.

Here U is the unitary evolution of the joint system consisting of the environment and the local system a. Taking, without loss of generality, $\rho_{\text{env}} = |\phi\rangle\langle\phi|_{\text{env}}$ and evaluating the partial trace:

$$\begin{aligned}\mathscr{E}(\rho_a) &= \sum_j \langle j|_{\text{env}} U\big[|\phi\rangle\langle\phi|_{\text{env}} \otimes \rho_a\big] U^\dagger |j\rangle_{\text{env}} \\ &= \sum_j E_j \rho E_j^\dagger\end{aligned} \tag{1.2-38}$$

one obtains the so-called operator-sum representation of the map $\mathscr{E}$, which turns out to often be a more convenient representation. Here, the operators $E_j = \langle j|U|\phi\rangle$, referred to as the Kraus operators of map, act on the system a. Since the map $\mathscr{E}$ is trace preserving[9] it follows that that

$$\sum_j E_j E_j^\dagger = I. \tag{1.2-39}$$

The physical interpretation of this description is that the operation E_j is applied with probability $\text{Tr}[E_j \rho E_j]$ [15]. As an example of a quantum channel in this picture, the aforementioned depolarizing channel has an operator-sum representation given by[10]

$$E_1 = \sqrt{1 - \frac{3}{4}\lambda} I, \quad E_2 = \frac{\sqrt{\lambda}}{2} X, \quad E_3 = \frac{\sqrt{\lambda}}{2} Y, \quad E_4 = \frac{\sqrt{\lambda}}{2} Z. \tag{1.2-40}$$

To be a quantum channel the map $\mathscr{E}$ has to satisfy the following conditions:

1. (Trace preserving) $\quad \text{Tr}\big[\mathscr{E}(\rho)\big] = \text{Tr}\big[\rho\big]$ (1.2-41)
2. (Linearity) $\quad \mathscr{E}(\alpha\rho_1 + \beta\rho_2) = \alpha\, \mathscr{E}(\rho_1) + \beta\, \mathscr{E}(\rho_2)$ (1.2-42)
3. (Completely Positive) $\quad \rho_{ab} \geq 0 \implies (I_a \otimes \mathscr{E}_b)(\rho_{ab}) \geq 0$ (1.2-43)

and such transformations are often referred to as Completely Positive Trace Preserving (CPTP) maps. The three above conditions are equivalent to a map being able to be expressed as the reduced dynamics using (1.2-37), and this equivalence is a consequence of the Stinespring dilation theorem [49, 50]. The condition of complete positivity means that even if $\mathscr{E}$ only acts on a subsystem of a composite system the density operator of the subsystem remains positive semidefinite. There are several equivalent descriptions of CPTP maps, one being using the so-called χ matrix. In this description, instead of using the Kraus operators E_j one decomposes them in terms of a fixed set of operators $\{\widetilde{E}_m\}$ that form a basis for the set of operators in the space E_j acts on:

$$E_j = \sum_m e_{jm} \widetilde{E}_m. \tag{1.2-44}$$

Inserting this into (1.2-38) one obtains

$$\mathscr{E}(\rho) = \sum_{jmn} e_{jm} e_{jn}^* \widetilde{E}_m \rho \widetilde{E}_n^\dagger = \sum_{mn} \chi_{mn} \widetilde{E}_m \rho \widetilde{E}_n^\dagger \tag{1.2-45}$$

[9] The local dynamics have to conserve probability. One can also consider evolutions where probability is not conserved, for example if a particle disappears from the local system. In this case (1.2-39) can be relaxed to an inequality.

[10] The fact that this representation consists of four operators is why the channel cannot be realised through an interaction with a two-dimensional environment system.

where the coefficients

$$\chi_{mn} = \sum_j e_{jm} e_{jn}^* \tag{1.2-46}$$

are the matrix elements of the positive semidefinite and Hermitian matrix χ. This representation is particularly convenient when experimentally reconstructing a quantum channel, as will be discussed in Section 1.2.9. Yet another equivalent representation is given by the Choi matrix, defined as

$$\mathcal{C}_{\mathscr{E}} = (I \otimes \mathscr{E})(|\Omega\rangle\langle\Omega|), \tag{1.2-47}$$

where $|\Omega\rangle\langle\Omega|$ is the unnormalized maximally entangled state

$$|\Omega\rangle\langle\Omega| = \sum_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j| \,. \tag{1.2-48}$$

In other words, the Choi matrix can be thought of as the density matrix one obtains after applying the map $\mathscr{E}$ to one half of a maximally entangled state. The Choi matrix uniquely defines the map $\mathscr{E}$, and its action on a given state can be found as

$$\mathscr{E}(\rho) = \operatorname{Tr}_A\big[(\rho_A^T \otimes I_B)\mathcal{C}_{\mathscr{E}}\big]. \tag{1.2-49}$$

This can be seen by inserting (1.2-47) and (1.2-48) into (1.2-49):

$$\begin{aligned} \mathscr{E}(\rho) &= \sum_{ij} \operatorname{Tr}_A \left[\left(\rho_A^T \otimes I_B\right) \left(|i\rangle\langle j|_A \otimes \mathscr{E}(|i\rangle\langle j|_B) \right) \right] \\ &= \sum_{ijk} \langle k|_A \left(\rho_A^T \otimes I_B\right) \left(|i\rangle\langle j|_A \otimes \mathscr{E}(|i\rangle\langle j|_B) \right) |k\rangle_A \,. \end{aligned} \tag{1.2-50}$$

Since only the terms for which $j = k$ are non-zero, this reduces to

$$\begin{aligned} \sum_{ij} \langle j|\rho_A^T|i\rangle \mathscr{E}(|i\rangle\langle j|_B) &= \mathscr{E}\left(\sum_{ij} \rho_{ij} \, |i\rangle\langle j|_B \right) \\ &= \mathscr{E}(\rho). \end{aligned} \tag{1.2-51}$$

The Choi representation is a consequence of the Choi–Jamiołkowski isomorphism, which is a one-to-one map between quantum states and linear maps [51]. This correspondence has many practical uses, as properties of the map get transformed into properties of the state, that are often easier to study. For example, completely positive maps correspond to positive Choi operators, the completely depolarizing channel corresponds to the maximally mixed state, unitary transformations get mapped to pure states and so on [52].

One way to interpret (1.2-49) is as gate teleportation through an entangled state using an auxiliary system C, as described in Section 1.2.3. More concretely, consider a state ρ_A being teleported through the channel, together with an unnormalized maximally entangled state:

$$\rho_A \otimes |\Omega\rangle\langle\Omega|_{CB} \,, \tag{1.2-52}$$

and then apply the map $\mathscr{E}$ to the system B:

$$\begin{aligned}\rho_A \otimes [(I \otimes \mathscr{E})(|\Omega\rangle\langle\Omega|_{CB})] &= \rho_A \otimes \sum_{ij} |i\rangle\langle j|_C \otimes \mathscr{E}(|i\rangle\langle j|_B) \\ &= \rho_A \otimes \mathcal{C}_{\mathscr{E}}.\end{aligned} \tag{1.2-53}$$

To perform the gate teleportation one projects the system AC onto the state $|\Omega\rangle_{AC}$:

$$\begin{aligned}\langle\Omega|_{AC}\, \rho_A \otimes \mathcal{C}_{\mathscr{E}}\, |\Omega\rangle_{AC} &= \sum_{ij} \langle i|_A\, \rho_A\, |j\rangle_A \otimes \langle i|_C\, \mathcal{C}_{\mathscr{E}}\, |j\rangle_C \\ &= \sum_{ij} \rho_{ij}\, \langle i|_C\, \mathcal{C}_{\mathscr{E}}\, |j\rangle_C\,.\end{aligned} \tag{1.2-54}$$

Multiplying by $\langle j|j\rangle_C$ inside the sum yields

$$\begin{aligned}\sum_{ij} \rho_{ij} \langle j|j\rangle_C\, \langle i|_C\, \mathcal{C}_{\mathscr{E}}\, |j\rangle_C &= \sum_{ij} \langle j|_C\, \rho_{ij}\, |j\rangle\langle i|_C\, \mathcal{C}_{\mathscr{E}}\, |j\rangle_C \\ &= \sum_{j} \langle j|_C\, (\rho_C^T \otimes I)\mathcal{C}_{\mathscr{E}}\, |j\rangle_C \\ &= \mathrm{Tr}_C\big[(\rho_C^T \otimes I)\mathcal{C}_{\mathscr{E}}\big] \\ &= \mathscr{E}(\rho).\end{aligned} \tag{1.2-55}$$

This interpretation also highlights the fact that state is mapped from one Hilbert space $\mathcal{L}(\mathcal{H}_A)$ to a second one $\mathcal{L}(\mathcal{H}_B)$:

$$\mathscr{E} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B) \tag{1.2-56}$$

and the two systems A and B are often referred to as the input and output spaces of the map, respectively. Here $\mathcal{L}(\mathcal{H}_{A/B})$ denotes the space of linear operators on the spaces $\mathcal{H}_{A/B}$, meaning that

$$|\Psi\rangle_A \in \mathcal{H}_A, \quad |\Psi\rangle\langle\Psi|_A \in \mathcal{L}(\mathcal{H}_A). \tag{1.2-57}$$

The use of the unnormalized entangled states guarantees that the projection onto $|\Omega\rangle$ always succeeds, and the gate is teleported without any additional local operations.

The fact that the Choi matrix of CP maps can be thought of as an unnormalized density matrix motivates a notation using state vectors in the underlying state space. To distinguish them from regular state vectors they are written using a double ket. For example, the identity operator can be written [53]

$$|\mathbb{1}\rangle\!\rangle = \sum_j |j\rangle \otimes |j\rangle\,, \tag{1.2-58}$$

which is the same as the entangled state in (1.2-48). More generally, any map whose Choi matrix corresponds to a pure state, for example a unitary transformation

$$T(\rho) = U\rho U^\dagger, \tag{1.2-59}$$

can be written as a double ket:

$$|T\rangle\!\rangle = I \otimes T|\mathbb{1}\rangle\!\rangle. \tag{1.2-60}$$

If T is described by a matrix then (1.2-60) has the simple interpretation of $|T\rangle\!\rangle$ being a column vector consisting of the stacked columns of the corresponding matrix. For instance

$$U = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad |U\rangle\!\rangle = I \otimes U|\mathbb{1}\rangle\!\rangle = \begin{bmatrix} a \\ c \\ b \\ d \end{bmatrix}. \tag{1.2-61}$$

Naturally, the Choi matrix corresponding to the map T is given by the outer products of the state vectors:

$$\mathcal{C} = |T\rangle\!\rangle\langle\!\langle T|. \tag{1.2-62}$$

In the most general case, when $\mathcal{C}$ is not pure, one simply recovers (1.2-47) with $|\Omega\rangle\langle\Omega| = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|$.

1.2.8 Process matrices

The quantum channel formalism outlined in the previous section is sufficient to describe any quantum CPTP map. For characterizing such maps, however, it is useful to introduce additional structure, of which so-called process matrices are an example. Process matrices describe the causal relations within a quantum channel, and were introduced to study processes with an indefinite causal order [54]. Formally, a process matrix

$$W \in \mathcal{L}(\mathcal{H}_{A_{\text{in}}} \otimes \mathcal{H}_{A_{\text{out}}} \otimes \mathcal{H}_{B_{\text{in}}} \otimes \mathcal{H}_{B_{\text{out}}}) \tag{1.2-63}$$

is a positive semidefinite Hermitian matrix that obeys the normalization condition

$$\text{Tr}\big[W(P_A^T \otimes T_B^T)\big] = 1, \tag{1.2-64}$$

where

$$P_A \in \mathcal{L}(\mathcal{H}_{A_{\text{in}}} \otimes \mathcal{H}_{A_{\text{out}}}), \quad T_B \in \mathcal{L}(\mathcal{H}_{B_{\text{in}}} \otimes \mathcal{H}_{B_{\text{out}}}) \tag{1.2-65}$$

are any Choi matrices describing CPTP maps. A process matrix encodes the joint probability distribution for local operations performed in the laboratories A and B, and the measurement probabilities for these operations can be expressed through a generalization of the Born rule:

$$P(\mathcal{P}_A^i, \mathcal{T}_B^j) = \text{Tr}\big[W(P_A^i \otimes T_B^i)^T\big]. \tag{1.2-66}$$

Here, $\mathcal{P}_A^i$, $\mathcal{T}_B^j$ are completely positive, but not necessarily trace preserving maps, and P_A^i, T_B^i are their corresponding Choi matrices. A set of CP maps $\{\mathcal{P}_A^i\}$ with the property that the sum $\sum_i \mathcal{P}_A^i$ is a CPTP map is called an instrument. The concept of instruments encompasses both channels and measurements, and can be used to describe either or a combination of both. The maps $P_A^i = P_{A_{\text{in}}A_{\text{out}}}^i$ do not necessarily map states from Hibert spaces with the same dimension. For instance, an instrument for which $\dim(A_{\text{out}}) = 1$

represents a measurement after which the post-measurement state is discarded, and where P^i are the corresponding POVM elements.

The utility of process matrices lies in their ability to express the causal relation between the operations performed in local laboratories, which are in turn represented by instruments. The most simple case is one in which the operations in the laboratories A and B have a fixed causal order, such as A being able to signal to B but not vice versa. This is denoted by $A \prec B$, and any process $W^{A\prec B}$ can be expressed by an ordinary quantum circuit. A straightforward example of this is a process in which Alice sends a quantum state ρ to Bob through a shared channel C, which can be expressed in the following way [54]:

$$W^{A\prec B} = \rho_{A_{\text{in}}} \otimes \mathcal{C}_{A_{\text{out}}B_{\text{in}}} \otimes \mathbb{1}_{B_{\text{out}}}, \tag{1.2-67}$$

where $\mathcal{C}_{A_{\text{out}}B_{\text{in}}}$ is the Choi matrix of the channel. The necessary conditions for a matrix W to describe a causally ordered process $A \prec B$ are [53]

$$W \geq 0 \tag{1.2-68}$$

$$\text{Tr}[W] = d_{A_{\text{out}}} d_{B_{\text{out}}} \tag{1.2-69}$$

$$\frac{\mathbb{1}^{B_{\text{out}}}}{d_{B_{\text{out}}}} \otimes \text{Tr}_{B_{\text{out}}}[W] = W \tag{1.2-70}$$

$$\frac{\mathbb{1}^{A_{\text{out}}}}{d_{A_{\text{out}}}} \otimes \text{Tr}_{B_{\text{out}}B_{\text{in}}A_{\text{out}}}[W] = \text{Tr}_{B_{\text{out}}B_{\text{in}}}[W], \tag{1.2-71}$$

where $d_{\{\cdot\}}$ is the dimension of the indicated system. It's easy to see that the process matrix in (1.2-67) satisfies the first and third conditions, while the third follows from the fact that for a map $\mathcal{E} : \mathcal{L}(A) \to \mathcal{L}(B)$ with Choi operator $\mathcal{C}_{A,B}$ one has

$$\text{Tr}[\mathcal{C}_{\mathcal{A},\mathcal{B}}] = d_A. \tag{1.2-72}$$

Note that in (1.2-67) A_{out} is the input system of the Choi operator. The fourth and last condition is equivalent to

$$\mathbb{1}_{B_{\text{in}}} \otimes \mathbb{1}_{A_{\text{out}}} \otimes \text{Tr}[\mathcal{C}_{B_{\text{in}}A_{\text{out}}}] = d_{A_{\text{out}}} \text{Tr}_{B_{\text{in}}}[\mathcal{C}_{B_{\text{in}}A_{\text{out}}}], \tag{1.2-73}$$

which is always satisfied when $\mathcal{C}$ is a CPTP map. This is easy to see by writing the Choi matrix using the Kraus operators E_m:

$$\begin{aligned} \text{Tr}_{A_{\text{out}}}[\mathcal{C}_{B_{\text{in}}A_{\text{out}}}] &= \sum_{ijk} \langle k|_{A_{\text{out}}} |i\rangle\langle j|_{A_{\text{out}}} |k\rangle_{A_{\text{out}}} \otimes \sum_m E_m |i\rangle\langle j|_{A_{\text{in}}} E_m^\dagger \\ &= \sum_{km} E_m |k\rangle\langle k|_{A_{\text{in}}} E_m^\dagger \\ &= \sum_m E_m E_m^\dagger = \mathbb{1}_{A_{\text{in}}}. \end{aligned} \tag{1.2-74}$$

Process matrices can also be used to describe processes which quantum circuits cannot, and which do not have a definite causal structure. The most well known example of this is the quantum switch [55], whose corresponding process matrix cannot be written as

a mixture of causally ordered process matrices, and is said to have an indefinite causal structure. While the quantum switch has seen experimental realisations [11, 56] it is not the most general causal structure that can be described in this formalism. There also exists valid process matrices that violate so-called causal inequalities, analogous to how entangled states can exhibit non-locality and violate Bell inequalities [54, 57]. At the moment, however, it is not fully known whether such processes are allowed within quantum mechanics [58, 59].

1.2.9 Quantum process tomography

In analogy with quantum state tomography discussed in Section 1.2.5, quantum process tomography is a method for reconstructing quantum channels. The Choi representation of quantum channels suggests that this can be done by sending one half of an entangled state through a channel, and then performing quantum state tomography on the full state [60]. This is indeed possible and has been demonstrated for small systems [61], however the requirement of a maximally entangled state makes the technique experimentally challenging. A more practical method is to instead send several different states through the channel, and to perform quantum state tomography on the resulting output states. The set of input states must be chosen such that the map is fully determined by the input-output relations for those states, and such sets are called tomographically complete.

If the channel can be assumed to be unitary, which is often the case experimentally, the process tomography can be simplified, since a unitary transformation only has d^2 free parameters, as opposed to the $d^4 - d^2$ free parameters of the χ matrix. Since the global phase of a unitary transformation is not directly measurable, process tomography on a unitary in practice only involves determining d^2-1 parameters. This can be done by first sending $d-1$ computational basis states through the channel and performing quantum state tomography on the output states, thereby yielding the columns of the unitary matrix up to a relative phase. The phases between the columns can be determined by sending an additional $d-1$ states from a superposition basis.

For two-dimensional unitary transformations there is a straightforward geometric approach for reconstructing the corresponding matrix. Consider the initial and final state vectors $|H\rangle$, $|+\rangle$ and $U|H\rangle = |H'\rangle$, $U|+\rangle = |+'\rangle$ on the Bloch sphere. Let the corresponding Bloch vectors be $\vec{v}_H$, $\vec{v}_{H'}$, $\vec{v}_+$ and $\vec{v}_{+'}$, then construct the two difference vectors

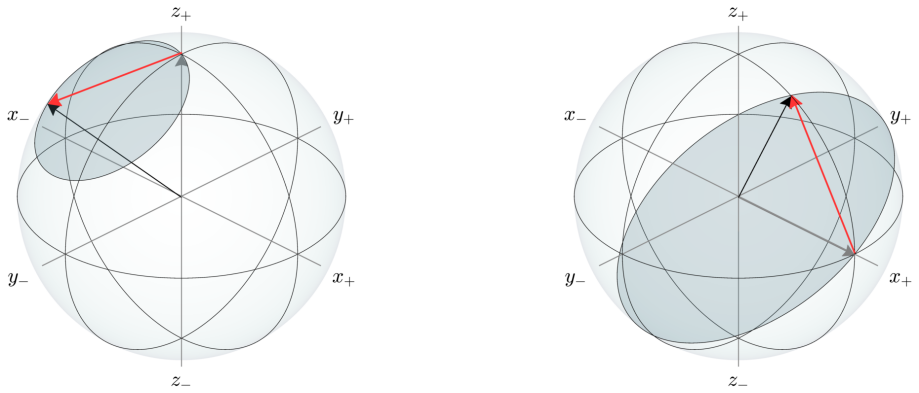
$$\Delta\vec{v}_H = \vec{v}_H - \vec{v}_{H'}, \qquad \Delta\vec{v}_+ = \vec{v}_+ - \vec{v}_{+'}. \tag{1.2-75}$$

The rotation axis $\vec{n}$ of the unitary is then simply the normal of the plane spanned by $\Delta\vec{v}_H$ and $\Delta\vec{v}_+$:

$$\vec{n} = \frac{\Delta\vec{v}_H \times \Delta\vec{v}_+}{\|\Delta\vec{v}_H \times \Delta\vec{v}_+\|} \tag{1.2-76}$$

This is illustrated in Figs. 1.3 and 1.4. The rotation angle θ can be determined from the input-output relations of either state by first projecting the initial and final Bloch vectors into the plane of rotation:

$$\vec{p}_H = \vec{v}_H - (\vec{v}_H \cdot \vec{n})\vec{n} \qquad \vec{p}_{H'} = \vec{v}_{H'} - (\vec{v}_{H'} \cdot \vec{n})\vec{n}, \tag{1.2-77}$$



(a) $U\,|H\rangle$ (b) $U\,|+\rangle$

Figure 1.3: **Reconstruction of a unitary.** The rotation axis of a unitary can be experimentally reconstructed by first constructing the difference between the initial and final probe states under the unitary transformation. These difference vectors are shown in red for initial states $|H\rangle$ and $|+\rangle$ in (a) and (b) respectively. The difference vectors lie in two parallel planes, indicated by the shaded regions, whose normal defines the rotation axis of the unitary.

and then finding the angle between these two vectors (see Fig. 1.4):

$$\theta_H = \cos^{-1}\left(\frac{\vec{p}_H \cdot \vec{p}_{H'}}{\|\vec{p}_H\|\,\|\vec{p}_{H'}\|}\right)\frac{\vec{n}\cdot(\vec{p}_H \times \vec{p}_{H'})}{\|\vec{n}\cdot(\vec{p}_H \times \vec{p}_{H'})\|}. \tag{1.2-78}$$

The second term in the expression above corrects the sign of the angle θ to match the definition of $\vec{n}$. By calculating the corresponding angle for the second input state and interpolating between them to maximize the fidelity for both states one typically obtains a sufficiently good matrix that any further numerical optimization is redundant. In the case where either difference vector $\Delta\vec{v}_H$ or $\Delta\vec{v}_+$ is small it can suffer from a large uncertainty, however this does not affect the direction of the normal $\vec{n}$, since the small difference vector will be approximately confined to a plane orthogonal to the second difference vector. The corresponding rotation angle θ will exhibit a similar uncertainty, but a good estimate of this angle can always be found from the complementary state. In the event that both difference vectors are small the direction of the normal will be very unstable, but this is expected since this condition implies that the transformation is close to the identity map, which does not have a well defined rotation axis.

In the case of a more general channel acting on a d-dimensional Hilbert space one needs to probe it with d^2 states whose corresponding density matrices form a basis for the space of $d \times d$ matrices. Perhaps surprisingly, such a basis does exist and an example for qubit states is

$$|H\rangle\langle H|\,, \quad |V\rangle\langle V|\,, \quad |+\rangle\langle +|\,, \quad |L\rangle\langle L|\,, \tag{1.2-79}$$

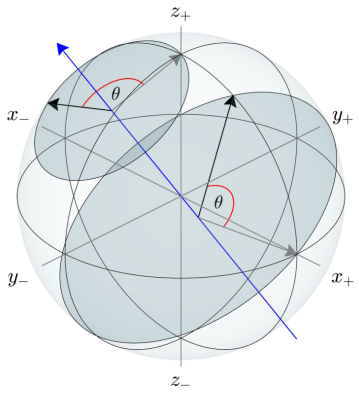


Figure 1.4: **Rotation angle.** The eigenphase $\theta : U|\lambda_\pm\rangle = e^{\mp i\theta/2}|\lambda_\pm\rangle$ of the unitary U can be found as the angle between the initial and final Bloch vectors projected into the plane of rotation. The blue arrow indicates the rotation axis of the transformation, which is normal to the two shaded planes. Experimentally the angle θ cannot be distinguished from $2\pi - \theta$ by performing tomography on the unitary alone.

with the off-diagonal elements in the computational basis being given by

$$\begin{aligned} |1\rangle\langle 0| &= |+\rangle\langle +| - i\,|L\rangle\langle L| + \frac{1-i}{2}(|H\rangle\langle H| + |V\rangle\langle V|), \\ |0\rangle\langle 1| &= |+\rangle\langle +| + i\,|L\rangle\langle L| - \frac{1+i}{2}(|H\rangle\langle H| + |V\rangle\langle V|). \end{aligned} \tag{1.2-80}$$

Since quantum channels are linear, the input-output relations for the basis states completely determine the map. More concretely, if $\{\sigma_j\}$ is the aforementioned matrix basis with which the channel is probed, then any densiy matrix can be written

$$\rho = \sum_j p_j \sigma_j, \tag{1.2-81}$$

and the action of the map on the basis states is

$$\mathscr{E}(\sigma_j) = \sum_k \lambda_{jk}\sigma_k. \tag{1.2-82}$$

Consequently, the action on a generic state ρ can be expanded as

$$\mathscr{E}(\rho) = \mathscr{E}\left(\sum_j p_j \sigma_j\right) = \sum_{jk} \lambda_{jk}\sigma_k. \tag{1.2-83}$$

The coefficients λ_{jk} can be found by projecting $|\mathscr{E}(\sigma_k)\rangle\!\rangle$ onto the vectorized basis. In the general case this vectorized basis of the matrices can be non-orthogonal, and the decomposition in such a basis can be computed as

$$\lambda_k = \begin{bmatrix} \lambda_{1k} \\ \vdots \\ \lambda_{dk} \end{bmatrix} = (B^T B)^{-1} B^T |\mathscr{E}(\sigma_k)\rangle\!\rangle, \tag{1.2-84}$$

where the matrix B is simply the matrix containing the vectorized basis elements:

$$B = \begin{bmatrix} |\sigma_1\rangle\rangle & \dots & |\sigma_d\rangle\rangle \end{bmatrix}. \tag{1.2-85}$$

The form of the χ-matrix (1.2-46) depends on the choice of operators $\widetilde{E}_j$ used to represent it. After these have been fixed, one can also decompose the action of them on the probe states $\{\sigma_j\}$:

$$\widetilde{E}_m \sigma_j \widetilde{E}_n^\dagger = \sum_k \beta_{jk}^{mn} \sigma_k. \tag{1.2-86}$$

The coefficients β_{jk}^{mn} can be obtained in analogy with (1.2-84) and can be represented as a matrix

$$\beta = \begin{bmatrix} \beta_{11}^{11} & \dots & \beta_{11}^{1d} & \beta_{11}^{21} & \dots & \beta_{11}^{2d} & \dots & \beta_{11}^{d1} & \dots & \beta_{11}^{dd} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ \beta_{1d}^{11} & \dots & \beta_{1d}^{1d} & \beta_{1d}^{21} & \dots & \beta_{1d}^{2d} & \dots & \beta_{1d}^{d1} & \dots & \beta_{1d}^{dd} \\ \beta_{21}^{11} & \dots & \beta_{21}^{1d} & \beta_{21}^{21} & \dots & \beta_{21}^{2d} & \dots & \beta_{21}^{d1} & \dots & \beta_{21}^{dd} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ \beta_{2d}^{11} & \dots & \beta_{2d}^{1d} & \beta_{2d}^{21} & \dots & \beta_{2d}^{2d} & \dots & \beta_{2d}^{d1} & \dots & \beta_{2d}^{dd} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ \beta_{d1}^{11} & \dots & \beta_{d1}^{1d} & \beta_{d1}^{21} & \dots & \beta_{d1}^{2d} & \dots & \beta_{d1}^{d1} & \dots & \beta_{d1}^{dd} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ \beta_{dd}^{11} & \dots & \beta_{dd}^{1d} & \beta_{dd}^{21} & \dots & \beta_{dd}^{2d} & \dots & \beta_{dd}^{d1} & \dots & \beta_{dd}^{dd} \end{bmatrix}. \tag{1.2-87}$$

By combining the decompositions (1.2-83) and (1.2-86) with the definition of the map (1.2-45) one gets

$$\mathcal{E}(\sigma_k) = \sum_{mn} \chi_{mn} \widetilde{E}_m \sigma_k \widetilde{E}_n^\dagger = \sum_{lmn} \chi_{mn} \beta_{kl}^{mn} \sigma_l = \sum_l \lambda_{kl} \sigma_l, \tag{1.2-88}$$

and since the matrices $\{\sigma_l\}$ are linearly independent by virtue of being a basis the expression above has to hold separately for each l:

$$\sum_{mn} \chi_{mn} \beta_{kl}^{mn} = \lambda_{kl} \iff \beta |\chi\rangle\rangle = |\lambda\rangle\rangle. \tag{1.2-89}$$

The matrix elements of χ can then be found by inverting the matrix β:

$$|\chi\rangle\rangle = \beta^{-1} |\lambda\rangle\rangle. \tag{1.2-90}$$

The explicitly constructed matrix χ is not guaranteed to be Hermitian or positive semidefinite. To address this, the Hermitian part of the matrix can be found by taking

$$\chi_H = \frac{\chi + \chi^\dagger}{2}. \tag{1.2-91}$$

Diagonalizing χ_H and setting all the negative eigenvalues to zero:

$$D = U^\dagger \chi_H U, \quad \chi_H^+ = U D^+ U^\dagger, \tag{1.2-92}$$

then ensures a Hermitian and positive semidefinite matrix χ_H^+, which can be used as the initial guess for a maximum likelihood optimization in analogy with quantum state tomography.

1.3 Quantum communication

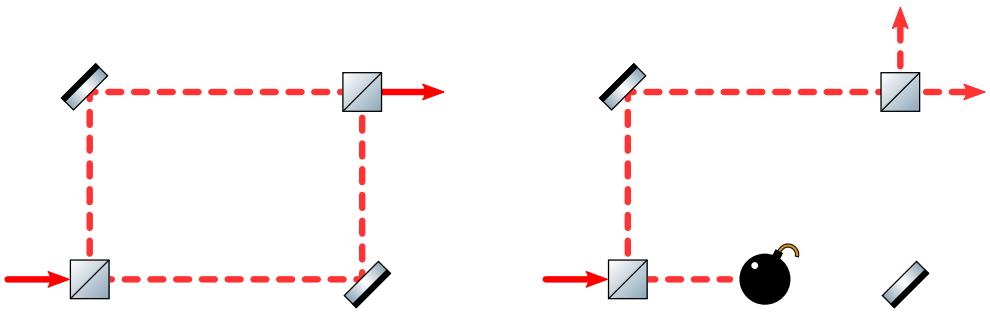
Due to the impossibility of fully characterizing a single copy of a quantum state discussed in Section 1.2.2, two parties wanting to exchange quantum information with each other cannot use a classical transmission channel, but instead need to exchange quantum states. This naturally leads to the idea of communicating using quantum bits instead of classical bits, and this is the field of quantum communication. Allowing two parties to exchange qubit states opens up a host of surprising applications, not only in the transmission of unknown quantum states through for example the previously described quantum teleportation [7, 29, 30], but in the transmission of classical information as well. An example of the latter is superdense coding [19, 62], which allows two parties to exchange two bits of classical information by only exchanging a single qubit.

1.3.1 Quantum key distribution

The most important communication task enabled by using quantum states is quantum key distribution (QKD), which is a family of methods for two parties to establish a secure cryptographic key without an eavesdropper being able to gain information about the key. It was first described in a 1984 paper by C.H. Bennett and G. Brassard and the associated protocol is known as BB84 [5]. The protocol works by one party, commonly referred to as Alice, generating a random bit, which is then randomly encoded either in the X or Z basis of a quantum state. She then transmits this state to Bob, who randomly picks one of the two bases to measure it in. If Alice and Bob pick the same basis the bit Alice encoded will be revealed to Bob with certainty, and if they pick different bases Bob's measurement result will be uncorrelated with Alice's encoded bit. After many states have been sent and measured Alice and Bob openly share their bases for each round, and discard all events where they picked different bases. Alice and Bob now share a random string that can be used as a basis for secure communication, for example using a one-time pad encoding [63]. Any eavesdropper attempting to measure the particles before they reach Bob will be unable to re-prepare the same state again after the measurement with unity probability, and will therefore induce imperfect correlations between Bob's measurement results and Alice's encoding in the rounds where they picked the same measurement basis. By revealing part of the shared random string they can find errors stemming from this imperfect correlation and thereby detect the presence of an eavesdropper.

1.3.2 Counterfactual communication

While many quantum communication schemes are studied for their potential applications in real-world scenarios, there are also ones that serve to push our understanding of quantum mechanics; counterfactual communication (CFC) is an example of the latter. At a fundamental level, all forms of communication involve the encoding of information in some degrees of freedom of a collection of particles, which are then transmitted from one party to another. Classically, a particle carrying a single bit of information can only be used to transmit that same amount of information, whereas in quantum mechanics superdense coding allows a single bit particle to transmit two bits of information. In contrast, counterfactual communication is a form of communication in which



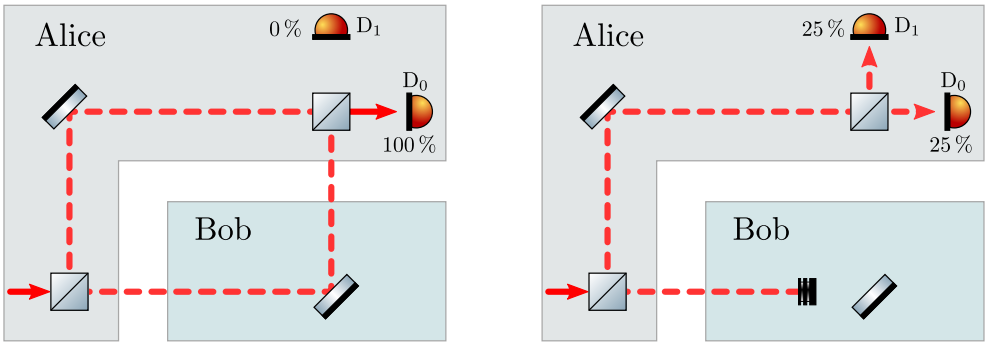
(a) Mach-Zehnder interferometer (b) Elizur-Vaidman bomb tester

Figure 1.5: **Interaction-free measurement** (a) A photon entering a balanced Mach-Zehnder interferometer from the lower left port will exit in the upper right port with probability 1. (b) If an object blocks one of the paths, a photon entering the interferometer has a 25 % chance to exit the interferometer in either port. Detection of a single photon in the upper output port thereby signals the presence of an object in one of the arms, without the photon having ever interacted with the object in question. This phenomenon has been termed interaction-free measurement.

no particles are exchanged between the two parties communicating. The basis for counterfacutal communication are so-called interaction-free measurements, first illustrated by the Elizur-Vaidman bomb tester thought experiment [64], in which they imagined a bomb which explodes if touched by a single photon. They posed the question if the presence of such a bomb could be detected using light, without detonating the bomb in the process, and went on to describe a method of accomplishing just that.

Consider the Mach-Zehnder interferometer pictured in Fig. 1.5a; a photon entering the interferometer from the bottom left, always exits in the upper right mode due to the destructive interference of the two photon paths leading to the upper port. If the bomb is placed in one arm of the interferometer the situation changes. As illustrated in Fig. 1.5b, there is a 50 % chance that the photon will be transmitted through the first beamsplitter and detonate the bomb, however there is an equal probability that the photon is reflected upwards and when it hits the second beamsplitter it now has a 50 % probability of exiting the interferometer in the upper port, since the wavefunction in the lower part of the interferometer is blocked by the bomb and the photon does not interfere with itself. If a detector placed in the upper output port of the interferometer detects a photon that implies that the bomb was present in the interferometer, and moreover that the photon traversed the upper path and therefore did not interact with the bomb. Consequently, 25 % of the time this device successfully detects the presence of the bomb without detonating it. This process of indirectly detecting an object has been dubbed interaction-free measurement [65].

From this setup a simple communication protocol can be realised, illustrated in Fig. 1.6. Both beamsplitters reside within Alice's laboratory, but one arm of the interferometer passes through Bob's laboratory. If Bob does nothing the detector D_0 will always click, but if he blocks his path there is a 25 % probability of the detector D_1 clicking. Bob can therefore deterministically transmit a logical 0 by leaving the path unblocked, and probabilistically transmit a logical 1 by blocking his path. While this method of communication is very noisy, with a 50 % error probability for the logical



(a) Transmission of a 0-bit

(b) Transmission of a 1-bit

Figure 1.6: **Interaction-free communication.** Using a Mach-Zehnder interferometer a basic communication protocol can be realised, in which Bob transmits a classical bit to Alice without Bob ever sending a particle to Alice. The two shaded areas denote the laboratories of Alice and Bob, and the remaining space constitutes a transmission channel. Depending on whether not Bob blocks the photon path through his laboratory he can, probabilistically, transmit a 0- or a 1-bit to Alice. (a) If Bob leaves his mirror unblocked, the photon sent by Alice will always cause a click in the detector D_0. (b) Whenever Bob blocks the photon path through his lab there is a 25 % probability that the photon will leave the second beamsplitter in the upper port, and make detector D_1 click.

1, it is nevertheless remarkable that Bob is able to transmit information to Alice without directly sending any particles to her. However, since the photon does enter Bob's laboratory and then propagates back to Alice, it can be argued that the protocol is not truly counterfactual even though Bob never interacts with the particle. Salih *et al.* showed that by considering a more complex interferometer consisting of several nested Mach-Zehnder interferometers, illustrated in Fig. 1.7, it becomes possible to perform counterfactual communication with an arbitrarily high probability of success [66]. Some authors have however questioned the counterfactuality of the protocol, sparking a long-running debate over the issue [67, 68, 69, 70, 71, 72]. The disagreement stems from the presence of a so-called weak trace in Bob's laboratory, and the difficulty in elucidating exactly what the past of a quantum particle is [73].

1.3.3 Weak measurements

The weak trace of a particle is related to the concept of weak measurements, which are a kind of measurement that only induce a minimal perturbation on the system being measured [74]. This is in contrast to projective measurements, sometimes called strong measurements in this context, that maximally disturb the system being measured, collapsing it into an eigenstate of the measurement operator. A weak measurement can be performed on a system by letting it weakly interact with a probe, or pointer system, for a limited amount of time, and then measuring the probe. Consider a probe system M with a normal distribution in the position basis [75]:

$$\phi(x) = \frac{1}{\sqrt[4]{2\pi\sigma^2}} e^{-\frac{x^2}{4\sigma^2}}. \tag{1.3-1}$$

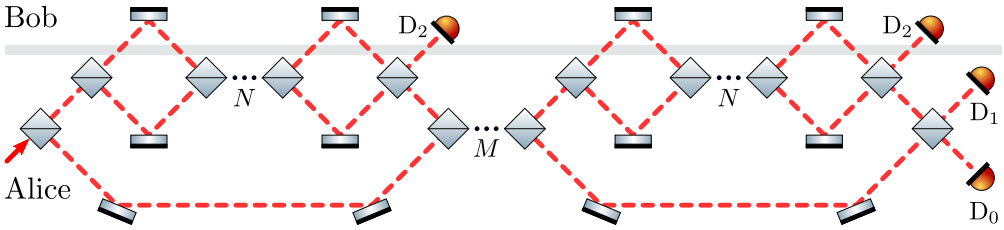


Figure 1.7: **Salih CFC protocol.** In Ref. [66] Salih et al. proposed a protocol for direct counterfactual communication that consists of several nested Mach-Zehnder interferometers. The grey shaded area in the figure represent the transmission channel that separates the laboratories of Alice and Bob. The idea of the protocol is the following: the inner set of N chained MZIs are balanced such that any photon entering them will always generate a click in one of the M detectors labelled D_2, whenever Bob does not block his mirrors. That means that when the mirrors are unblocked, the inner chained MZIs effectively work to block the upper path through the outer chain of M MZIs. However, given that the reflectivity of the outer beamsplitters is sufficiently high the photon will, in this scenario, travel to detector D_0 with near certainty. On the other hand, when Bob blocks his mirrors the photon will be confined to the lower part of the inner MZI chain, which also has high reflectivity beam-splitters. By choosing the reflectivities of both sets of beamsplitters carefully, one can ensure that the outer path and lower inner path interfere destructively in the D_0 port, and constructively in the D_1 port, ensuring that the detector D_1 clicks with near certainty when Bob blocks his mirrors.

The state vector of the probe in the position basis is then given by an intergral over the wavefunction:

$$|\phi(x)\rangle_M = \int_{-\infty}^{\infty} \phi(x)\,|x\rangle_M\,dx. \tag{1.3-2}$$

Let, as an example, the state to be measured be a spin-$\frac{1}{2}$ system:

$$|\Psi\rangle_S = \alpha\,|\uparrow\rangle_S + \beta\,|\downarrow\rangle_S\,, \tag{1.3-3}$$

written in the eigenbasis of $\frac{1}{2}Z$, and where the eigenvalues are $\lambda_\uparrow = \frac{1}{2}$ and $\lambda_\downarrow = -\frac{1}{2}$. To perform the weak measurement the two systems are coupled through the interaction Hamiltonian

$$H(t) = g(t)\frac{1}{2}Z_S \otimes \mathcal{P}_M, \tag{1.3-4}$$

where $g(t)$ is an envelope function for the interaction that satisfies

$$\int_0^T g(t)dt = 1, \tag{1.3-5}$$

and

$$\mathcal{P} = -i\hbar\frac{\partial}{\partial x} \tag{1.3-6}$$

is the generator of translations for the probe system [18, 76]:

$$\begin{aligned}\exp\Big[-a\frac{i}{\hbar}\mathcal{P}\Big]\,|\phi(x)\rangle &= \sum_{n=0}\frac{1}{n!}\Big(-a\frac{\partial}{\partial x}\Big)^n\,|\phi(x)\rangle \\ &= |\phi(x-a)\rangle\,.\end{aligned} \tag{1.3-7}$$

Since the Hamiltonian (1.3-4) is time dependent, the corresponding time-evolution operator from $t = 0$ to $t = T$ is given by [18]

$$\begin{aligned}\mathcal{U}(0,T) &= \exp\left[-\frac{i}{\hbar}\int_0^T H(t)dt\right] \\ &= \exp\left[-i\int_0^T g(t)\frac{1}{2}Z_S \otimes \frac{1}{\hbar}\mathcal{P}_M dt\right] \\ &= \exp\left[-i\frac{1}{2}Z_S \otimes \frac{1}{\hbar}\mathcal{P}_M\right].\end{aligned} \tag{1.3-8}$$

Then, using the notation

$$\frac{1}{2}Z\,|\lambda\rangle = \lambda\,|\lambda\rangle\,, \tag{1.3-9}$$

the effect of the time evolution under the Hamiltonian in (1.3-4) for a time T on a state $|\lambda\rangle_S \otimes |\phi(x)\rangle_M$ is that of translating the position of the probe by an amount λ:

$$\mathcal{U}(0,T)\,|\lambda\rangle_S \otimes |\phi(x)\rangle_M = |\lambda\rangle_S \otimes |\phi(x-\lambda)\rangle_M\,. \tag{1.3-10}$$

This can be seen by, as in (1.3-7), Taylor expanding the time evolution:

$$\begin{aligned}\mathcal{U}(0,T)\,|\lambda\rangle_S \otimes |\phi(x)\rangle_M &= \exp\Big[-i\frac{1}{2}Z_S \otimes \frac{1}{\hbar}\mathcal{P}_M\Big]\,|\lambda\rangle_S \otimes |\phi(x)\rangle_M \\ &= |\lambda\rangle_S \otimes |\phi(x)\rangle_M - \lambda\,|\lambda\rangle_S \otimes \frac{\partial}{\partial x}\,|\phi(x)\rangle_M \\ &+ \lambda^2\,|\lambda\rangle_S \otimes \frac{1}{2}\Big(-\frac{\partial}{\partial x}\Big)^2\,|\phi(x)\rangle_M + \cdots \\ &= |\lambda\rangle_S \otimes \sum_{n=0}\frac{1}{n!}\Big(-\lambda\frac{\partial}{\partial x}\Big)^n\,|\phi(x)\rangle_M \\ &= |\lambda\rangle_S \otimes |\phi(x-\lambda)\rangle_M\end{aligned} \tag{1.3-11}$$

In the literature, these types of interactions are frequently presented in the Heisenberg picture, in which one can reach the same conclusion by starting from the fact that the change in the time-dependent position operator $\mathcal{X}$ is given by

$$\mathcal{X}(T) - \mathcal{X}(0) = \int_0^T \frac{d}{dt}\mathcal{X}(t)dt. \tag{1.3-12}$$

Then, using Heisenberg's equation of motion

$$\frac{d}{dt}\mathcal{X} = \frac{i}{\hbar}[H, \mathcal{X}(t)], \tag{1.3-13}$$

and the fact that $\mathcal{X}$ and $\mathcal{P}$ are conjugate operators that satisfy

$$[\mathcal{X}, \mathcal{P}] = i\hbar \tag{1.3-14}$$

(1.3-12) can be expanded as

$$\begin{aligned}\int_0^T \frac{d}{dt}\mathcal{X}(t)dt &= \int_0^T \frac{i}{\hbar}[g(t)\frac{1}{2}Z\otimes\mathcal{P}, I\otimes\mathcal{X}(t)]dt \\ &= \int_0^T \frac{i}{\hbar}g(t)\frac{1}{2}Z\otimes[\mathcal{P},\mathcal{X}(t)]dt \\ &= \int_0^T \frac{1}{2}g(t)Zdt = \frac{1}{2}Z.\end{aligned} \tag{1.3-15}$$

Taking the expectation value of (1.3-12) for an initial state $|\lambda\rangle\otimes|\phi(x)\rangle$ it then follows that the change in the position operator is once again a translation by λ:

$$\begin{aligned}\langle\lambda,\phi(x)|I\otimes[\mathcal{X}(T)-\mathcal{X}(0)]|\lambda,\phi(x)\rangle &= \langle\lambda,\phi(x)|\frac{1}{2}Z\otimes I|\lambda,\phi(x)\rangle \\ &= \langle\lambda,\phi(x)|I\otimes\lambda I|\lambda,\phi(x)\rangle,\end{aligned} \tag{1.3-16}$$

and we have

$$\mathcal{X}(T)-\mathcal{X}(0)=\lambda. \tag{1.3-17}$$

Returning to the superposition state in (1.3-3) and including the pointer system:

$$|\Psi\rangle_S\otimes|\phi(x)\rangle_M = \left(\alpha|\uparrow\rangle_S+\beta|\downarrow\rangle_S\right)\otimes|\phi(x)\rangle_M, \tag{1.3-18}$$

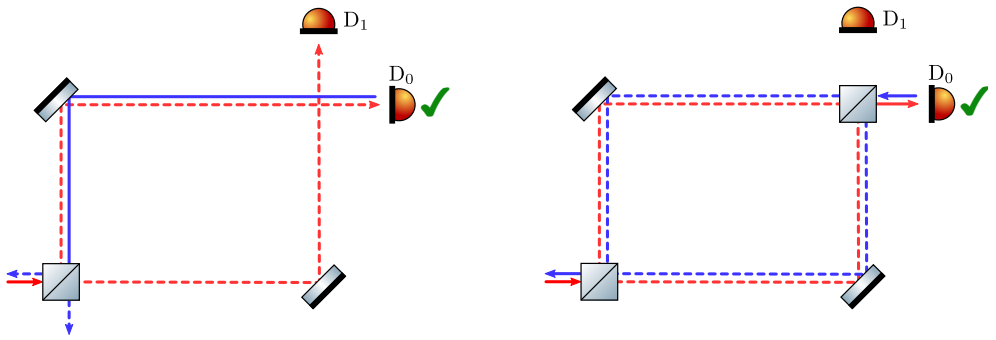
we now see that the action of the Hamiltonian on this state is to entangle the position of the pointer with the state of the system S:

$$\mathcal{U}(0,T)|\Psi\rangle\otimes|\phi(x)\rangle = \alpha|\uparrow\rangle\otimes|\phi(x-\frac{1}{2})\rangle+\beta|\downarrow\rangle\otimes|\phi(x+\frac{1}{2})\rangle. \tag{1.3-19}$$

If the variance σ of the pointer's position is large compared to the induced shift, then the two position distributions correlated with the $|\uparrow\rangle$ and $|\downarrow\rangle$ states largely overlap, and a projective measurement of the probe's position only minimally influences the state of the spin system. This is what justifies calling the measurement weak. In the other extreme, when the variance σ is small and the position distributions do not overlap, the state (1.3-19) becomes maximally entangled; the position is fully correlated with the spin state, and a strong, von Neumann measurement is recovered. By construction, a weak measurement does not reveal much information about the system under scrutiny. To resolve the shift in the pointer it is therefore necessary to repeat the measurement a large number of times.

1.3.4 Weak traces in counterfactual communication

Weak measurements have been proposed as a method for analysing the past of quantum particles [73, 77]. More concretely, if a weak measurement performed on a particle in a given region of space results in a resolvable pointer shift, then the particle is said to have left a weak trace there. It has been argued that the presence of a weak trace of a particle along a trajectory implies that the particle travelled along said trajectory, however this remains a point of contention [78, 79, 80, 81, 82]. A convenient



(a) Keeping which-path information (b) Erasing which-path information

Figure 1.8: **Weak trace in the TSV formalism.** Wheeler's delayed choice experiment is a thought experiment that probes the nature of particle-wave duality. In (a) a photon impinges on a beamsplitter, and is subsequently detected, forcing the photon to 'decide' which path it took, suggesting that it behaved as a particle at the time. By placing a second beamsplitter before the detector, as shown in (b), the photon will no longer behave like a particle, but will instead interfere with itself. Therefore the nature of the photon, particle or wave, is not decided at the first beamsplitter, but at the point of measurement. The trajectory of the photon can be analysed using the two-state-vector formalism, which includes not only the forwards propagating wavefunction shown in red, but also the post-selected backwards propagating wavefunction, shown in blue. The photon has a weak trace only in the regions where the forwards and backwards propagating are both present. In (a) this happens only along one arm of the interferometer, and the photon therefore has a definite trajectory, as one would expect. When the second beamsplitter is added, however, there is a weak trace of the photon in both arms of the interferometer.

method for analysing whether or not a weak trace of a photon in an interferometer is present is given by the so-called two-state vector formalism [83, 84]. In this framework a quantum state is described, in the time between two measurements at an initial time t_i and a final time t_f, by a two-state vector:

$$\langle\Phi|\ \ |\Psi\rangle\,, \tag{1.3-20}$$

where $|\Psi\rangle$ is the pre-measurement state from time $t = t_i$ evolving forwards in time, and $\langle\Phi|$ is the post-measurement state at time $t = t_f$ evolving backwards in time. A particle leaves a weak trace in a region of space if the forwards and backwards evolving components have a non-zero overlap in that region of space:

$$\langle\Psi|\Phi\rangle \neq 0, \tag{1.3-21}$$

no matter how small. This gives a visual method for evaluating the weak trace inside an interferometer, by drawing the particle trajectories from the initial and final states; in the regions where both the forward and backward particle paths are present the weak trace is non-vanishing. An example of this for Wheeler's delayed choice experiment is shown in Fig. 1.8. A similar analysis applied to the Salih protocol for counterfactual communication reveals that when Bob transmits a 0 bit there is a weak trace of the photon inside his laboratory, as illustrated in Fig. 1.9. A protocol for counterfactual communication that does not leave a weak trace was proposed by Arvidsson-Shukur and Barnes in [85]; its experimental realisation was the subject of Publication 1. This

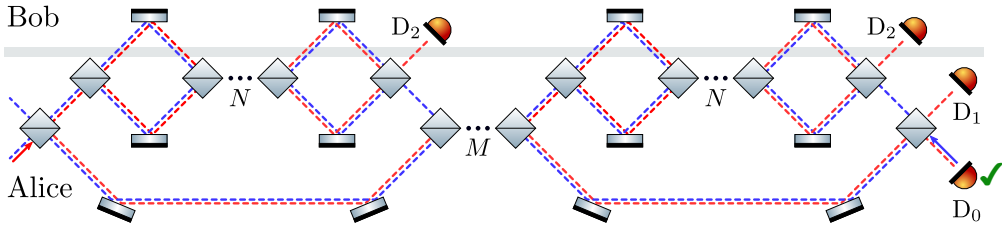


Figure 1.9: **Weak trace in the Salih CFC protocol.** Applying the two-state-vector formalism to the Salih et al. CFC protocol reveals the fact that even though no photon ever travels from Bob to Alice when he leaves his mirrors unblocked, both the forwards and backwards propagating wavefunctions are present in Bob's laboratory, meaning that there is a weak trace of the photon there. The backwards propagating wavefunction from the post-selected state at D_0 is shown in blue, while the forwards propagating wavefunction is shown in red.

protocol differs in its notion of counterfactuality in that particles travel from Alice to Bob, and particles are therefore exchanged. However, no particles that enter Bob's laboratory ever travel back to Alice. The Arvidsson-Shukur and Barnes protocol is sometimes referred to as a type-II protocol, to distinguish from the type-I protocols in which Alice and Bob never exchange particles.

1.3.5 QKD using counterfactual communication

While counterfactual communication does not offer any apparent advantage over more traditional forms of communication, both quantum and classical, the property of the photon remaining in Alice's laboratory can be used to establish a secure key [86]. This is due to the simple fact that if the photon does not travel to Bob's laboratory then it cannot be intercepted by an eavesdropper. It's possible to perform QKD with both type-I and type-II counterfactual communication protocols. In the case of type-II protocols the QKD method consists of two counterfactual communication channels and works by Alice randomly sending the photon into one of the channels. Bob simultaneously and independently picks one of the channels to block. If they both pick the same channel, and Bob blocks the mirrors in the channel Alice sent the photon into, then the photon never enters Bob's laboratory and returns to Alice. All the rounds where the photon went to Bob are disregarded in analogy with the BB84 protocol, and in the remaining rounds Alice and Bob know that their guesses were correlated without revealing what those guesses were.

2 Photonic quantum information processing

In this chapter some important concepts in the quantum theory of light will be introduced. This includes both the description of quantum states of light as well their transformation using various optical devices. Particular focus will be given to the manipulation and measurement of polarization-encoded qubit states, as well as the description of single-photon sources based on spontaneous parametric down-conversion

2.1 Introduction

There are a multitude of different physical systems that have been investigated as potential candidates for quantum information processing, all with their different strengths and weaknesses. The challenge of finding a suitable physical platform consists in reconciling two contradictory criteria, namely that the system should have very long coherence times, which requires it to be very well isolated from its classical environment, while at the same time allowing for strong interactions between the qubits themselves. The problem lies within the fact that a strong interaction between the qubits in most cases leads to strong interactions with the environment as well. As an example, qubits encoded in the nuclear spin of individual silicon atoms exhibit very long coherence times, exceeding several tens of milliseconds [87], due to the nuclear spins being extremely well isolated from their environment. However, this same property makes the implementation of qubit-qubit interactions challenging, with the best gate fidelities being in the mid 90s [88]. On the other hand, architectures based on quantized flux or charge in a superconducting circuit, often simply referred to as superconducting qubits, struggle to achieve coherence times even two orders of magnitude shorter than the ones in silicon [89], but have enabled demonstrations of high two-qubit gate fidelities [90, 91].

Photons are quite different from the matter based systems discussed above in that they, for all intents and purposes, do not interact with one another. In classical electrodynamics this is a direct consequence of the linearity of Maxwell's equations. When moving to a quantized description of light using quantum electrodynamics, photon-photon scattering becomes possible through the creation of virtual electron-positron pairs [92]. However, being a fourth order process, the corresponding scattering cross section is negligible and is further suppressed by the condition of energy conservation when the photon energies are far away from the rest mass of the electron. For low-energy photons this interaction is therefore unobservable in practice. This poses a challenge for photonic quantum information processing as a nonlinear photon-photon interaction is

required to generate entanglement. On the other hand the interactions between photons and macroscopic objects is generally not sensitive to the microscopic properties of the object in question, meaning that photons suffer little to no decoherence. For instance, the reflection of a photon off a mirror is caused by a collective response in the mirror and does not depend on microscopic and hard to control effects such as stray surface charges, and while the photon does impart some momentum to the mirror the two do not become entangled since the phonons in the mirror will have a warm thermal distribution, and a tiny perturbation to this state results in a new distribution with almost unity overlap with the initial one. The state of a mirror is therefore left unchanged by a photon reflecting off it.[1]

The strong coherence properties of photons led them to be the platform used for many pioneering experiments in quantum information, including the violation of Bell inequalities [3, 95], superdense coding [19], quantum teleportation [7, 30], entanglement swapping [31] and measurement-based quantum computation [96]. Particularly in the case of quantum communication protocols photons are the only viable information carrier, since matter-based ones are infeasible to transmit or transport. Thanks to ultra-low loss optical fibers single photon are able to coherently transmit quantum information over distances upwards of a hundred kilometers. These fibers have such remarkably low absorption that they beat most state-of-the-art quantum memories in terms of storage time and extraction efficiency [97]. These unique properties ensure that photonic quantum information will play a key role in future quantum technologies.

The rest of this chapter will introduce theoretical tools and experimental methods used in photonic quantum information processing, beginning with how quantum states of light are described, before moving on to how they can be generated and manipulated.

2.2 A quantum description of light

While there are many physical scenarios in which the interaction between quantum objects, such as atoms, and light can be described in a semi-classical way by treating light as a classical electromagnetic wave altering the Hamiltonian of the system in question [98, 99], there are also many effects that only become apparent when the full quantized nature of the electromagnetic field is considered. These quantum effects are most apparent for very weak intensities of light, as states with more photons are akin to macroscopic objects in which the quantum effects quickly wash out.

2.2.1 Second quantization

Any multi-particle quantum system can be described by a many-body wavefunction expressed in terms of products of single-particle wavefunctions of the form

$$\Psi(\vec{r}_1, \ldots, \vec{r}_N) = \prod_{k=1}^{N} \psi_{a_k}(\vec{r}_k). \tag{2.2-1}$$

[1] In the field of quantum optomechanics one specifically endeavours to engineer systems in which a single photon can dramatically change for example the vibrational modes of a macroscopic object, and become entangled with phonons inside the material [93, 94].

According to the spin-statistics theorem, such many-body wavefunctions must be either symmetric or anti-symmetric under particle exchange [100, 101]:

$$\text{Fermions}: \quad \Psi(\vec{r}_1,\ldots,\vec{r}_i,\ldots,\vec{r}_j,\ldots,\vec{r}_N) = -\Psi(\vec{r}_1,\ldots,\vec{r}_j,\ldots,\vec{r}_i,\ldots,\vec{r}_N) \tag{2.2-2}$$

$$\text{Bosons}: \quad \Psi(\vec{r}_1,\ldots,\vec{r}_i,\ldots,\vec{r}_j,\ldots,\vec{r}_N) = +\Psi(\vec{r}_1,\ldots,\vec{r}_j,\ldots,\vec{r}_i,\ldots,\vec{r}_N) \tag{2.2-3}$$

Since photons are bosons any multi-photon wavefunction has to be symmetric under particle exchange. The correct symmetry can be achieved by taking a superposition of permutations of the wavefunction in (2.2-1). For states consisting only of a few particles the symmetrised wavefunction is easy to write down explicitly; for example, the wavefunction of a two-photon state can be written as

$$\Psi(\vec{r}_1,\vec{r}_2) = \frac{\psi_{a_1}(\vec{r}_1)\psi_{a_2}(\vec{r}_2) + \psi_{a_2}(\vec{r}_1)\psi_{a_1}(\vec{r}_2)}{\sqrt{2}}, \tag{2.2-4}$$

however for higher numbers of particles this quickly becomes impractical as the number of permutations in the wavefunction rapidly grows. For this reason it becomes necessary to introduce a new formalism that more succinctly describes many-body systems, and this formalism is called second quantization. Instead of using a wavefunction that is written in terms of which identical particle occupies which identical state, one instead condenses this information by only counting the number of identical particles in a given state. In analogy with the quantum harmonic oscillator this is done by introducing the creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$, respectively. These operators have the effect of adding or removing a single quanta of light in a given mode. One can show that in order to preserve the correct symmetry of the wavefunction, these operators need to obey the following commutation relations [18, 101]

$$[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}, \tag{2.2-5}$$

$$[\hat{a}_i, \hat{a}_j] = [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0, \tag{2.2-6}$$

where the subscript indicates the mode, and modes with different indices are orthogonal. The space on which the creation and annihilation operators act is called the Fock space, whose basis states are states with a definite photon number. On such states the creation and annihilation operators have the following effect:

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1}\,|n+1\rangle \tag{2.2-7}$$

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle\,. \tag{2.2-8}$$

The creation and annihilation operators are not Hermitian and do therefore not correspond to any measurable quantity, nor can they be physically realised despite claims to the contrary [102, 103]. It is, however, possible to construct observables using these operators. One example of such an observable is the number operator, which measures the number of photons in a given mode:

$$\hat{N} = \hat{a}^\dagger \hat{a} \tag{2.2-9}$$

$$\hat{N}\,|n\rangle = n\,|n\rangle\,. \tag{2.2-10}$$

In this formalism it is possible to write down anti-symmetric states, for example the singlet Bell state, expressed here in terms of horizontal and vertical photon polarization:

$$\frac{1}{\sqrt{2}}\left(\hat{a}_H^\dagger \hat{b}_V^\dagger - \hat{a}_V^\dagger \hat{b}_H^\dagger\right)|0\rangle = \frac{|HV\rangle - |VH\rangle}{\sqrt{2}}, \tag{2.2-11}$$

however the corresponding wave-functions are nevertheless guaranteed to be symmetric, since they are compositions of symmetric wave-functions.

2.2.2 Classical states of light

Classically there are essentially two different kinds of light, the first one is thermal light emitted from a black body, and the second one is coherent light, such as the light generated by a laser. The difference between the two is that in thermal light the different frequency components do not have a well defined phase relationship between one another, whereas coherent light, as the name suggests, does have a well defined phase, with the degree of coherence depending on the specifics of the light. Both of these types of light can be described in the formalism of second quantization as well. In the case of a thermal state, the lack of coherence is reflected in the state being a mixture of different Fock states [104]:

$$\rho_{\text{th}} = \sum_{n=0}^{\infty} P(n)\,|n\rangle\langle n|\,, \tag{2.2-12}$$

where

$$P(n) = \frac{\langle N\rangle^n}{(1+\langle N\rangle)^{n+1}}, \tag{2.2-13}$$

is the probability to detect n photons. The thermal state can equivalently expressed as

$$\rho_{\text{th}} = (1-\xi)\sum_{n=0}^{\infty} \xi^n\,|n\rangle\langle n|\,, \tag{2.2-14}$$

where

$$\xi = \exp\left[-\frac{\hbar\omega}{k_B T}\right], \tag{2.2-15}$$

is related to the temperature of the emitting black body.

The second kind of classical light is coherent light, usually generated by a laser. In an idealized picture coherent light is perfectly monochromatic and is an infinitely extending sine wave. In the framework of second quantization coherent states are the eigenstates of the annihilation operator:

$$\hat{a}^\dagger\,|\alpha\rangle = \alpha\,|\alpha\rangle\,. \tag{2.2-16}$$

It is easy to verify that the following photon number distribution describes exactly these eigenstates:

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\,|n\rangle\,. \tag{2.2-17}$$

Here $|\alpha|^2$ is the mean photon number.

2.2.3 Photon statistics

Since the states of light in the previous section were written in the Fock basis, and therefore in terms of their photon number distributions, it is natural to investigate these distributions, and in particular their variances. For the coherent state, the variance is easy to calculate from the photon number probability distribution, which in turn is found by taking the expectation value of the projector onto the n-th Fock state:

$$\begin{aligned} P(n) &= \langle\alpha|(|n\rangle\langle n|)|\alpha\rangle \\ &= \frac{|\alpha|^{2n}}{n!}e^{-|\alpha|^2} \\ &= \frac{\langle N\rangle^n}{n!}e^{-\langle N\rangle}. \end{aligned} \tag{2.2-18}$$

This is a Poissonian distribution, and coherent light is for this reason sometimes referred to as Poissonian light. The variance of the photon number distribution is

$$\Delta N^2 = \langle N\rangle. \tag{2.2-19}$$

This variance in the photon number is the origin of the square-root scaling of shot noise in laser light. For thermal light, the probability to detect n photons can be found from (2.2-14) as

$$P(n) = \mathrm{Tr}\big[\,|n\rangle\langle n|\,\rho_{\mathrm{th}}\big] = (1-\xi)\xi^n. \tag{2.2-20}$$

This is known as a geometric distribution, and it has a mean and variance of

$$\langle N\rangle = \frac{\xi}{\xi - 1} \tag{2.2-21}$$

$$\Delta N^2 = \frac{1-\xi}{\xi^2}. \tag{2.2-22}$$

Inverting (2.2-21) to express ξ in terms of $\langle n\rangle$:

$$\xi = \frac{\langle N\rangle}{1+\langle N\rangle}, \tag{2.2-23}$$

and inserting this into (2.2-22) one finds, after some manipulation

$$\Delta N^2 = \langle N\rangle^2 + \langle N\rangle. \tag{2.2-24}$$

This variance in the photon number is qualitatively different from coherent light, since the relative uncertainty in the photon number does not increase with the intensity of the light. More quantitatively, a coherent light source has a signal-to-noise ratio (SNR) given by

$$\mathrm{SNR}_{\mathrm{coherent}} = \sqrt{\frac{\langle N\rangle^2}{\Delta N^2}} = \sqrt{\langle N\rangle}, \tag{2.2-25}$$

while for a thermal light source the signal-to-noise ratio is

$$\mathrm{SNR}_{\mathrm{th}} = \sqrt{\frac{\langle N\rangle^2}{\Delta N^2}} \approx 1. \tag{2.2-26}$$

This approximation is valid for warm thermal states where $\langle N\rangle^2 \gg \langle N\rangle$. The two classical types of light discussed above both have a variance in the photon number equal to or greater than $\langle N\rangle$, however it is self-evident that the photon number states have zero variance:

$$|n\rangle: \quad \Delta N^2 = 0, \tag{2.2-27}$$

since they are per definition definite photon number states. This is a signature of non-classicality, and states that exhibit a variance below $\langle N\rangle$ cannot be described by classical electrodynamics [105]. Note that the vanishing variance for the Fock states does not imply that these states can be used to achieve an infinite signal-to-noise ratio, since any process continuously encoding information in the photon number will invariably create a state described by a photon number superposition, which does have a variance.

2.2.4 Second-order coherence

For classical fields of light one can quantify the coherence in terms of a correlation function [104, 106]:

$$g^{(1)}(\tau) = \frac{\langle E^*(t)E(t+\tau)\rangle}{\langle |E(t)|^2\rangle}, \tag{2.2-28}$$

where τ is a time delay with respect to the time t. This function, known as the first-order correlation function, describes the visibility of interference between a light field and a time delayed version of itself. In the case of for example coherent light, one finds that $|g^{(1)}(\tau)| = 1$, since the light perfectly interferes with itself even at arbitrary delays, and has infinite coherence length. It is also possible to define related functions involving higher powers of the electric field, for example the second-order correlation function [104]:

$$g^{(2)}(\tau) = \frac{\langle E^*(t)E^*(t+\tau)E(t+\tau)E(t)\rangle}{\langle E^*(t)E(t)\rangle^2} = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)^2\rangle}, \tag{2.2-29}$$

where I is the intensity of the light. In analogy to how the $g^{(1)}$ function quantifies how correlated the phases of a field are at different points in time or space, the second-order correlation function quantifies how correlated the intensities are. To measure this quantity one typically employs a so-called Hanbury-Brown-Twiss interferometer [107], which transforms the problem of measuring the correlation between intensities at two different times, to a problem of measuring correlations between intensities at two different points in space and time.

A second-order correlation value of $g^{(2)}(\tau) = 1$ indicates that there is no correlation between the intensity fluctuations in the two detectors, and coherent light has exactly this property. Thermal light, on the other hand, shows a positive correlation at zero delay, more specifically

$$g^{(2)}(0) = 2, \tag{2.2-30}$$

and these intensity correlations gradually vanish as the time delay is increased, at a rate that depends on the coherence time of the light.

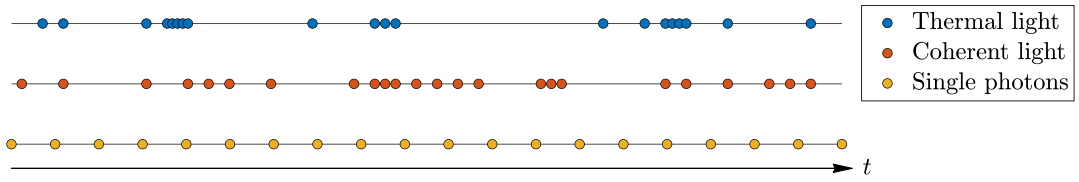


Figure 2.1: **Examples of photon statistics.** The figure illustrates three different types of photon statistics, by showing their distribution in time. Thermal states (blue dots), also known as bunched states, have the property that the detection of one photon increases the conditional probability of detecting another one shortly thereafter. Coherent states (red dots), on the other hand, show no such correlation, and the photon detection probability at any given time is independent of whether a photon was detect recently. Finally, anti-bunched light in the form of single photons (yellow dots) has the property that the detection of a photon ensures that another one will not be detected immediately thereafter.

In the quantum picture the second-order correlation function can be defined as [104, 108]

$$g^{(2)}(\tau) = \frac{\langle \hat{a}^\dagger(t)\hat{a}^\dagger(t+\tau)\hat{a}(t+\tau)\hat{a}(t)\rangle}{\langle \hat{a}^\dagger(t)\hat{a}(t)\rangle^2}, \tag{2.2-31}$$

written in the Heisenberg picture where the operators are time dependent. In the photon picture the second-order correlation function can be thought of as describing how likely one is to detect a photon at time $t+\tau$, given that one was that one was detected at time t. For coherent light the fact that $g^{(2)}(\tau) = 1$ means that the probability to detect a photon at time $t+\tau$ is independent of whether or not a photon was detected at time t. However, for thermal light the intensity correlation implied by $g^{(2)}(0) = 2$ means that the probability to detect a photon at $t+\tau$ for a small τ is greater when one was detected at time t. In other words, the photons tend to bunch together. For example, if one does a non-destructive measurement on a thermal state that reveals that there are photons present:

$$\frac{(I - |0\rangle\langle 0|)\rho_{\text{th}}(I - |0\rangle\langle 0|)}{\text{Tr}[|0\rangle\langle 0|\,\rho_{\text{th}}]} = \rho'_{\text{th}}, \tag{2.2-32}$$

the effect is that the expectation value of the photon number increases:

$$\langle N\rangle_{\rho'_{\text{th}}} = \langle N\rangle_{\rho_{\text{th}}} + 1, \tag{2.2-33}$$

since

$$\begin{aligned}\rho'_{\text{th}} &= \frac{1}{1-P(0)}\sum_{n=1}^{\infty}\frac{\langle N\rangle^n}{(1+\langle N\rangle)^{n+1}}\,|n\rangle\langle n| \\ &= \sum_{n=1}^{\infty}\frac{\langle N\rangle^{n-1}}{(1+\langle N\rangle)^n}\,|n\rangle\langle n| \\ &= \sum_{n=0}^{\infty}\frac{\langle N\rangle^n}{(1+\langle N\rangle)^{n+1}}\,|n+1\rangle\langle n+1|\,.\end{aligned} \tag{2.2-34}$$

This is a reflection of the bunched nature of the light. For classical light it is straightforward to show that [104]

$$g^{(2)}(0) \geq 1, \tag{2.2-35}$$

however the same is not true for quantum states of light. The quantum second-order correlation function (2.2-31) can be expanded at $\tau = 0$ as

$$g^{(2)}(0) = \frac{\langle \hat{a}^\dagger \hat{a}^\dagger \hat{a}\hat{a} \rangle}{\langle \hat{a}^\dagger \hat{a} \rangle^2} = \frac{\langle \hat{a}^\dagger (\hat{a}\hat{a}^\dagger - 1)\hat{a} \rangle}{\langle N \rangle^2} = \frac{\langle N^2 - N \rangle}{\langle N \rangle^2} = 1 + \frac{\Delta N^2 - \langle N \rangle}{\langle N \rangle^2}. \tag{2.2-36}$$

Taking for example the single-photon state $|1\rangle$ and inserting the variance and photon number one finds

$$g^{(2)}(0) = 0, \tag{2.2-37}$$

and in general the Fock states obey [104]

$$g^{(2)}(0) = 1 - \frac{1}{n}. \tag{2.2-38}$$

Photon-number states are anti-bunched, and in a Hanbury-Brown-Twiss interferometer this means that the detection of a photon in one arm lowers the probability of another photon simultaneously being detected in the other arm. For a pure single-photon states in particular this probability reduces to zero. Different kinds of light, bunched, coherent and anti-bunched, are figuratively illustrated in Fig. 2.1. Due to the aforementioned fact that only quantum states of light can exhibit a second-order correlation value of less than one, this is often used as a metric for non-classicality of light. Similarly, when characterising sources of single photons the $g^{(2)}(0)$-value is often a proxy for the quality of the single photons, since the state $|1\rangle$ is the only one for which $g^{(2)}(0) = 0$.

2.2.5 Phase-space representation

While any state of light can be represented as either a coherent superposition, or a mixture of Fock states, alternative representations can often provide new insights. One alternative quantum description of light is the so-called phase-space representation. Let us first introduce the quadrature operators [26, 109]

$$\hat{x} = \frac{1}{\sqrt{2}}(\hat{a}^\dagger + \hat{a}), \quad \hat{p} = \frac{i}{\sqrt{2}}(\hat{a}^\dagger - \hat{a}). \tag{2.2-39}$$

These operators, analogous to the position and momentum operators of the harmonic oscillator, represent the in- and out-of-phase components of the electric field, respectively. Or, equivalently, the real and imaginary parts of the complex amplitude of the annihilation operator. Since they are Hermitian operators they correspond to observable quantities, and can be measured for example using a technique called homodyne detection (see Fig. 2.2) [110]. These operators have (improper) eigenstates

$$\hat{x}\,|x\rangle = x\,|x\rangle\,, \quad \hat{p}\,|p\rangle = p\,|p\rangle\,. \tag{2.2-40}$$

Using these eigenstates as a continuous basis, we can now write down a transformation between a density matrix ρ in the Fock space and its phase-space representation:

$$W(x,p) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \langle x - \frac{1}{2}q|\rho|x + \frac{1}{2}q\rangle e^{ipq} dq. \tag{2.2-41}$$

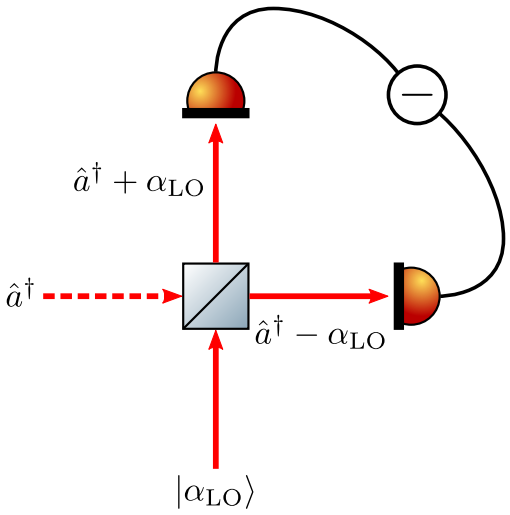


Figure 2.2: **Balanced homodyne detection.** The quadratures of the electric field of a quantum state can be measured using a method called balanced homodyne detection. It consists of interfering the quantum state in question with a strong coherent beam, referred to as the local oscillator (LO) on a balanced beam-splitter. Since the local oscillator can be treated classically the effect of the beamsplitter is simply to add a displacement of $\alpha_{\mathrm{LO}} = |\alpha_{\mathrm{LO}}|e^{i\theta}$ to the creation operator of the input quantum mode (a $1/\sqrt{2}$ scaling factor has been omitted in the figure). The photodiodes then generate a photocurrent proportional to the number operators of the two modes $I_\pm \propto N_\pm = (\alpha^*_{\mathrm{LO}} \pm \hat{a}^\dagger)(\alpha_{\mathrm{LO}} \pm \hat{a})/2$. Subtracting the two photocurrent yields a quadrature operator: $I_+ - I_- \propto \alpha^*_{\mathrm{LO}}\hat{a} + \alpha_{\mathrm{LO}}\hat{a}^\dagger = |\alpha_{\mathrm{LO}}|(\hat{a}e^{-i\theta} + \hat{a}^\dagger e^{i\theta})$.

The above map is called the Wigner transformation, after Eugene Wigner, and the function $W(x,p)$ is known as the Wigner function or Wigner quasiprobability distribution [111, 112, 113]. The reason for the latter name is the fact that the function can take on negative values, which cannot be interpreted as probabilities. Classical states of light always have positive Wigner functions, and negativity of the Wigner function is therefore an indicator of non-classicality [112]. However, the converse statement is not true, and there exist non-classical states of light exhibiting non-local correlations, that nevertheless have positive Wigner functions. The Wigner function does, however, have positive marginal distributions, and these correspond to the probability distributions for the $\hat{x}$ and $\hat{p}$ quadrature operators [114]:

$$P(x) = \langle x|\rho|x\rangle = \int_{-\infty}^{\infty} W(x,p)dp \tag{2.2-42}$$

$$P(p) = \langle p|\rho|p\rangle = \int_{-\infty}^{\infty} W(x,p)dx. \tag{2.2-43}$$

In the phase-space description, the vacuum state $|0\rangle$ is a Gaussian minimum uncertainty state, with $\Delta x^2 = \Delta p^2 = \frac{1}{2}$. The coherent states can be represented as displacements of the vacuum state, effected by the displacement operator

$$\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}, \tag{2.2-44}$$

where $\alpha = x + ip$ is a complex number. The displacement operator acting on the vacuum

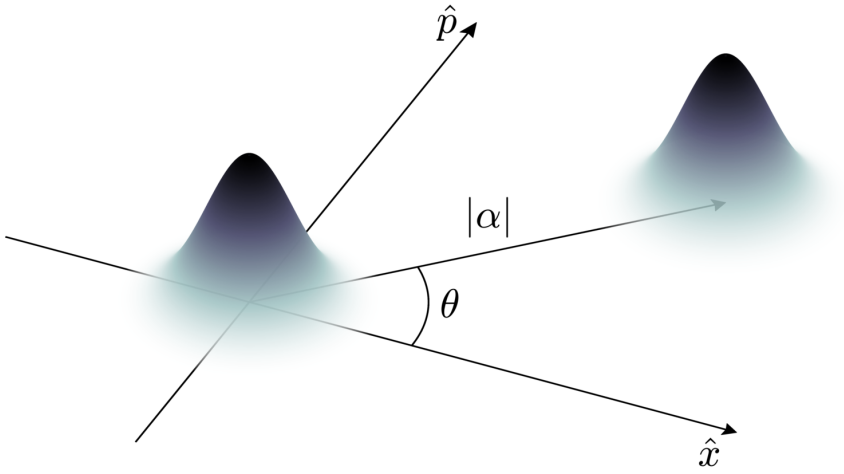


Figure 2.3: **Phase-space displacement.** The displacement operator $\hat{D}(\alpha) = \exp[\alpha\hat{a}^\dagger + \alpha^*\hat{a}]$ has the effect of displacing the Wigner function of a state by an amount $|\alpha|$ along a direction defined by the complex argument of $\alpha = |\alpha|e^{i\theta}$. The figure shows the action of the displacement operator on the vacuum state. The resulting state is a coherent state with average photon number $|\alpha|^2$.

has the following effect:

$$\hat{D}(\alpha)\,|0\rangle = |\alpha\rangle\,. \tag{2.2-45}$$

In the phase-space picture this corresponds to displacing the vacuum by an amount $|\alpha|$, along a direction defined by the complex argument of the same variable, hence the name. This is illustrated in Fig. 2.3.

The photon-number states, in contrast to the coherent states, have a characteristic negativity. Their Wigner functions are given by a Gaussian function weighted by a Laguerre polynomial [115, 116]:

$$W_n(x,p) = \frac{(-1)^n}{\pi} e^{-(x^2+p^2)} L_n(2(x^2+p^2)), \tag{2.2-46}$$

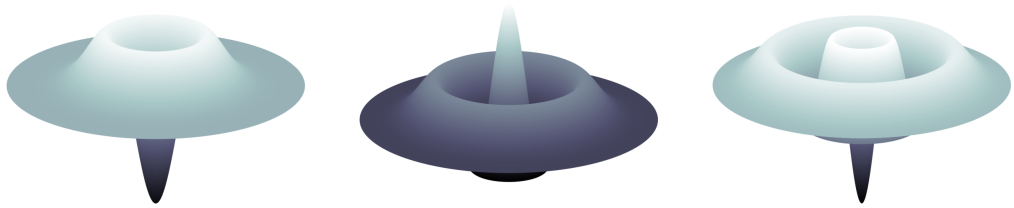
where L_n is the n-th order Laguerre polynomial. A few Fock-state Wigner functions for low photon numbers are shown in Fig. 2.4.

2.2.6 Squeezed states

As mentioned in the previous section, there are states which cannot be described by classical electrodynamics, which nevertheless have strictly non-negative Wigner functions. An important example of such states are the squeezed states. Consider the so-called single-mode squeezing operator [117]

$$\hat{S}(\zeta) = \exp\Big[\frac{\zeta^*\hat{a}^2 - \zeta(\hat{a}^\dagger)^2}{2}\Big], \tag{2.2-47}$$

where $\zeta = re^{i\theta}$ is the squeezing parameter. The name comes from the fact that this operator compresses the Wigner function along one axis, and stretches it along another.



(a) $|1\rangle$-state (b) $|2\rangle$-state (c) $|3\rangle$-state

Figure 2.4: **Wigner functions of Fock states.** The Wigner functions of Fock states, shown here for the first three non-vacuum states, display a characteristic negativity, arising from the Laguerre-polynomial factor. The number of zero-crossings of the Wigner function is equal to the photon number.

The axis along which the state is squeezed is given by the parameter θ. Setting for example $\theta = 0$ yields an operation that transforms the quadrature operators in the following way:

$$\hat{x} \mapsto e^{-r}\hat{x}, \qquad \hat{p} \mapsto e^{r}\hat{p}. \tag{2.2-48}$$

The squeezing transformation also alters the variances of the quadrature operators:

$$\Delta x^2 \mapsto \frac{\Delta x^2}{R^2}, \qquad \Delta p^2 \mapsto R^2 \Delta p^2, \tag{2.2-49}$$

where $R = e^r$. If the initial state is in a minimum uncertainty state, then the squeezing operator preserves this property since the products of the variances in the quadratures remains unchanged. However, squeezing can be used to decrease the variance in a single quadrature below the Heisenberg limit. This fact has enabled precision sensing below the shot-noise limit, and squeezed vacuum was recently employed in the gravitational wave detector LIGO to increase its sensitivity [118, 119]. The effect of the squeezing on the variance of the quadrature operators is quite easily understood visually, since the variance in the quadrature operators is simply the width of the Wigner function projected onto the corresponding coordinate axis. Applying the squeezing operator on the vacuum state $|0\rangle$ one obtains the single-mode squeezed-vacuum state [117]

$$|\mathrm{SV}\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} (-e^{-i\theta} \tanh r)^n \frac{\sqrt{(2n)!}}{2^n n!} |2n\rangle \,. \tag{2.2-50}$$

The Wigner function of this state is illustrated in Fig. 2.5. As can be read out from (2.2-50), only even photon numbers occur in the state; a consequence of the fact that the lowest order of the creation operator in the Taylor expansion of (2.2-47) is two. Also note that in the limit of infinite squeezing, i.e. $r \to \infty$, the mean photon number diverges, meaning that the state has infinite energy. A closely related state to the single-mode squeezed vacuum is the two-mode squeezed vacuum, which is defined as the state one gets when the vacuum is acted on by the two-mode squeezing operator

$$\hat{S}_2(\zeta) = \exp\Big[\frac{\zeta^* \hat{a}\hat{b} - \zeta \hat{a}^\dagger \hat{b}^\dagger}{2}\Big]. \tag{2.2-51}$$

(a) $|\zeta| = 0$ (b) $|\zeta| = 0.7$ (c) $|\zeta| = 1.2$

Figure 2.5: **Squeezed vacuum states.** The Wigner function of a squeezed vacuum state can be expressed as $W(\alpha, \zeta) = 1/\pi \exp\big[-2|\alpha \cosh r + \alpha^* e^{i\theta} \sinh r|^2\big]$, where $\zeta = re^{i\theta}$ is the squeezing parameter and $\alpha = x + ip$ [120]. This corresponds to a 2D-Gaussian that is compressed along the axis of squeezing, and stretched along the other axis. This is illustrated above, for three different states with progressively more squeezing. The first state from the left, with $|\zeta| = 0$, is a normal vacuum state, while the last two have squeezing parameters of $|\zeta| = 0.7$ and $|\zeta| = 1.2$, repectively.

The resulting state is

$$|\text{TMSV}\rangle = \hat{S}_2(\zeta)\,|0\rangle = \frac{1}{\cosh r} \sum_{n=0}^{\infty} (-e^{-i\theta} \tanh r)^n\,|n, n\rangle\,. \tag{2.2-52}$$

This state is entangled in the photon number, since this quantity is perfectly correlated between the two modes. It is also entangled in the field quadratures. Setting $\theta = 0$ for simplicity, the quadrature operators of the squeezed modes obey [117]

$$\hat{x}_a(r) \pm \hat{x}_b(r) = (\hat{x}_a(0) \pm \hat{x}_b(0))e^{\pm r}, \tag{2.2-53}$$

$$\hat{p}_a(r) \pm \hat{p}_b(r) = (\hat{p}_a(0) \pm \hat{p}_b(0))e^{\mp r}. \tag{2.2-54}$$

In the limit of infinite squeezing the position quadratures become perfectly correlated, while the momentum quadratures become perfectly anti-correlated:

$$\hat{x}_a(\infty) - \hat{x}_b(\infty) = 0 \tag{2.2-55}$$

$$\hat{p}_a(\infty) + \hat{p}_b(\infty) = 0. \tag{2.2-56}$$

States with these correlations are referred to as EPR states, after the famous paper by Einstein, Podolsky and Rosen describing a thought experiment in which two particles become entangled in position and momentum [1, 26].

A natural consequence of the entanglement between the two modes is that the local description of either mode yields a mixed state. Tracing out one mode of the two-mode

squeezed vacuum one finds

$$\begin{aligned}\rho_a &= \mathrm{Tr}_b[\rho_{\mathrm{TMSV}}] \\ &= \sum_{m=0}^{\infty} \langle m|_b \left(\frac{1}{\cosh^2(r)} \sum_{n=0}^{\infty} \tanh^{2n}(r)\, |n,n\rangle\langle n,n|_{ab} \right) |m\rangle_b \\ &= \frac{1}{\cosh^2(r)} \sum_{n=0}^{\infty} \tanh^{2n}(r)\, |n\rangle\langle n|_a \\ &= (1-\tanh^2(r)) \sum_{n=0}^{\infty} \tanh^{2n}(r)\, |n\rangle\langle n|_a \,.\end{aligned} \tag{2.2-57}$$

Making the identification $\tanh^2(r) = \xi$ it is easy to see that this is the same state as in (2.2-14). Locally, the two-mode squeezed vacuum therefore looks like a thermal state, with a temperature that depends on the degree of squeezing.

2.2.7 Linear optics

Most of the optical devices commonly used to manipulate light, such as lenses, mirrors, beamsplitters etc. are linear devices. Classically this means that their response to the electric field is linear, and does not depend on the intensity of the light. In the quantum picture these devices are described by Hamiltonians that are bilinear in the creation and annihilation operators [121]:

$$H = \sum_{jk} A_{jk} \hat{a}_j^\dagger \hat{a}_k. \tag{2.2-58}$$

Let

$$N_{\mathrm{tot}} = \sum_l \hat{a}_l^\dagger \hat{a}_l, \tag{2.2-59}$$

be the total number operator for the system. Then consider the commutator between this number operator and a single term in the Hamiltonian (2.2-58):

$$[\hat{a}_j^\dagger \hat{a}_k, N_{\mathrm{tot}}] = \sum_l \hat{a}_j^\dagger \hat{a}_k \hat{a}_l^\dagger \hat{a}_l - \hat{a}_l^\dagger \hat{a}_l \hat{a}_j^\dagger \hat{a}_k. \tag{2.2-60}$$

Using the identity

$$\hat{a}_k \hat{a}_l^\dagger = \hat{a}_l^\dagger \hat{a}_k + \delta_{lk}, \tag{2.2-61}$$

this can be simplified to

$$\begin{aligned}[\hat{a}_j^\dagger \hat{a}_k, N_{\mathrm{tot}}] &= \sum_l \hat{a}_j^\dagger (\hat{a}_l^\dagger \hat{a}_k + \delta_{lk}) \hat{a}_l - \hat{a}_l^\dagger (\hat{a}_j^\dagger \hat{a}_l + \delta_{jl}) \hat{a}_k \\ &= \sum_l \delta_{lk} \hat{a}_j^\dagger \hat{a}_l - \delta_{jl} \hat{a}_l^\dagger \hat{a}_k + \hat{a}_j^\dagger \hat{a}_l^\dagger \hat{a}_k \hat{a}_l - \hat{a}_l^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}k \\ &= \sum_l \delta_{lk} \hat{a}_j^\dagger \hat{a}_l - \delta_{jl} \hat{a}_l^\dagger \hat{a}_k.\end{aligned} \tag{2.2-62}$$

Evaluating the sum reduces this to

$$[\hat{a}_j^\dagger \hat{a}_k, N_{\text{tot}}] = \hat{a}_j^\dagger \hat{a}_k - \hat{a}_j^\dagger \hat{a}_k = 0. \tag{2.2-63}$$

This shows that the number operator commutes with every term in the Hamiltonian and therefore with the Hamiltonian itself:

$$[H, N_{\text{tot}}] = 0. \tag{2.2-64}$$

This in turn means that the photon number is conserved, and this can be seen as the defining feature of linear optics. Formally, all linear-optical transformations are generated by Hamiltonians of the form (2.2-58). For example, the action of a beamsplitter is generated by the Hamiltonian [26, 121]

$$H_{\text{BS}} = \theta(e^{i\varphi}\hat{a}^\dagger \hat{b} + e^{-i\varphi}\hat{a}\hat{b}^\dagger), \tag{2.2-65}$$

where θ is a real parameter that determines the splitting ratio of the beamsplitter and φ is a relative phase shift between the two modes. The corresponding unitary operator obtained by exponentiating the Hamiltonian:

$$U_{H_{\text{BS}}} = e^{-iH_{\text{BS}}}, \tag{2.2-66}$$

acts on the infinite-dimensional Fock space, and does therefore not have an exact matrix representation. To find how this unitary acts, consider a state generated by some polynomial of the creation operator acting on vacuum:

$$|\Psi\rangle = \sum_n \alpha_n (\hat{a}^\dagger)^n |0\rangle . \tag{2.2-67}$$

Using the fact that

$$U_{H_{\text{BS}}}^\dagger |0\rangle = |0\rangle , \tag{2.2-68}$$

the state after the beamsplitter can be written

$$\begin{aligned} U_{H_{\text{BS}}} |\Psi\rangle &= U_{H_{\text{BS}}} \sum_n \alpha_n (\hat{a}^\dagger)^n U_{H_{\text{BS}}}^\dagger |0\rangle \\ &= \sum_n \alpha_n (U_{H_{\text{BS}}} \hat{a}^\dagger U_{H_{\text{BS}}}^\dagger)^n |0\rangle . \end{aligned} \tag{2.2-69}$$

Hence, the action of the beamsplitter on the creation operator of the mode a is given by

$$\hat{a}^\dagger \mapsto U_{H_{\text{BS}}} \hat{a}^\dagger U_{H_{\text{BS}}}^\dagger. \tag{2.2-70}$$

Using the Baker–Campbell–Hausdorff lemma [18]

$$e^A B e^{-A} = \sum_n \frac{[(A)^n, B]}{n!}, \qquad [(A)^n, B] \equiv [\underbrace{A, \cdots [A, [A}_{\text{n times}}, B]] \cdots] \tag{2.2-71}$$

this can be expanded as

$$\hat{a}^\dagger - i[H_{BS}, \hat{a}^\dagger] + \frac{i^2}{2!}[H_{BS}, [H_{BS}, \hat{a}^\dagger]] - \frac{i^3}{3!}[H_{BS}, [H_{BS}, [H_{BS}, \hat{a}^\dagger]]] + \dots \tag{2.2-72}$$

It is straightforward to verify that

$$[H_{\mathrm{BS}}, \hat{a}^\dagger] = \theta e^{-i\varphi}\hat{b}^\dagger, \qquad [H_{\mathrm{BS}}, \hat{b}^\dagger] = \theta e^{i\varphi}\hat{a}^\dagger, \tag{2.2-73}$$

and using these commutator relations (2.2-72) simplifies to

$$\begin{aligned} \hat{a}^\dagger \mapsto \quad & \hat{a}^\dagger\left(1 + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^4}{4!} + \cdots\right) \\ & -\hat{b}^\dagger e^{-i\varphi}\left(\frac{i\theta}{1!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^5}{5!} + \cdots\right). \end{aligned} \tag{2.2-74}$$

These infinite series are just the Maclaurin series of the cosine and sine functions respectively, and the transformation can therefore be written in a succinct way as

$$\hat{a}^\dagger \mapsto \cos\theta\, \hat{a}^\dagger - ie^{-i\varphi}\sin\theta\, \hat{b}^\dagger. \tag{2.2-75}$$

Performing the same calculation for the second mode gives

$$\hat{b}^\dagger \mapsto -ie^{i\varphi}\sin\theta\, \hat{a}^\dagger + \cos\theta\, \hat{b}^\dagger. \tag{2.2-76}$$

This transformation can now, in contrast to (2.2-66), be described by a simple 2×2 matrix:

$$U_{\mathrm{BS}} = \begin{bmatrix} \cos\theta & -ie^{-i\varphi}\sin\theta \\ -ie^{i\varphi}\sin\theta & \cos\theta \end{bmatrix}, \tag{2.2-77}$$

written in the basis

$$\hat{a}^\dagger = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \hat{b}^\dagger = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{2.2-78}$$

Similar derivations can be done for other linear-optical operations and elements, such as the phase shift. The fact that these operations can be represented using matrices greatly simplifies calculations involving higher photon number states, since

$$(\hat{a}^\dagger)^n \mapsto (U_{H_{\mathrm{BS}}}\hat{a}^\dagger U_{H_{\mathrm{BS}}}^\dagger)^n = (U_{\mathrm{BS}}\hat{a}^\dagger)^n. \tag{2.2-79}$$

This means that the behaviour of a multi-photon state in a linear-optical circuit is encoded in the description of single photons in the same circuit.

Returning to the beamsplitter and setting $\theta = \pi/4$ in (2.2-77) one obtains the balanced, or $50:50$-beamsplitter, for which the convention $\varphi = \pi/2$ is usually picked:

$$U_{\mathrm{BS}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}. \tag{2.2-80}$$

This operator is commonly written in a form with only real entries, which can be done with the following transformation:

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \tag{2.2-81}$$

the physical meaning of which is applying a phase shift to one input and one output mode. The balanced beamsplitter is one of the most commonly used optical devices, and is the basic building block of interferometers.

Another frequently occurring optical element is the polarizing beamsplitter (PBS), whose input-output relations are given by

$$\begin{aligned} \hat{a}_H^\dagger &\mapsto \hat{a}_H^\dagger, \qquad \hat{a}_V^\dagger \mapsto \hat{b}_V^\dagger, \\ \hat{b}_H^\dagger &\mapsto \hat{b}_H^\dagger, \qquad \hat{b}_V^\dagger \mapsto \hat{a}_V^\dagger, \end{aligned} \tag{2.2-82}$$

where the subscripts H/V denote the photon polarization. This transformation correlates the output spatial mode with the input polarization state, and can therefore be used to perform a Von Neumann measurement on the photon polarization. This will be discussed in greater detail in Section 2.3.2.

2.2.8 Simulating linear optics

According to (2.2-79) a linear-optical circuit will transform any creation operator into a polynomial of creation operators over all the modes, and in general an n-photon state described by a product of n creation operators acting on the vacuum will be transformed into a state acted on by some n-th order polynomial of the creation operators on all the modes. This gives an explicit way of calculating the output amplitudes of any linear-optical circuit for any Fock-state input. However, this symbolic manipulation of polynomials is inconvenient in a numerical setting. An alternative but equivalent way of computing the output amplitudes of any linear-optical circuit was shown in Ref. [122]. There the authors proved that the transition amplitude from some Fock state to another is given by the permanent of the matrix describing the single-photon input-output relations. The permanent of a matrix $A = (a_{i,j})$ is defined as

$$\mathrm{Per}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i,\sigma(i)}, \tag{2.2-83}$$

where S_n is the set of all permutations of all the numbers $0, 1, \ldots n$. From this definition one can see that the permanent is very similar to the determinant, but does not include alternating signs.

Given a unitary U that describes the single-photon input-output relations of a linear circuit, the transition amplitude

$$\langle S|\varphi(U)|T\rangle \tag{2.2-84}$$

from an initial Fock state T:

$$|T\rangle = |t_1, \ldots, t_n\rangle\,, \tag{2.2-85}$$

to a final state S:

$$|S\rangle = |s_1, \ldots, s_n\rangle\,, \tag{2.2-86}$$

is given by

$$\langle S|\varphi(U)|T\rangle = \frac{\mathrm{Per}(U_{S,T})}{\sqrt{s_1! \cdots s_n! t_1! \cdots t_n!}}. \tag{2.2-87}$$

Here $U_{S,T}$ is a matrix constructed from U by repeating columns and rows from U, and $\varphi(U)$ is the unitary acting in the 2^n-dimensional n-photon space, which does not need to be explicitly constructed. The matrix $U_{S,T}$ is found by first building the matrix U_T by taking t_i copies of the i-th column of U. $U_{S,T}$ is then obtained by taking s_i copies of the i-th row of U_T. As an example, consider the matrix

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \tag{2.2-88}$$

and the input/output states

$$|T\rangle = |1,1,2\rangle\,, \qquad |S\rangle = |1,2,1\rangle\,. \tag{2.2-89}$$

The matrix U_T is then

$$U_T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \tag{2.2-90}$$

and $U_{S,T}$ is

$$U_{S,T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{2.2-91}$$

Finally, the transition amplitude is given by

$$\frac{\mathrm{Per}(U_{S,T})}{\sqrt{1!2!1!1!1!2!}} = \frac{2}{\sqrt{2\cdot 2}} = 1. \tag{2.2-92}$$

In this simple example the transition amplitude is 1 because the matrix U only swaps the modes 2 and 3 and therefore does not create a superposition state. Note that the action of repeating rows and columns of U commutes, meaning that one can equivalently begin by creating the matrix U_S through taking s_i copies of the i-th row of U. The matrix $U_{S,T}$ can also be constructed through direct matrix multiplication:

$$U_{S,T} = I_S^T U I_T, \tag{2.2-93}$$

where the superscript denotes the transpose, and where I_S (I_T) is a matrix consisting of s_i (t_i) ones in each row, preceded by $\sum_{k=0}^{i-1} s_k$ ($\sum_{k=0}^{i-1} t_k$) zeros. For the example above these matrices are

$$I_S = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad I_T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \tag{2.2-94}$$

In general, the full output state given a Fock-state input is obtained by evaluating (2.2-87) for every possible output state containing the same number of photons as the input. For a system with n photons and m modes the total number of possible output

states is $\binom{m+n-1}{n}$ [122], which quickly becomes large even for a modest number of photons and modes. In practice, however, one is often only interested in particular types of output states, for example ones in which all the photons are anti-bunched and end up in different modes. This makes the approach using the permanent particularly convenient, since one can directly compute the relevant amplitudes, which is a simpler problem than the full evolution of the input state.

The approach also lends itself well to simulations of non-trivial input states, such as entangled states, due to the fact that amplitudes of different input states can be coherently added after they're calculated. For example the input state

$$|\Psi\rangle = \frac{|1,0,\cdots,0\rangle + |0,\cdots,0,1\rangle}{\sqrt{2}}, \tag{2.2-95}$$

can be decomposed as

$$|\Psi\rangle = a_1 |T_1\rangle + a_2 |T_2\rangle \tag{2.2-96}$$

The transition amplitudes can be calculated for each term separately:

$$\langle S_i|\varphi(U)|T_j\rangle = \alpha_{ij} = \frac{\mathrm{Per}(U_{S_i,T_j})}{\sqrt{s_1!\cdots s_n! t_1!\cdots t_n!}}, \tag{2.2-97}$$

and the transition amplitude of the full input state is then simply recovered through the linearity of the inner product:

$$\langle S_i|\varphi(U)|T\rangle = a_1\langle S_i|\varphi(U)|T_1\rangle + a_2\langle S_i|\varphi(U)|T_2\rangle = \sum_{ij} \alpha_{ij} a_j, \tag{2.2-98}$$

$$\varphi(U)\,|\Psi\rangle = \sum_{ij} \alpha_{ij} a_j\,|S_i\rangle \tag{2.2-99}$$

Similarly, for mixed states

$$\rho = \sum_i p_i\,|T_i\rangle\langle T_i|\,, \tag{2.2-100}$$

one can calculate the transition amplitudes for each state vector $|T_i\rangle$ and add their squared norms, giving the output mixture

$$\rho \mapsto \sum_{ij} p_j |\langle S_i|\varphi(U)|T_j\rangle|^2\,|S_i\rangle\langle S_i|\,. \tag{2.2-101}$$

The discussion in this section has implicitly assumed that all the photons input into the linear optical circuit are fully indistinguishable in every degree of freedom except for the explicitly included ones, represented by the indexed modes. In practice this assumption is rarely fully justified, as independently produced photons tend to exhibit small variations in various degrees of freedom, for example their frequency. One way to include the effects of partial indistinguishability is to explicitly include more degrees of freedom [123]. For example, in the case of two photons in two modes one can include an additional two modes per photon that model distinct internal degrees of freedom of the two photons. If these modes do not couple, then photons in the auxiliary modes never

interfere, and by choosing an input state that is a superposition over the original and auxiliary modes the degree to which the two photons interfere can be controlled.

As the number of photons increases, however, the number of these auxiliary modes quickly grows, as one needs to consider every possible combination of interfering and non-interfering input photons. This motivates the development of more efficient simulation techniques that do not explicitly consider these degrees of freedom, and a significant amount of literature exists on this subject [124, 125, 126, 127]. One drawback that all methods operating with a reduced dimensionality suffer from is the inability to simulate entangled input states. This is a consequence of the fact that these methods produce the output probabilities, rather than the output amplitudes. This, in turn, is necessary because a given Fock-state vector specifying the number of photons in each mode is not a well defined quantity in the case where the photons are partially distinguishable. Instead, these methods effectively compute the partial trace over the hypothetical auxiliary modes, and the corresponding output states are then no longer pure.

2.2.9 Hong-Ou-Mandel interference

Armed with the knowledge of how to model linear-optical systems we can now consider a simple scenario of two indistinguishable photons in two different spatial modes, both incident on the same balanced beamsplitter:

$$|\Psi\rangle = \hat{a}^\dagger \hat{b}^\dagger |0\rangle = |1,1\rangle . \tag{2.2-102}$$

The beamsplitter transforms the creation operators as

$$\hat{a}^\dagger \mapsto \frac{1}{\sqrt{2}}(\hat{a}^\dagger + \hat{b}^\dagger), \qquad \hat{b}^\dagger \mapsto \frac{1}{\sqrt{2}}(\hat{a}^\dagger - \hat{b}^\dagger), \tag{2.2-103}$$

and the state after the beamsplitter can therefore be written as

$$\begin{aligned} |\Psi\rangle &= \frac{1}{2}(\hat{a}^\dagger + \hat{b}^\dagger)(\hat{a}^\dagger - \hat{b}^\dagger)\,|0\rangle \\ &= \frac{1}{2}(\hat{a}^\dagger\hat{a}^\dagger - \hat{a}^\dagger\hat{b}^\dagger + \hat{b}^\dagger\hat{a}^\dagger - \hat{b}^\dagger\hat{b}^\dagger\,|0\rangle \\ &= \frac{1}{2}(\hat{a}^\dagger\hat{a}^\dagger - \hat{b}^\dagger\hat{b}^\dagger)\,|0\rangle \\ &= \frac{|2,0\rangle + |0,2\rangle}{\sqrt{2}}. \end{aligned} \tag{2.2-104}$$

This simple calculation suggests that two indistinguishable photons always bunch on a beamsplitter. This effect, pictured in Fig. 2.6, is called Hong-Ou-Mandel (HOM) interference [128,129]. The bunching can be seen as a reflection of the bosonic nature of photons. If one were to perform the same experiment with fermionic particles, the opposite behaviour would be seen. Indeed, the anti-symmetric state $|\Psi^-\rangle = (|H,V\rangle - |V,H\rangle)/\sqrt{2}$ will always anti-bunch on a beamsplitter, and more generally the beamsplitter acts as a symmetry sorter through which all symmetric two-qubit states bunch. Hong-Ou-Mandel interference is a purely non-classical effect and cannot be observed with thermal or coherent states. It is worth noting that HOM-interference does not depend on the phase

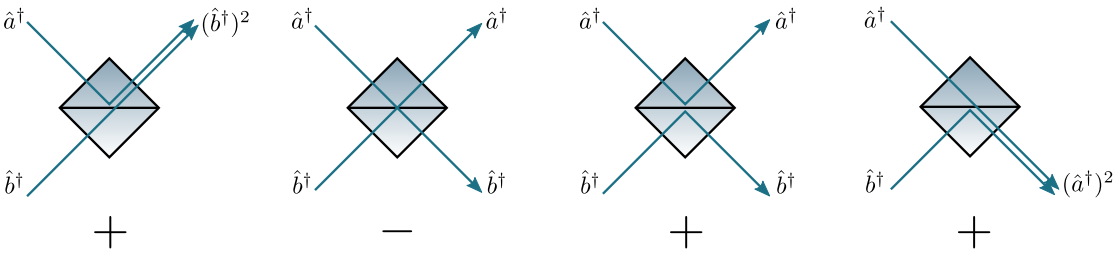


Figure 2.6: **Hong-Ou-Mandel interference.** Visual representation of the four amplitudes after the beamsplitter, corresponding to all the ways that the two photons can be reflected and transmitted. If the two photons are indistinguishable then the amplitude corresponding to both of them being reflected interferes destructively with the amplitude for both of them being transmitted. This in turn leads to only the first and last amplitudes remaining, and consequently the photons always bunch.

of the photons, since any phase in (2.2-102) would simply be a global phase. While HOM-interference is not a true two-photon interaction, it is nevertheless a critical ingredient in photonic quantum information processing and underlies all of the linear-optical two-photon operations discussed in Section 2.3.3. The interference between the two amplitudes corresponding to both photons being reflected or transmitted relies on these two outcomes being indistinguishable, which in turn requires the photons to be indistinguishable. For example, if the two photons have orthogonal polarizations then the second line in (2.2-104) instead reads:

$$\frac{1}{2}(\hat{a}_H^\dagger \hat{a}_V^\dagger - \hat{a}_H^\dagger \hat{b}_V^\dagger + \hat{b}_H^\dagger \hat{a}_V^\dagger + \hat{b}_H^\dagger \hat{b}_V^\dagger \,|0\rangle\,, \tag{2.2-105}$$

and no cancellation occurs. Ensuring a high degree of indistinguishablility between photons is therefore of great importance in photonic quantum information. In practice, photons produced by a single-photon source exhibit some partial degree of distinguishability in one or more degrees of freedom. In practice, photons are also localized to within their coherence length, and need to arrive at the beamsplitter with a time difference much smaller than this coherence length. When scanning the relative time delay between two photons the photons continuously transition from being distinguishable to being indistinguishable in time. By plotting the rate of coincidence detection events between the two output ports of the beamsplitter one gets what is known as a Hong-Ou-Mandel dip, and the visibility of this curve corresponds to the indistinguishability of the photons in the remaining degrees of freedom. More realistic models of HOM-interference will be discussed in Section 2.4.5.

2.3 Single-photon quantum information

Within the field of photonic quantum information there are two distinct approaches to encoding the actual quantum information [130]. One utilizes the continuous degrees of freedom of the electric field quadratures, often referred to as a continuous variable (CV) encoding [26, 114], while the second approach uses discrete degrees of freedom of single photon states, and this is called a discrete variable (DV) encoding [131,132]. Both

approaches have their own advantages and disadvantages. The former has the strength that a simple balanced beamsplitter can act as a deterministic maximally entangling operation, and this simplicity has enabled the experimental realisation of exceptionally large entangled states [133,134,135]. While CV quantum information tries to exploit the intrinsically bosonic nature of the photon, this approach suffers from the problem that one almost always wants to encode discrete information in the system. This is possible through the use of so-called GKP states [136], which are states with periodic Wigner functions that mimic qubit states. These states, however, are unphysical, requiring infinite energy, and while finite energy approximations of GKP states can be shown to be a useful computational resource [137], they nevertheless remain challenging to produce and have not been realised in a photonic system to date.

The discrete variable encoding takes the opposite approach, ignoring the bosonic nature of the photons by directly constructing qubits out of them. This, however, presents a new problem, namely that manipulations of photons do not necessarily preserve the qubit states. For example, two photons impinging on different ports of a beamsplitter might exit in the same port, and the resulting state is typically not a qubit state. An additional challenge is that of realising two-qubit gates, since photons do not interact. In the CV approach the entanglement that can be generated deterministically is photon-number entanglement. While maximally entangled states in the photon number can in some cases be deterministically generated in the discrete variable paradigm as well [138], the photon number turns out to be an impractical degree of freedom to work with. For example, a measurement in a diagonal photon-number basis cannot conserve photon number, and consequently it cannot be performed using only linear optics [10].

In practice the most commonly used discrete degrees of freedom of single photons are the polarization [96, 139, 140], path [141, 142, 143], time [144, 145] and oribital angular momentum (OAM) [146, 147], though more encodings exist [148, 149]. Several degrees of freedom can be combined to encode higher-dimensional states in a single photon, and these are often easier to entangle [150, 151, 152, 153]. For instance, a simple polarizing beamsplitter does a controlled-NOT operation between the polarization and path, and therefore acts as an entangling operation between the two degrees of freedom. An OAM encoding offers the advantage of an infinite-dimensional Hilbert space that is purely internal to the photon [154], and this has enabled the realisation of several highly entangled high-dimensional states [155, 156, 157]. However, a major drawback is that OAM states are comparatively fragile, and are not preserved in standard optical fibers. Consequently the first three encodings are by far the most common in practical applications, with polarization and time encodings seeing significant use in quantum communication technologies, while path encodings are currently being pursued in photonic quantum computation due to its inherent compatibility with integrated photonic devices [158, 159].

The experiments presented in this thesis primarily use polarization encoded states, with occasional uses of the path and time degrees of freedom. Photon polarization is a particularly appealing degree of freedom due to its high robustness to noise and ease of manipulation. The robustness of the encoding comes from a fact that the polarization basis states can only be rotated, or acquire relative phases, by propagating through birefringent media. Many optical components exhibit small or negligible birefringence, and

components that are birefringent do not display a significant time-dependence, meaning that their influence can be pre-corrected. This is of course only approximately true, as there are effects which induce decoherence in polarization states as well. For example, propagation through sufficiently long optical fibers will introduce a stochastic time dependent polarization rotation that acts as depolarizing noise, strongly birefringent materials introduce temporal walk-off that can decohere a polarization state when the walk-off is on the order of the coherence time of the light, and very broadband photons can experience decoherence arising from wavelength dependent birefringence.

2.3.1 Manipulating polarization qubits

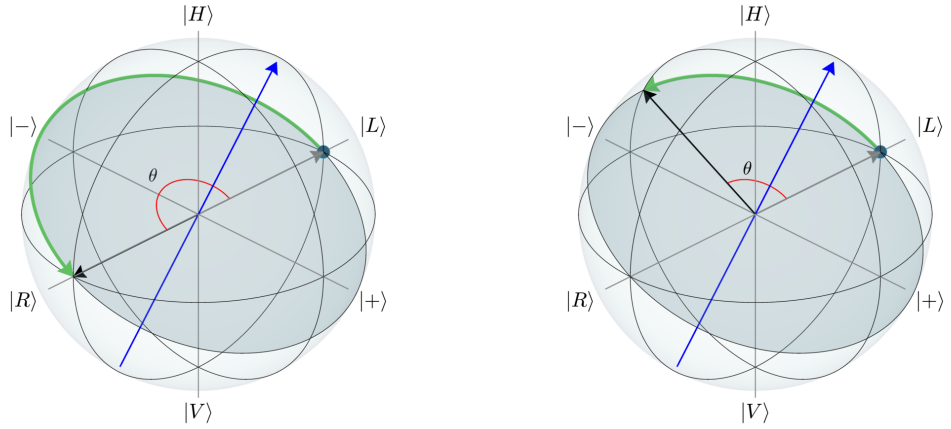
As discussed in Section 1.1.5 a general qubit rotation can be decomposed as three rotations about some fixed axes on the Bloch sphere. For a polarization qubit, such a rotation is realised by propagating through a birefringent material, where the axis of rotation is set by the orientation of the optical axis of the medium, and the rotation angle is set by the total birefringent phase shift acquired by propagating through the material. In practice there are therefore three different approaches for manipulating polarization qubits: one can either use devices with a tunable birefringence but with fixed optical axes, devices with fixed birefringence but whose optical axes can be rotated, or a combination of the two. Materials with tunable birefringence can be realised using for example the Pockels effect, or using liquid crystals. The approach using elements with fixed birefringence often turns out to be the most practical and cost-effective one. Abundantly available materials such as quartz exhibit linear birefrigence, and can be manufactured with small tolerances.

Since qubit rotations are typically parameterised in terms of Euler- or Tait-Bryan-rotations, which map to tunable birefringence, it is not entirely obvious how one would generate arbitrary rotations from elements with fixed birefringence, or how to optimally pick these optical elements. In a series of papers in the late 1980's, R. Simon and N. Mukunda proved that there exists a three-component gadget consisting of fixed linear retarders that is both optimal and universal for SU(2) [160, 161]. The device uses two linear retarders with a total birefringence of $\pi/2$, called quarter-wave plates, and one linear retarder with a phase-delay of π, called a half-wave plate. These components can be described in matrix form as[2]

$$\begin{aligned} H(\theta) = U_{\mathrm{HWP}}(\theta) &= e^{-i\pi/2} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \\ &= e^{-i\pi/2} \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \end{aligned} \tag{2.3-1}$$

$$\begin{aligned} Q(\theta) = U_{\mathrm{QWP}}(\theta) &= e^{-i\pi/4} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \\ &= e^{-i\pi/4} \begin{bmatrix} \cos^2\theta + i\sin^2\theta & (1-i)\cos\theta\sin\theta \\ (1-i)\cos\theta\sin\theta & i\cos^2\theta + \sin^2\theta \end{bmatrix}, \end{aligned} \tag{2.3-2}$$

[2]The exact form of these matrices depends on the convention used to represent the polarization states. The matrices here use the convention defined in (2.3-3).



(a) HWP (b) QWP

Figure 2.7: **Visual representation of the action of a HWP and a QWP.** The two figures show examples of how a half-wave plate (HWP) and quarter-wave plate (QWP) rotate a state on the Bloch sphere. In these examples, the initial state is a left-handed circularly polarized state and the rotation axis is parallel to the vector [1 0 2]. The HWP in (a) rotates the initial state $|L\rangle$, indicated by a faint arrow, by 180° to $|R\rangle$, while the QWP in (b) rotates $|L\rangle$ by 90° such that it ends up in the plane spanned by the linear polarizations. This happens independently of the rotation axis and therefore the physical orientation of the wave plate, although the exact linear state that the QWP rotates $|L\rangle$ to does depend on the rotation axis.

where θ is the angle between the horizontal axis and the fast axis of the wave plate.

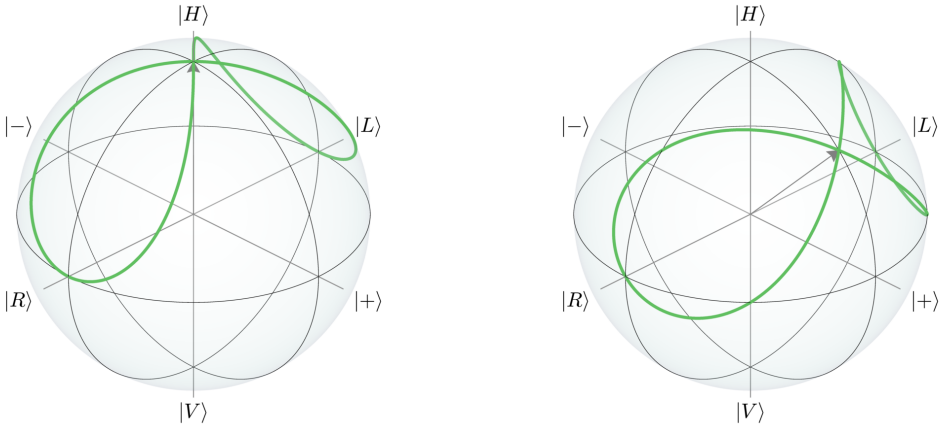
In their paper, Simon and Mukunda used algebraic methods to relate the angles of the wave plates to an Euler-angle decomposition of an arbitrary rotation [161]. This approach offers little physical insight, and here a visual method for determining the correct wave-plate angles will be presented. The following convention for the polarization states and Pauli matrices will be used:

$$\begin{aligned} |0\rangle &= |H\rangle & |+\rangle &= \frac{|H\rangle + |V\rangle}{\sqrt{2}} & |L\rangle &= \frac{|H\rangle + i\,|V\rangle}{\sqrt{2}} \\ |1\rangle &= |V\rangle & |-\rangle &= \frac{|H\rangle - |V\rangle}{\sqrt{2}} & |R\rangle &= \frac{|H\rangle - i\,|V\rangle}{\sqrt{2}} \end{aligned} \tag{2.3-3}$$

$$\begin{aligned} Z\,|H\rangle &= |H\rangle & X\,|+\rangle &= |+\rangle & Y\,|L\rangle &= |L\rangle \\ Z\,|V\rangle &= -\,|V\rangle & X\,|-\rangle &= -\,|-\rangle & Y\,|R\rangle &= -\,|R\rangle\,. \end{aligned} \tag{2.3-4}$$

In other words, the horizontal and vertical polarization states are the computational basis states, diagonal polarization states are real superpositions of the former and the circular polarization states are complex superpositions of the horizontal and vertical polarizations. In this convention, half- and quarter-wave plates are rotations around an axis that lies in the $X-Y$-plane of the Bloch sphere, with aotation angles of 180° and 90° respectively, and the direction of the axis is determined by the physical orientation of the optical axis. In other words, they are rotations of the form

$$U_{\mathrm{WP}} = \cos\frac{\theta}{2} I - i \sin\frac{\theta}{2}(n_x X + n_z Z), \qquad n_x^2 + n_z^2 = 1 \tag{2.3-5}$$



(a) Initial state $|H\rangle$ (b) Initial state $U_{\mathrm{HWP}}(12.25°)\,|H\rangle$

Figure 2.8: **Image of a QWP for two linearly polarized states.** The green curves show the image of a QWP, i.e. the curve traced out as the angle of the wave plate is rotated from 0° to 180°, for two different initial states in the plane of linear polarization. Consequently these curves contain all the possible final states that can be reached from the respective initial states using a single quarter-wave plate. Equivalently, a QWP can bring any state on the green curves back to the initial state in each figure. Rotating the initial state in the linear plane has the effect of rotating these green curves about the axis defined by the circular polarizations. It is therefore clear that by choosing an appropriate initial state in the linear plane, one can find a curve that intersects any point on the Bloch sphere.

with

$$\theta_{\mathrm{HWP}} = \pi, \qquad \theta_{\mathrm{QWP}} = \frac{\pi}{2}. \tag{2.3-6}$$

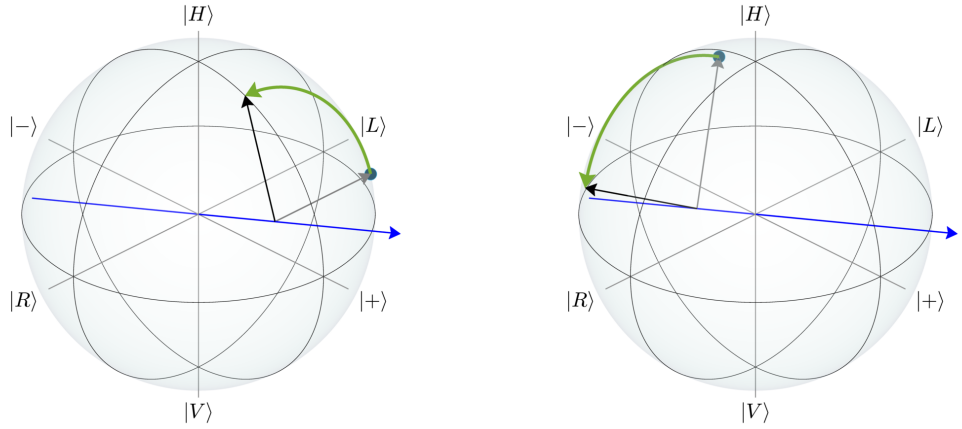
This is visually illustrated in Fig. 2.7. We now turn to the problem of realising general qubit rotations using wave plates. Consider an arbitrary unitary $U \in \mathrm{SU}(2)$. The problem of implementing U can equivalently be thought of as implementing the inverse of a unitary $V = U^\dagger$. Suppose V acts on the states $|H\rangle$ and $|L\rangle$ as

$$V\,|H\rangle = |\Psi\rangle\,, \qquad V\,|L\rangle = |\Phi\rangle\,. \tag{2.3-7}$$

Implementing U is then equivalent to finding some unitary that maps

$$|\Psi\rangle \to |H\rangle\,, \qquad |\Phi\rangle \to |L\rangle\,. \tag{2.3-8}$$

To this end, it is useful to first note two facts about wave plates: firstly, that a HWP always maps circularly polarized states to circularly polarized states with the opposite handedness (see Fig. 2.7b), and secondly that a QWP can take any state on the Bloch sphere and map it to the linear plane. The second property is less obvious, but can be understood by thinking of the inverse process, that is, a QWP acting on a state in the linear plane. This is illustrated in Fig. 2.8, which shows all the possible states that a QWP can map a given linearly polarized state to. Two identical linear retarders can always be made to implement an identity transformation by orienting them at 90° to each other, since this configuration ensures that there is no net birefringence. It follows



(a) Initial state $V\,|L\rangle$ (b) Initial state $V\,|H\rangle$

Figure 2.9: **First QWP of a universal three-component gadget.** The first step of implementing an arbitrary unitary U is to map the state $|\Phi\rangle = U^\dagger\,|L\rangle = V\,|L\rangle$ to the linear plane using a quarter-wave plate. This is shown in (a) where the grey arrow indicates the initial state $|\Psi\rangle = V\,|L\rangle$ and the solid arrow shows the resulting state in the linear plane. (b) shows the seemingly arbitrary action of the same QWP on $U^\dagger\,|H\rangle$. As for all linear retarders, the rotation axis, indicated by the blue arrow, lies in the plane of linear polarization.

that:

$$Q(\theta + 90°) = Q(\theta)^{-1}. \tag{2.3-9}$$

Therefore the curves in Fig. 2.8 show all the initial states which a QWP can map back to a given linearly polarized state. It is easy to see visually that the set of all such curves for all linearly polarized states covers the whole Bloch sphere. Therefore, given an arbitrary state on the Bloch sphere, there is always some initial linear state which a QWP can map to the desired state, and consequently a QWP can bring any state back to the linear plane. Returning to the problem of implementing the map in (2.3-8) using wave plates, one begins by using a QWP to map the state $|\Phi\rangle$ to some linearly polarized state $|\Phi_\mathrm{L}\rangle$:

$$Q(\theta)\,|\Phi\rangle = |\Phi_\mathrm{L}\rangle = a\,|H\rangle + b\,|V\rangle\,, \quad a, b \in \mathbb{R}. \tag{2.3-10}$$

This is illustrated in Fig. 2.9. In Appendix A it is shown that the angle θ is given by

$$\theta = \frac{1}{2}\mathrm{atan2}\big(\mathrm{Tr}[X\,|\Phi\rangle\langle\Phi|], \mathrm{Tr}[Z\,|\Phi\rangle\langle\Phi|]\big). \tag{2.3-11}$$

As illustrated in Fig. 2.10, a second QWP can be used to map $|\Phi_\mathrm{L}\rangle$ to $|R\rangle$:

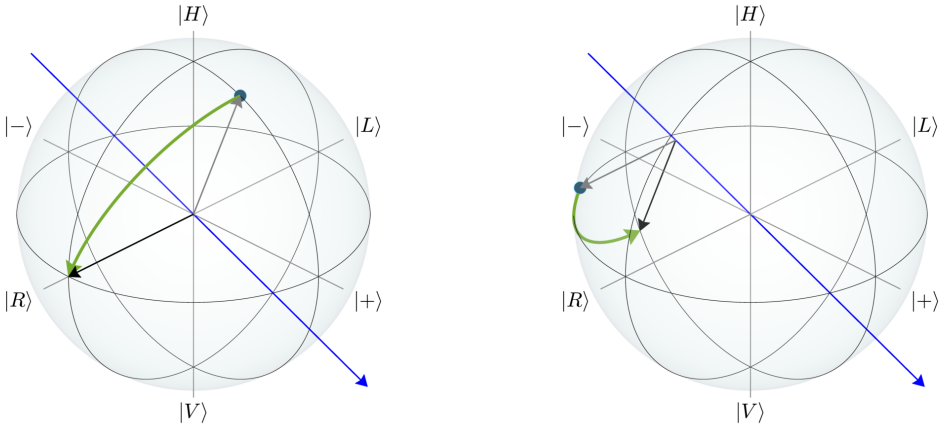
$$Q(\varphi)\,|\Phi_\mathrm{L}\rangle = |R\rangle\,, \tag{2.3-12}$$

where the angle φ can be found as

$$\varphi = \frac{1}{2}\mathrm{atan2}\big(\mathrm{Tr}[X\,|\Phi_\mathrm{L}\rangle\langle\Phi_\mathrm{L}|], \mathrm{Tr}[Z\,|\Phi_\mathrm{L}\rangle\langle\Phi_\mathrm{L}|]\big) + 45°. \tag{2.3-13}$$

Since the states

$$|\Psi\rangle = V\,|H\rangle\,, \qquad |\Phi\rangle = V\,|L\rangle\,. \tag{2.3-14}$$



(a) Initial state $Q(\theta)V\,|L\rangle$ (b) Initial state $Q(\theta)V\,|H\rangle$

Figure 2.10: **Second QWP of a universal three-component gadget.** The second quarter-wave plate in the universal SU(2) three-component polarization gadget maps the states in the figure to linear and circular polarizations. (a) The QWP angle is chosen such that the state $Q(\theta)V\,|L\rangle$, which lies in the linear plane and should ultimately get mapped to $|L\rangle$, gets mapped to $|R\rangle$. (b) This choice of QWP angle simultaneously ensures that $Q(\theta)V\,|H\rangle$ gets mapped to the plane of linear polarization.

are orthogonal on the Bloch sphere, and

$$Q(\varphi)Q(\theta)\,|\Phi\rangle = |R\rangle\,, \tag{2.3-15}$$

it follows that

$$Q(\varphi)Q(\theta)\,|\Psi\rangle = |\Psi_{\mathrm{L}}\rangle\,, \tag{2.3-16}$$

is a linearly polarized state. The transformation in (2.3-8) can therefore be completed using a HWP:

$$H(\gamma)Q(\varphi)Q(\theta)\,|\Phi\rangle = H(\gamma)\,|R\rangle = |L\rangle \tag{2.3-17}$$
$$H(\gamma)Q(\varphi)Q(\theta)\,|\Psi\rangle = H(\gamma)\,|\Psi_{\mathrm{L}}\rangle = |H\rangle\,. \tag{2.3-18}$$

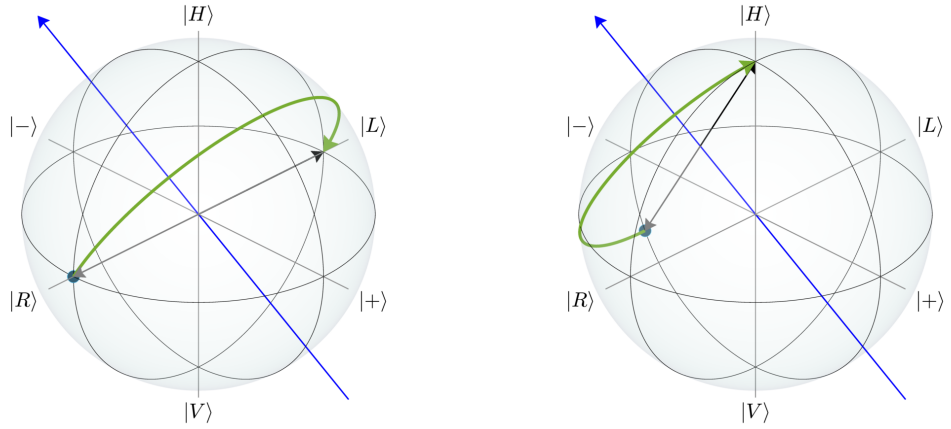
This last rotation is shown in Fig. 2.11. Since the HWP will map $|R\rangle$ to $|L\rangle$ independently of its angle, the angle only depends on the linearly polarized state $|\Psi_{\mathrm{L}}\rangle$ and can be expressed as

$$\gamma = \frac{1}{4}\mathrm{atan2}\big(\mathrm{Tr}[X\,|\Psi_{\mathrm{L}}\rangle\langle\Psi_{\mathrm{L}}|], \mathrm{Tr}[Z\,|\Psi_{\mathrm{L}}\rangle\langle\Psi_{\mathrm{L}}|]\big). \tag{2.3-19}$$

With these angles the three wave plates realise the desired transformation:

$$H(\gamma)Q(\varphi)Q(\theta) = U = V^{\dagger}, \tag{2.3-20}$$

up to an irrelevant global corresponding to a 2π rotation on the Bloch sphere. The order of the three wave plates does not affect the universality of the gadget, however a re-ordering of the wave plates leads to a different set of angles. These can be calculated



(a) Initial state $Q(\varphi)Q(\theta)V\,|L\rangle$ (b) Initial state $Q(\varphi)Q(\theta)V\,|H\rangle$

Figure 2.11: **HWP in a universal three-component gadget.** The action of the HWP in the Simon Mukunda gadget, which acts last, after the two QWPs. This wave plate has the effect of (a) mapping $Q(\varphi)Q(\theta)V\,|L\rangle = |R\rangle$ to $|L\rangle$, and (b) mapping $Q(\varphi)Q(\theta)V\,|H\rangle$ to $|H\rangle$, thereby completing the desired transformation.

using the following permutation identities:

$$Q(\alpha)H(\beta) = H(\beta)Q(2\beta - \alpha) \tag{2.3-21}$$
$$H(\beta)Q(\alpha) = Q(2\beta - \alpha)H(\beta). \tag{2.3-22}$$

2.3.2 Measuring polarization qubits

The ease of manipulating polarization qubits also carries over into measuring them. While some photo-detection processes exhibit a polarization dependence, the most straightforward way to measure a polarization state is with a Von Neumann measurement. This is done by coupling the polarization basis states to different spatial modes, since measuring which spatial mode the photon occupies then also reveals its polarization [46]. As mentioned in Section 2.2.7, such a Von Neumann measurement can be implemented using a polarizing beamsplitter (PBS), which typically transmits horizontally polarized light and reflects vertically polarized light (see Fig. 2.12). A bare PBS therefore effects a measurement in the computational basis, corresponding to the observable Z. Detecting a photon in the transmitted port of the PBS reveals the eigenvalue $+1$, and detecting one in the reflected port reveals -1. Measurements of other single-qubit observables can be implemented by performing single-qubit rotations before the PBS. Given an observable A with eigenstates $|a_\pm\rangle$ and corresponding eigenvalues ± 1 one needs to perform an operation U such that

$$U\,|a_+\rangle = |H\rangle\,, \qquad U\,|a_-\rangle = |V\rangle\,. \tag{2.3-23}$$

In other words, the eignstate with the positive eigenvalue should be mapped to $|H\rangle$, and the state with the negative eigenvalue should get mapped to $|V\rangle$. The logic behind this is simple: a Von Neumann measurement of the observable A is one where the positive

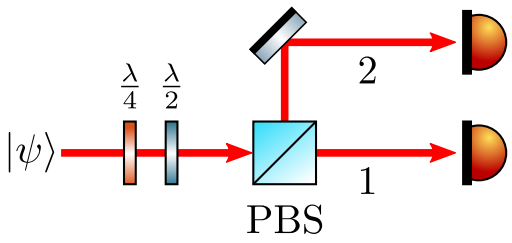


Figure 2.12: **Polarization qubit measurement.** A Von Neumann measurement of a polarization qubit can be realised using a polarizing beamsplitter that reflects vertically polarized light and transmits horizontally polarized light. This correlates the polarization and path of the photon: $|\psi\rangle = \alpha\,|H\rangle_1 + \beta\,|V\rangle_1 \mapsto \alpha\,|H\rangle_1\,|0\rangle_2 + \beta\,|0\rangle_1\,|V\rangle_2$. A measurement of the photon position then reveals the polarization state. Using a quarter-wave and half-wave plate before the PBS any qubit measurement basis can be selected.

eigenstate $|a_+\rangle$ is always transmitted through the PBS, since this is the path associated with the positive eigenvalue, and vice versa for $|a_-\rangle$.

The problem of mapping an arbitrary state to $|H\rangle$ is simpler than performing a general qubit rotation, since the relative phase of the basis states does not matter. It is therefore sufficient to use only a single quarter-wave plate and one half-wave plate together with a PBS to realise any qubit observable. In analogy with the method for general unitaries presented in the previous section, a QWP first rotates the state $|a_+\rangle$ to the linear plane:

$$Q(\theta)\,|a_+\rangle = \alpha\,|H\rangle + \beta\,|V\rangle\,, \quad \alpha, \beta \in \mathbb{R}, \tag{2.3-24}$$

and a half-wave plate rotates the resulting state in the linear plane to $|H\rangle$:

$$H(\varphi)Q(\theta)\,|a_+\rangle = |H\rangle\,. \tag{2.3-25}$$

The angles θ and φ can be found using the methods outlined in previous section. It is worth noting that for some observables these angles are not unique, even when disregarding the natural periodicity of wave-plate rotations. For implementing the commonly used Pauli observables one set of possible angles is shown in Table 2.1.

Although generalized measurements remain more challenging and significantly less widely used than straightforward projective measurements, photons also admit relatively simple implementations of single-qubit POVMs in comparison to other physical systems. This is due to the fact that they are easy to place in spatial superpositions, and an appropriate superposition over two or more modes can be used to realise aPOVM. The photonic realisation of a POVM can be understood using a simple constructive method [162]. Firstly, note that it suffices to consider general implementations of rank-1 POVMs, as any higher rank POVM can be realised using a rank-1 POVM with more outcomes. Then consider a general rank-1 POVM M with POVM elements

$$E_m = p_m\,|\Psi_m\rangle\langle\Psi_m|\,. \tag{2.3-26}$$

Let $|\Psi_m^\perp\rangle$ be the state orthogonal to $|\Psi_m\rangle$:

$$\langle\Psi_m^\perp|\Psi_m\rangle = 0. \tag{2.3-27}$$

Wave-plate angles		
Observable	QWP angle	HWP angle
X	45°	22.5°
-X	45°	67.5°
Y	0°	67.5°
-Y	0°	22.5°
Z	0°	0°
-Z	0°	45°

Table 2.1: **Qubit-tomography wave-plate angles.** Single-qubit tomography can be performed using six (three) different wave-plate settings when using a polarizer (PBS). When using a PBS the positive and negative eigenvalues of the observables are measured in parallel. The table shows one possible set of wave plate angles to implement the three Pauli observables.

Then clearly

$$\mathrm{Tr}[E_m \left|\Psi_m^{\perp}\right\rangle\!\left\langle\Psi_m^{\perp}\right|] = 0. \tag{2.3-28}$$

In other words, given that the state being measured is $|\Psi_m^{\perp}\rangle$ the outcome m will never be observed. This is a sufficient condition for the realisation of a POVM and motivates a simple algorithm for its implementation: consider a circuit with as many modes as the N outcomes of the POVM. Suppose the state $|\Psi_1^{\perp}\rangle$ is input into the circuit in modes $N-1$ and N. Then perform a local operation U_1 that maps this state to mode N:

$$|\Psi_1^{\perp}\rangle = a_1 \left|N-1\right\rangle + b_1 \left|N\right\rangle, \qquad U_1 \left|\Psi_1^{\perp}\right\rangle = \left|N\right\rangle, \tag{2.3-29}$$

where $|N\rangle = \hat{a}_N^{\dagger} |0\rangle$ represents a single photon in the N-th mode. This choice of U_1 also ensures that

$$U_1 \left|\Psi_1\right\rangle = \left|N-1\right\rangle. \tag{2.3-30}$$

Next, apply a unitary V_1 that couples modes $N-1$ and 1 with a strength given by p_1:

$$V_1 \left|N-1\right\rangle = \sqrt{p_1}\left|1\right\rangle + \sqrt{1-p_1}\left|N-1\right\rangle. \tag{2.3-31}$$

Associating a photo-detection event in mode 1 with the outcome $m = 1$ then ensures that the following conditions for the POVM are satisfied:

$$\mathrm{Tr}[E_1 \left|\Psi_1\right\rangle\!\left\langle\Psi_1\right|] = p_1, \qquad \mathrm{Tr}[E_1 \left|\Psi_1^{\perp}\right\rangle\!\left\langle\Psi_1^{\perp}\right|] = 0. \tag{2.3-32}$$

For the second POVM outcome, one considers projections of the propagated states

$$V_1 U_1 \left|\Psi_2^{\perp}\right\rangle, \tag{2.3-33}$$

$$V_1 U_1 \left|\Psi_2\right\rangle, \tag{2.3-34}$$

onto the modes $N-1$ and N:

$$(\langle N-1| + \langle N|) V_1 U_1 \left|\Psi_2^{\perp}\right\rangle = \left|\Psi_{2,\mathrm{proj}}^{\perp}\right\rangle, \tag{2.3-35}$$

$$(\langle N-1| + \langle N|) V_1 U_1 \left|\Psi_2\right\rangle = \left|\Psi_{2,\mathrm{proj}}\right\rangle \tag{2.3-36}$$

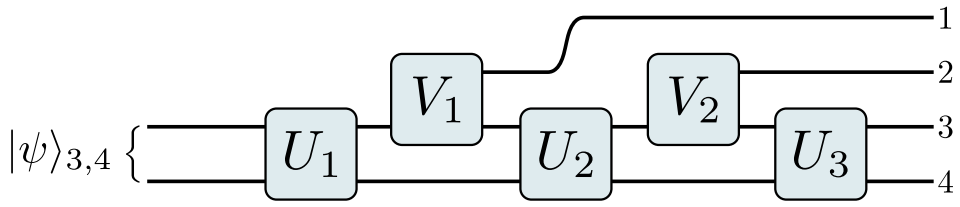


Figure 2.13: **Four-outcome POVM on a path qubit.** A four-outcome POVM can be realised using only the path degree of freedom of a single photon. As described in the main text, the unitary U_1 ensures that the state $|\Psi_1^\perp\rangle$ is never detected in mode 1, which is equivalent to the corresponding POVM element having the correct orientation on the Bloch sphere. The unitary V_1 couples the modes 1 and 3 in a way that ensures $\mathrm{Tr}[E_1 |\Psi_1\rangle\langle\Psi_1|] = p_1$, thereby giving the correct magnitude of the POVM element E_1.

and performs the same procedure, mapping

$$U_2 |\Psi_{2,\mathrm{proj}}^\perp\rangle \propto |N\rangle\,, \tag{2.3-37}$$

and

$$\||\Psi_{2,\mathrm{proj}}\rangle\|^2 V_2 |N-1\rangle = \sqrt{p_2}\,|2\rangle + \sqrt{\||\Psi_{2,\mathrm{proj}}\rangle\|^2 - p_2}\,|N-1\rangle\,. \tag{2.3-38}$$

In general U_k is a two-mode unitary that satisfies

$$U_k \prod_{m=0}^{k-1} V_m U_m P |\Psi_k^\perp\rangle \propto |N\rangle\,, \quad k = 1 \ldots N-1 \tag{2.3-39}$$

$$U_0 = \mathbb{1}, \tag{2.3-40}$$

where

$$\||\Psi_{k,\mathrm{proj}}\rangle\|^2 V_k |N-1\rangle = \sqrt{p_k}\,|k\rangle + \sqrt{\||\Psi_{k,\mathrm{proj}}\rangle\|^2 - p_k}\,|N-1\rangle\,, \quad k \geq 1 \tag{2.3-41}$$

$$|\Psi_{k,\mathrm{proj}}\rangle = P \prod_{m=0}^{k-1} V_m U_m P |\Psi_k\rangle \tag{2.3-42}$$

$$V_0 = \mathbb{1}, \tag{2.3-43}$$

and

$$P = \langle N-1| + \langle N|\,. \tag{2.3-44}$$

Note that $N-1$ unitaries U_k and $N-2$ unitaries V_k suffice due to the completeness of the POVM elements. A circuit implementing a four outcome POVM obtained through this constructive method is illustrated for a path-encoded qubit in Fig. 2.13, as this encoding is easy to understand. Polarization qubits can be measured by such a circuit by placing a PBS at its input, thereby mapping the polarization state to a path-encoded state. However, a POVM on a polarization qubit can also be realised in a more direct way, with a combination of path and polarization modes. In such a scheme, the unitaries U_m and V_m act on the polarization degree of freedom, and control the coupling to the spatial modes. Typically this is done using beam displacers, which are birefringent pieces of glass that physically displace one polarization component with respect to another one,

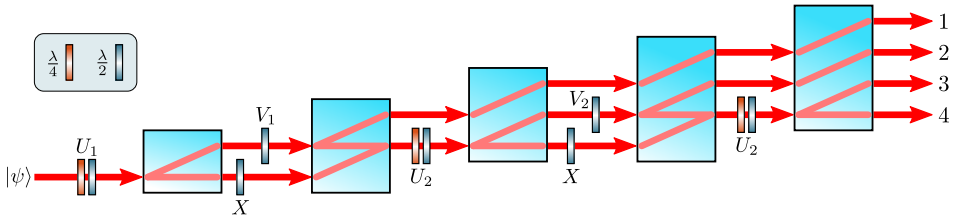


Figure 2.14: **Four-outcome POVM on a polarization qubit.** Using beam-displacers and wave plates POVMs can be directly realized on a polarization qubit by coupling it to multiple paths. Similarly to the path implementation, the unitary U_k ensures that the state $|\Psi_k^{\perp}\rangle$ gets mapped to $|H\rangle$, and is not displaced by the beam displacer. The unitaries V_k, in conjunction with the beam displacers realize the tunable coupling to the path degree of freedom. The X gates are necessary when using beam displacers that all displace in the same direction, and this configuration ensures there is no decoherence due to temporal walk-off.

but leave the two propagating in parallel directions. However, it can also be performed using tunable partially-polarizing beamsplitters.

In Fig. 2.14 a path-polarization realisation of the unitary V_1U_1 is shown. Initially the photon is in a polarization state

$$|\Psi_1^{\perp}\rangle = a_1 |H\rangle_1 + b_1 |V\rangle_1 , \tag{2.3-45}$$

where the subscript denotes the spatial mode. A set of wave plates implements the polarization unitary U_1 that maps:

$$U_1 |\Psi_1^{\perp}\rangle = |H\rangle_1 . \tag{2.3-46}$$

The above state is entirely transmitted through the subsequent beam-displacer:

$$|H\rangle_1 \mapsto |H\rangle_2 , \tag{2.3-47}$$

after which a second polarization unitary effects the transformation

$$V_1 |V\rangle_1 = \sqrt{p_1} |V\rangle_1 + \sqrt{1-p_1} |H\rangle_1 . \tag{2.3-48}$$

Finally, the second beam-displacer couples the polarization state above to two different paths:

$$\sqrt{p_1} |V\rangle_1 + \sqrt{1-p_1} |H\rangle_1 \mapsto \sqrt{p_1} |V\rangle_1 + \sqrt{1-p_1} |H\rangle_2 . \tag{2.3-49}$$

This approach has been utilized in several experiments [163,164,165], however it presents two challenges. The first is that the lateral displacement produced by birefringent crystals is typically limited to a few mm, due to low birefringence and constraints on the crystal size. This small beam separation complicates the implementation of the polarization rotation V_m that acts on only one of the two spatial modes. A second challenge is that the two beam displacers act as a Mach-Zehnder interferometer, and while the parallel-path geometry is phase insensitive to displacements of either crystal, the interferometer does nevertheless not have full passive stability. An alternative approach using a partially-polarizing beamsplitter (PPBS) is shown in Fig. 2.15. In this scheme the unitary U_m, as before, acts on the polarization state in a single spatial mode, and the polarization-path coupling unitary V_m is part of the PPBS itself. The realisation of a passively phase-stable PPBS is discussed in Chapter 3.

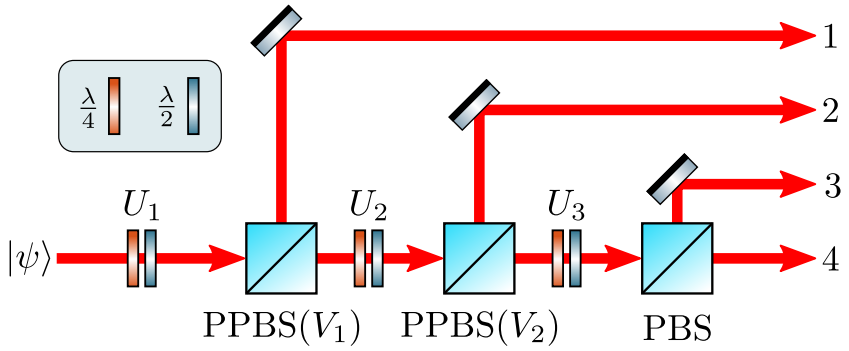


Figure 2.15: **Four-outcome POVM using tunable PPBSes.** Given access to tunable PPBSes the construction of a POVM on a polarization qubit can be simplified. As in the case of the beam-displacer based scheme in Fig. 2.14 the unitaries U_k act directly on the polarization using wave plates, but the path-polarization coupling unitaries V_k are the tunable PPBSes themselves.

2.3.3 Two-photon operations

As previously mentioned, the kind of bilinear Hamiltonians in (2.2-58) that characterise linear optics do not allow for deterministic entangling gates between degrees of freedom other than the photon number. This is because such Hamiltonians do not describe interactions between photons, and therefore cannot realise conditional operations. As an example, consider the controlled-phase gate

$$U_{\text{C-Phase}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{2.3-50}$$

This is a maximally entangling two-qubit gate:

$$\begin{aligned} U_{\text{C-Phase}} |+,+\rangle &= \frac{1}{2} U_{\text{C-Phase}} (|H,H\rangle + |H,V\rangle + |V,H\rangle + |V,V\rangle) \\ &= \frac{1}{2} (|H,H\rangle + |H,V\rangle + |V,H\rangle - |V,V\rangle) \\ &= \mathbb{1} \otimes U_{\text{Had}} \frac{|H,H\rangle + |V,V\rangle}{\sqrt{2}} \\ &= \mathbb{1} \otimes U_{\text{Had}} |\Phi^+\rangle , \end{aligned} \tag{2.3-51}$$

where $U_{\text{Had}} = (X + Z)/\sqrt{2}$ denotes the Hadamard operator. To realise this gate one needs a Hamiltonian of the form

$$H_{\text{C-Phase}} = \phi \hat{a}_V^\dagger \hat{a}_V \hat{b}_V^\dagger \hat{b}_V, \tag{2.3-52}$$

however a linear-optical phase shift is local to each mode:

$$H_{\text{linop}} = \phi_a \hat{a}_V^\dagger \hat{a}_V + \phi_b \hat{b}_V^\dagger \hat{b}_V, \tag{2.3-53}$$

and the resulting unitary operator is therefore simply a tensor product of two local phase shifts on modes a and b, and clearly does not generate any entanglement.

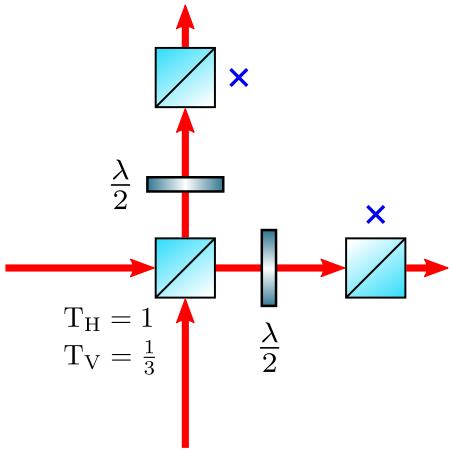


Figure 2.16: **Linear-optical C-Phase gate.** A linear-optical controlled-phase gate for polarization can be realised using three partially-polarizing beamsplitters that fully transmit horizontally polarized light, and transmit vertically polarized light with a probability of 1/3. The two half-wave plates exchange the H and V polarizations, and therefore effectively swap the transmission coefficients of the second two PPBSes. The gate succeeds whenever exactly one photon is detected in the transmission ports of the last two PPBSes, and this occurs with probability 1/9.

While deterministic entanglement generation is not possible with linear optics alone, it is nevertheless possible to generate entanglement probabilistically. In general terms, this is done by preparing a state that contains a set of amplitudes that constitute an entangled state in such a way that the rest of the state, that contributes to making it separable, can be removed using a projective measurement. A simple polarizing beamsplitter is capable of probabilistically entangling two photons in polarization. Consider the state:

$$\frac{1}{2}(\hat{a}_H^\dagger + \hat{a}_V^\dagger)(\hat{b}_H^\dagger + \hat{b}_V^\dagger)\,|0\rangle\,, \tag{2.3-54}$$

incident on a PBS. The state after the PBS is

$$|\psi\rangle = \frac{1}{2}\big((\hat{a}_H^\dagger\hat{b}_H^\dagger + \hat{a}_V^\dagger\hat{b}_V^\dagger + \hat{a}_H^\dagger\hat{a}_V^\dagger + \hat{b}_H^\dagger\hat{b}_V^\dagger)\,|0\rangle\,. \tag{2.3-55}$$

The two last terms represent amplitudes where both photons exited in the same spatial mode. Projecting on exactly one photon being in each mode gives a two-photon Bell-state, with probability 1/2:

$$\begin{aligned}\frac{1}{2}\big(\langle H,H| + \langle V,V| + \langle H,V| + \langle V,H|\big)\,|\psi\rangle =& \frac{1}{2}\big(|H,H\rangle + |V,V\rangle\big)\\ =& \frac{1}{2}\,|\Phi^+\rangle\,.\end{aligned} \tag{2.3-56}$$

Since non-destructive measurements of the photon number are challenging to realise, direct photo-detection is normally used instead. This has the consequence that the entangled state is destroyed in the same instant it is created, and this is referred to as entanglement by post-selection. In other words, non-classical correlations can be observed by disregarding certain measurement outcomes.

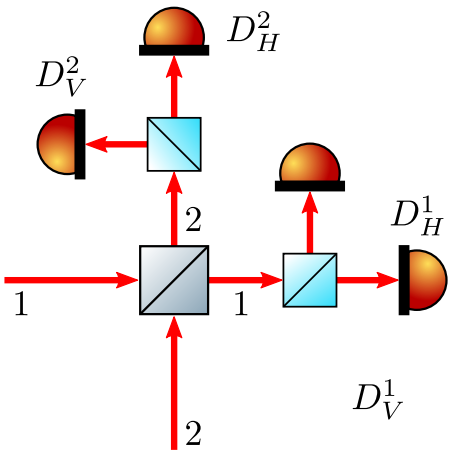


Figure 2.17: **Linear-optical Bell measurement.** Using linear optics a probabilistic Bell measurement can be realised, shown here for a polarization state $|\psi\rangle_{1,2}$. It consists of one balanced beamsplitter and two polarizing beamsplitters. The anti-symmetric state $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|H,V\rangle - |V,H\rangle)$ has the property that it never bunches on a beamsplitter, and it is the only two-qubit state with this property. Detecting exactly one photon in mode 1 and one in mode 2 therefore projects the input state onto $|\Psi^-\rangle_{1,2}$. The remaining three Bell states all bunch on the beamsplitter, but the state $|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|H,V\rangle + |V,H\rangle)$ will bunch on either PBS. A detection in both D_H^1 and D_V^1, or D_H^2 and D_V^2, therefore projects the input state onto $|\Psi^+\rangle_{1,2}$. By performing local polarization operations before the beamsplitter it is possible to project onto the other Bell states, however for any given configuration the setup is only capable of projecting onto two of them.

While the PBS interaction described above can generate a maximally entangled two-qubit state, it does not correspond to a gate in the circuit model even after post-selection, since there are certain input states such as $|H,V\rangle$ that never yield a qubit state as output, because the photons are guaranteed to bunch. Using slightly more complicated schemes, however, it is possible to build valid gates [39]. One example of this is the linear-optical C-Phase gate [166], consisting of three partially-polarizing beamsplitters, shown in Fig. 2.16 [167]. The last two PPBSes balance the amplitudes of the terms that bunch less frequently or not at all, with the $|V,V\rangle$ term that bunches more frequently. This gate has a success probability of only 1/9, which is in fact optimal [168], however in certain specific situations, such as when generating a four-qubit cluster state, it is possible to remove the last two PPBSes and balance the output amplitudes by intentionally preparing an imbalanced input state, and thereby boost the success probability to 1/4. Despite yielding a well defined circuit model gate, the need for post-selection nevertheless means that multiple instances of these gates generally cannot be concatenated, since the state after the gate is only known to be a valid qubit-state if the photons are measured.

A related idea to the probabilistic entangling gate is the projection of a state onto an entangled basis. When the basis is the two-qubit basis formed by the Bell states this is known as a Bell measurement. A probabilistic Bell measurement can be realised using a single balanced beamsplitter [62, 169], which has the property that it sorts two-qubit states according to their symmetry: anti-symmetric states always anti-bunch, meaning one photon exits in each mode, and symmetric states always bunch, i.e. both photons

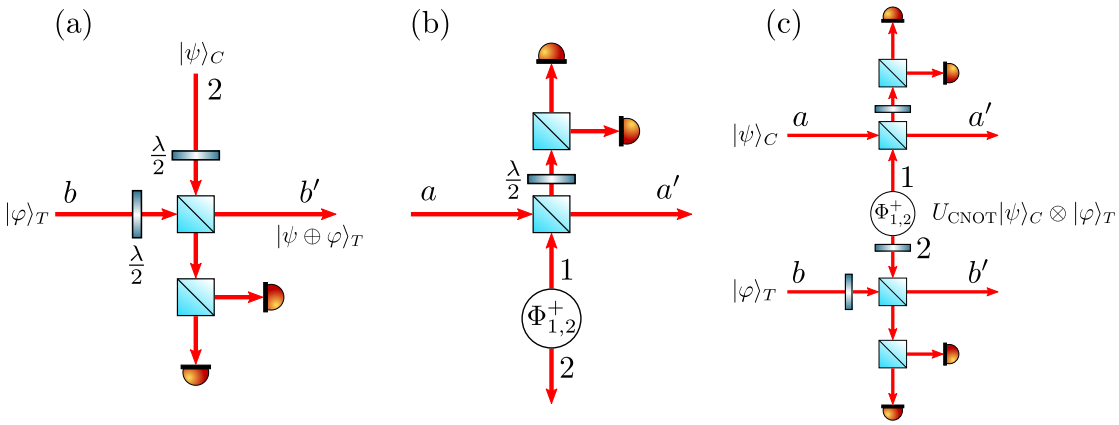


Figure 2.18: **Heralded C-NOT gate.** A heralded linear-optical C-NOT gate can be constructed using polarizing beamsplitters. It consists of two parts, the first one shown in (a) is a destructive C-NOT that succeeds whenever one and only one photon is detected in the two photodetectors at the output of the second PBS. The two half-wave plates change the basis of the PBS from the H/V basis to the $+/-$ basis. The control and target qubits of the gate are indicated by the subscripts C and T respectively. The second part shown in (b) is called a quantum encoder, and also succeeds whenever one and only one photon is detected. The gate has the effect of copying the state in mode a into two different photons, as described in the main text, and it requires a maximally entangled $|\Phi^+\rangle$ state as a resource. Combining these two parts, as shown in (c), leads to the realisation of a heralded linear-optical C-NOT gate that succeeds whenever exactly one photon is detected in both the upper and lower sets of photodetectors, which happens with probability 1/4.

exit in the same mode. Detecting one photon in each mode therefore projects the state before the beamsplitter onto the $|\Psi^-\rangle$ state, under the assumption that a valid two-qubit state was input. If the qubits are encoded in the polarization, performing a polarization resolving measurement in each output arm of the beamsplitter can furthermore distinguish the state $|\Psi^+\rangle$ from the remaining two Bell states, as this is the only symmetric Bell-state that is anti-correlated in the H, V basis. This is shown in Fig. 2.17. The last two states, $|\Phi^+\rangle$ and $|\Phi^-\rangle$, are both symmetric and correlated in the H, V basis, meaning that they both lead to photon bunching at the beamsplitter and at either PBS, and can therefore not be distinguished. Formulated as a discrimination task between the different Bell states, this scheme based on a single beamsplitter only succeeds 50 % of the time, and it can be shown that this is in fact optimal for a two-photon Bell measurement using only linear optics [170, 171].

By giving an experimenter more resources in the form of additional photons that are consumed in the process, commonly referred to as ancillas, it becomes possible to boost the success probability up to 75 % [172]. This approach can also be applied to two-qubit logic gates. The ancilla photons can not only boost the success probability, but can crucially be used herald the successful application of the gate, removing the need for destructive post-selection [39, 173]. An example of such a gate is the heralded C-NOT gate, which has a success probability of 1/4 and has seen multiple experimental realisations [174, 175, 176]. This gate, illustrated in Fig. 2.18, uses a two ancilla photons in a maximally entangled Bell-state, and consists of two different parts. The first part,

shown in Fig. 2.18a, is a destructive C-NOT gate that is heralded, but consumes the control qubit in the process and succeeds with probability 1/2. To preserve the control qubit state, a second part that entangles this state with one of the ancilla photons is needed. This heralded entangling operation, shown in Fig. 2.18b, is called a quantum encoder, and also consumes one photon. The quantum encoder realises the following transformation:

$$\alpha \left|H\right\rangle_a + \beta \left|V\right\rangle_a \rightarrow \alpha \left|H\right\rangle_{a'} \left|H\right\rangle_2 + \beta \left|V\right\rangle_{a'} \left|V\right\rangle_2 , \tag{2.3-57}$$

where the modes are labelled according to Fig. 2.18b. Note that this copying is not equivalent to quantum cloning. It is, however, probabilistic, by virtue of relying on a specific parity check outcome and succeeds with probability 1/2 [173]. Combining the quantum encoder with the destructive C-NOT allows the control qubit to be preserved, and since both constituent elements of the gate are heralded, the full gate is as well. When allowing for feed-forward operations to correct heralding-outcome-dependent phase shifts this gate succeeds with a probability of 1/4, dropping to 1/16 when no feed-forward is used.

In the context of linear-optical quantum computation several other types of gates also exist. Of particular interest are so-called fusion gates, which as the name suggests are used to fuse different states into a large cluster state that can be used as a computational resource [177]. It has been shown that fusion gates with a success probability exceeding certain bounds can enable deterministic quantum computation [159, 178, 179, 180].

2.4 Single-photon generation

Since the photon statistics of classical light sources such as black bodies and lasers do not approach those of the single-photon state even in the limit of low average photon numbers, it is not possible to produce single photons simply by attenuating traditional light sources, and a different method is needed. The best way, on paper, to generate a single-photon state is using an idealized two-level system. By driving such a system with a laser it can be deterministically excited to the upper level. Once in the excited state it will start to undergo spontaneous decay, emitting exactly one photon in the process. Such idealized two-level systems are, unfortunately, hard to find or engineer and real-world single-photon sources have many trade-offs.

A perfect single-photon source should emit one and only one photon, into well defined spatial, frequency and momentum modes, and do so on demand with 100 % efficiency [181, 182]. One could imagine using an optical transition of single atom as a single-photon source, but already there the difficulties begin. A single atom is very sensitive to its surroundings, and unless carefully isolated will therefore exhibit time-varying characteristics that can reduce the spectral purity of the light. Another problem is that spontaneous emission is spatially isotropic, meaning that the photon will be emitted in a random direction. Additionally, no atom is a two-level system and unwanted decay paths can alter the emission spectrum, among other things. Despite these challenges, great strides have been made towards realising true on-demand single-photon sources [181, 182, 183, 184, 185].

While neutral atoms have been used to generate photons [186, 187], they are far from

the only physical platform that has been explored. Other examples include trapped ions [188,189], single molecules [190,191], color centers [192,193,194,195] and quantum dots [196,197,198,199]. Common to all these physical systems is that they are generally placed inside optical cavities in order to enhance the emission into a particular mode through the Purcell effect [200, 201], and require cooling to cryogenic temperatures in order to operate. Although quantum-dot sources in particular have in recent years begun to show performance approaching that of an ideal single-photon source, due to their high cost and technical complexity they are nevertheless not the preferred way to generate single photons in all situations. Instead, by far the most common way to generate approximate single-photon states is by using bulk nonlinearities in various crystals, using what is known as parametric fluorescence [129, 202, 203, 204]. While this process is facilitated by a matter system, the response of the medium is a collective one and the microscopic internal degrees of freedom do not couple to the generated light, a fact which greatly reduces the experimental complexity and allows for high reproducibility. How this fluorescence phenomenon can be used for single-photon generation will be explored in the following sections.

2.4.1 Nonlinear optics

While most bulk materials are approximately linear, they do in fact generally have a small but non-negligible nonlinear response to external electric fields. What this means more concretely is that an oscillating electric field will perturb the electrons in the solid and cause them to oscillate, and that this oscillation will not only occur at the driving frequency, but will have small contributions from other frequency components as well. As the electrons' positions inside the material are perturbed the material itself becomes polarized and a net dipole moment is induced. This is described by the polarization density, which is the induced electric dipole moment per unit volume. In classical electrodynamics, this polarization density as a function of the external electric field is usually expanded in a Taylor series [106, 205]:

$$\mathbf{P} = \epsilon_0 \chi^{(1)} \mathbf{E} + \epsilon_0 \chi^{(2)} \mathbf{E}^2 + \epsilon_0 \chi^{(3)} \mathbf{E}^3 + \cdots \tag{2.4-1}$$

where ϵ_0 is the vacuum permittivity, $\chi^{(n)}$ is the n-th order susceptibility tensor that describes the response of the medium, and the boldface indicates a vector-valued quantity. $\chi^{(n)}$ is a tensor with $n+1$ Cartesian components, which means that the linear response is simply described by a matrix. The notation $\epsilon_0 \chi^{(n)} \mathbf{E}^n$ is shorthand for the sum

$$P_j^{(n)} = \epsilon_0 \sum_{k\ldots l=1}^{3} \chi^{(2)}_{jk\ldots l} E_k \cdots E_l. \tag{2.4-2}$$

For example, the third-order term is

$$P_j^{(3)} = \epsilon_0 \sum_{klm=1}^{3} \chi^{(3)}_{jklm} E_k E_l E_m. \tag{2.4-3}$$

If the external field is oscillating at some frequency ω then nonlinear terms describe a polarization density that oscillates at higher harmonics of this frequency, and the time-varying dipole moment in the solid can therefore emit light at these other frequencies.

A common way to estimate the magnitude of $\chi^{(n)}$ is by noting that one would expect the size of the nonlinear terms to be comparable to the linear response when the external electric field strength is on the order of the atomic electric field strength, which can be estimated as $E_{\text{atom}} = e/(4\pi\epsilon_0 a_0^2)$, where e is the electron charge and a_0 is the Bohr radius [206, 207]. In a homogeneous, non-dispersive and isotropic dielectric medium $\chi^{(1)}$ is a scalar and

$$\frac{\epsilon}{\epsilon_0} = n^2 = 1 + \chi^{(1)}, \tag{2.4-4}$$

where n is the refractive index of the medium. This implies that $\chi^{(1)}$ is of order unity, and consequently the nonlinear terms can be estimated as

$$\chi^{(n)} \approx \frac{1}{E_{\text{atom}}^n} \approx (2 \times 10^{-12}\,\text{V}\,\text{m}^{-1})^n. \tag{2.4-5}$$

This simple estimate is surprisingly accurate for the low order nonlinear response, and it also suggests that the nonlinear response of dielectric materials will typically be very small.

The dynamics of an electromagnetic field inside a medium are governed by Maxwell's equations:

$$\nabla \cdot \mathbf{D} = \rho, \tag{2.4-6}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{2.4-7}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{2.4-8}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, \tag{2.4-9}$$

where $\mathbf{J}$ is a vector describing the free current density inside the material and ρ is the density of free charges. If the medium is non-magnetic then the $\mathbf{H}$ and $\mathbf{B}$ fields are proportional: $\mathbf{H} = \mathbf{B}/\mu_0$, where μ_0 is the vacuum permeability.

From this set of equations one can derive a single differential equation describing the propagation of waves inside a medium [106, 208]. We begin by taking the curl of (2.4-8):

$$\nabla \times (\nabla \times \mathbf{E}) = -\frac{\partial}{\partial t}(\nabla \times \mathbf{B}), \tag{2.4-10}$$

where the curl can be moved inside the time derivative because the fields are sufficiently continuous that the partial derivatives commute. Using the vector identity $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ and assuming that the medium is non-magnetic, this can be rewritten as

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\frac{\partial}{\partial t}(\nabla \times \mu_0 \mathbf{H}) \tag{2.4-11}$$

Assuming further that the free current density is zero and inserting (2.4-9) the equation becomes

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu_0 \frac{\partial^2 \mathbf{D}}{\partial t^2}. \tag{2.4-12}$$

For dielectric materials $\nabla \cdot \mathbf{D} = 0$, and if the medium is isotropic then the $\mathbf{E}$- and $\mathbf{D}$-fields are proportional meaning that $\nabla \cdot \mathbf{E} = 0$. In most applications of nonlinear optics the medium in question is not isotropic, however under the assumption of weak birefringence $\nabla \cdot \mathbf{E} \approx 0$ still holds, simplifying the equation to

$$\nabla^2 \mathbf{E} = \mu_0 \frac{\partial^2 \mathbf{D}}{\partial t^2}. \tag{2.4-13}$$

In a dielectric medium, the relationship between the displacement field $\mathbf{D}$ and electric field $\mathbf{E}$ is given by

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \tag{2.4-14}$$

and inserting this into (2.4-13) finally yields

$$\nabla^2 \mathbf{E} = \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2}. \tag{2.4-15}$$

From this equation one can see that the polarization of the medium acts as a source for the electric field, and that an oscillating polarization density will generate propagating electric waves. In processes that involve fields that all propagate in the same direction, called a collinear process, and in which these fields can be approximately described by plane waves, the vector-valued equation can be replaced by a one-dimensional version:

$$\frac{\partial^2 E}{\partial z^2} = \epsilon_0 \mu_0 \frac{\partial^2 E}{\partial t^2} + \mu_0 \frac{\partial^2 P}{\partial t^2}. \tag{2.4-16}$$

To see what kind of dynamics arise from this equation, we will examine a nonlinear process called difference-frequency generation (DFG). The quantum analogue of this process can be used to generate single photons. Consider a monochromatic electromagnetic wave:

$$E(t) = E_0 \mathrm{Re}\big[e^{i(-\omega t + kz + \phi)}\big] = \frac{1}{2}\Big(A e^{i(-\omega t + kz)} + A^* e^{i(\omega t - kz)}\Big), \tag{2.4-17}$$

where $A = E_0 e^{i\phi}$ is a complex electric field amplitude. Since we are interested in the nonlinear dynamics we need to evaluate nonlinear terms of the polarization density. In the case of DFG, the process is described by the second-order nonlinear response, which for the scalar equation is

$$P^{(2)} = \epsilon_0 d_{\mathrm{eff}} E_{\mathrm{tot}}^2, \tag{2.4-18}$$

where d_{eff} is an effective scalar nonlinear susceptibility along axis in question [207, 208], and E_{tot} is the total electric field. In the DFG process there are three interacting fields with different frequencies, and thus

$$E_{\mathrm{tot}} = E(\omega_p) + E(\omega_s) + E(\omega_i), \tag{2.4-19}$$

where the subscripts are shorthand for 'pump', 'signal' and 'idler', which are the conventional names for these fields. Substituting the corresponding complex amplitudes into (2.4-18) gives

$$P^{(2)} = \frac{1}{4}\epsilon_0 d_{\mathrm{eff}}\Big(A_p e^{i(-\omega_p t + k_p z)} + A_s e^{i(-\omega_s t + k_s z)} + A_i e^{i(-\omega_i t + k_i z)} + \mathrm{c.c.}\Big)^2, \tag{2.4-20}$$

where c.c. denotes the complex conjugate of the preceding terms. In DFG, the pump and idler fields combine to create photons in the signal field:

$$\hbar\omega_p + \hbar\omega_i \mapsto \hbar\omega_i + \hbar\omega_i + \hbar\omega_s, \tag{2.4-21}$$

where due to energy conservation $\omega_p = \omega_i + \omega_s$. For this reason, it suffices to consider terms in the expanded square of (2.4-20) that oscillate at the frequency $\omega_s = \omega_p - \omega_i$.[3] These are

$$P^{(2)}(\omega_s) = \frac{1}{4}\epsilon_0 d_{\text{eff}} A_p A_i^* e^{i(-\omega_s t + (k_p - k_i)z)} + A_p^* A_i e^{i(\omega_s t + (k_i - k_p)z)}. \tag{2.4-22}$$

Neglecting higher-order terms, the total polarization density at ω_s is

$$P(\omega_s) = \frac{1}{2}\epsilon_0 \chi^{(1)}(A_s e^{i(-\omega_s t + k_s z)} + \text{c.c.}) + P^{(2)}(\omega_s). \tag{2.4-23}$$

This expression, together with (2.4-17), can now be substituted into the one-dimensional wave equation (2.4-16). The left-hand term of this equation becomes

$$\frac{\partial^2}{\partial z^2}E(\omega_s) = \frac{1}{2}\frac{\partial^2}{\partial z^2}\Big(A_s e^{i(-\omega_s t + k_s z)} + A_s^* e^{i(\omega_s t - k_s z)}\Big) \tag{2.4-24}$$

$$= \frac{1}{2}\Big(- k_s^2 A_s + 2ik_s\frac{\partial A_s}{\partial z} + \frac{\partial^2 A_s}{\partial z^2}\Big)e^{i(-\omega_s t + k_s z)} + \text{c.c.} \tag{2.4-25}$$

The second-order spatial derivative in this expression is usually neglected. This is known as the slowly varying envelope approximation, and it is justified whenever $|\partial_z^2 A| \ll |k\partial_z A|$. In practice this approximation holds when the envelope does not vary significantly within a wavelength. The right-hand side of the wave equation can be expanded as

$$\begin{aligned}\frac{\partial^2}{\partial t^2}E(\omega_s) &= \frac{1}{2}\frac{\partial^2}{\partial t^2}\Big(A_s e^{i(-\omega_s t + k_s z)} + A_s^* e^{i(\omega_s t - k_s z)}\Big) \\ &= -\frac{1}{2}\omega_s^2 A_s e^{i(-\omega_s t + k_s z)} + \text{c.c.}\end{aligned} \tag{2.4-26}$$

$$\frac{\partial^2}{\partial t^2}P(\omega_s) = -\omega_s^2\Big(\frac{1}{2}\epsilon_0\chi^{(1)}A_s e^{i(-\omega_s t + k_s z)} + \epsilon_0 d_{\text{eff}} A_p A_i^* e^{i(-\omega_s t + (k_p - k_i)z)}\Big) + \text{c.c.} \tag{2.4-27}$$

These expressions can now be substituted into the one-dimensional wave equation (2.4-16), giving

$$\begin{aligned}&\frac{1}{2}\Big(- k_s^2 A_s + 2ik_s\frac{\partial A_s}{\partial z}\Big)e^{i(-\omega_s t + k_s z)} + \text{c.c} \\ &= -\omega_s^2\Big(\frac{1}{2}\epsilon_0\mu_0(\chi^{(1)} + 1)A_s e^{i(-\omega_s t + k_s z)} + \epsilon_0 d_{\text{eff}} A_p A_i^* e^{i(-\omega_s t + (k_p - k_i)z)}\Big) + \text{c.c.}\end{aligned} \tag{2.4-28}$$

Grouping all the terms oscillating at positive (negative) frequencies the differential equation becomes $C_+ e^{i\omega_s t} + C_- e^{-i\omega_s t} = 0$. It's clear that this equation only has the solution $C_+ = C_- = 0$, which means that the equality in (2.4-28) holds even if one discards the complex conjugate terms. Using the relations $n = \sqrt{1 + \chi^{(1)}}$, $1/c = \sqrt{\epsilon_0\mu_0}$, and $k = \omega n/c$ the A_s terms can, after some manipulation, be cancelled and wave equation reduces to

$$\frac{\partial A_s}{\partial z} = i\frac{\omega_s}{n_s c} d_{\text{eff}} A_p A_i^* e^{i(k_p - k_s - k_i)z}. \tag{2.4-29}$$

[3] Processes corresponding to other terms might still occur, depending on whether or not they satisfy energy and momentum conservation conditions.

Similar differential equations can be derived for the pump and idler amplitudes A_p and A_i leading to a system of coupled differential equations. However, if the interaction is weak the pump and idler fields can be treated as constant. The ordinary differential equation for the signal field can then be solved by integration, and the solution reads [209]

$$A_s = i\frac{\omega_s}{n_s c} L d_{\text{eff}} A_p A_i^* e^{i\Delta k L/2} \operatorname{sinc}\Big(\frac{\Delta k}{L}\Big), \tag{2.4-30}$$

where L is the length of the spatial region in which the interaction takes place and $\Delta k = k_p - k_s - k_i$ is the momentum mismatch between the three fields. Two things worth noting about this solution are, first of all, that if the amplitude in the idler field is zero then the signal amplitude remains zero as well, which as will be discussed in the next section is in contradiction with the quantum description. The second important property of the solution is the dependence on the momentum mismatch and interaction length, which prevents the signal field from growing monotonically, and also influences the spectral properties of the light generated in the process. This will be explored in more detail in Section 2.4.3.

2.4.2 Spontaneous Parametric Down-Conversion

Spontaneous parametric down-conversion, or SPDC, is a process in which a single photon spontaneously decays into a pair of photons with lower frequency:

$$|1\rangle_{\omega_p} \mapsto |1\rangle_{\omega_s} + |1\rangle_{\omega_i}\,. \tag{2.4-31}$$

This process is constrained by the energy and momentum conservation conditions

$$\omega_p = \omega_s + \omega_i \tag{2.4-32}$$

$$\vec{k}_p = \vec{k}_s + \vec{k}_i. \tag{2.4-33}$$

The spontaneous decay of a pump photon into a pair of photons corresponds to a difference-frequency-generation process where the amplitude of the idler field is zero. Classically this is forbidden, as discussed in the previous section, but in a semi-classical model SPDC can be thought of as being seeded by vacuum fluctuations that get amplified by the pump field [208]. In the quantized picture the SPDC process is described by a Hamiltonian of the form

$$H = H_0 + H_I, \tag{2.4-34}$$

with

$$H_0 = \hbar\omega_m \sum_m \Big(\hat{a}_m^\dagger \hat{a}_m + \frac{1}{2}\Big), \tag{2.4-35}$$

$$H_I = \kappa \hat{a}_p \hat{a}_s^\dagger \hat{a}_i^\dagger + \kappa^* \hat{a}_p^\dagger \hat{a}_s \hat{a}_i. \tag{2.4-36}$$

The H_0 term is the Hamiltonian of the free electric field, and H_I is the Hamiltonian describing the interaction between the fields. The first term in the interaction Hamiltonian describes the actual down-conversion process, while the second term, necessary to make the Hamiltonian hermitian, describes the time-reversed process called sum-frequency generation.

In most cases one is only interested in the interaction Hamiltonian, and it is therefore common to describe the process in the so-called interaction picture. This picture corresponds to a rotating reference frame in which the free evolution of the state is baked into the creation and annihilation operators, while the state vector evolves under the interaction Hamiltonian. More formally, let time-evolution operator under the free Hamiltonian be [18, 110]

$$U(t, t_0) = \exp\left[-\frac{i}{\hbar}\int_{t_0}^{t} H_0 d\tau\right], \tag{2.4-37}$$

and let $T = U^\dagger(t, t_0)$. The state vector in the interaction picture is then

$$|\Psi_I(t)\rangle = T\,|\Psi_S(t)\rangle\,, \tag{2.4-38}$$

where $|\Psi_S(t)\rangle$ is the Schrödinger-picture state vector. We are interested in the time evolution of the state in the interaction picture. Using the fact that

$$\frac{\partial}{\partial t}T = \frac{i}{\hbar}H_0 T, \tag{2.4-39}$$

together with the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\,|\Psi_S(t)\rangle = H\,|\Psi_S(t)\rangle\,, \tag{2.4-40}$$

the time evolution of the state can be written

$$\begin{aligned}\frac{\partial}{\partial t}\,|\Psi_I(t)\rangle &= \left(\frac{\partial}{\partial t}T\right)|\Psi_S(t)\rangle + T\frac{\partial}{\partial t}\,|\Psi_S(t)\rangle \\ &= \frac{i}{\hbar}H_0 T\,|\Psi_S(t)\rangle + T\frac{1}{i\hbar}(H_0 + H_I)\,|\Psi_S(t)\rangle\,.\end{aligned} \tag{2.4-41}$$

Multiplying both sides with $i\hbar$ and using the fact that $TH_0 = H_0 T$ the equation becomes

$$\begin{aligned}i\hbar\frac{\partial}{\partial t}\,|\Psi_I(t)\rangle &= (-TH_0 + TH_0 + TH_I)\,|\Psi_S(t)\rangle \\ &= TH_I T^\dagger\,|\Psi_I(t)\rangle\,.\end{aligned} \tag{2.4-42}$$

We see that the time evolution in the interaction picture obeys the Schrödinger equation with the interaction Hamiltonian in the rotating frame:

$$H_I' = TH_I T^\dagger, \tag{2.4-43}$$

and the free time evolution does not need to be taken into account explicitly. The transformed interaction Hamiltonian can be written on the same form as (2.4-36) with the transformed creation and annihilation operators

$$\hat{a}^\dagger \mapsto T\hat{a}^\dagger T^\dagger \tag{2.4-44}$$

$$\hat{a} \mapsto T\hat{a}T^\dagger, \tag{2.4-45}$$

and this will be done implicitly for the remainder of this chapter where necessary.

In SPDC the signal and idler intensities are intentionally kept low, and a strong coherent pump field remains, for all intents and purposes, unaffected by the process. The pump field can therefore be described classically, which amounts to replacing the annihilation operator of the pump mode with a complex classical variable [210]:

$$\hat{a}_p \to \alpha_p \in \mathbb{C}. \tag{2.4-46}$$

In the regime where this is valid, the SPDC interaction Hamiltonian (2.4-36) simply becomes the squeezing operator from (2.2-47)

$$H = \kappa\alpha\hat{a}_s^\dagger\hat{a}_i^\dagger + \kappa^*\alpha^*\hat{a}_s\hat{a}_i. \tag{2.4-47}$$

When the signal and idler fields occupy two different modes this is a two-mode squeezing Hamiltonian, and if they occupy the same mode it corresponds to the single-mode squeezing operator. Instead of solving the dynamics exactly, SPDC is usually treated perturbatively in the aforementioned interaction picture, where one considers a scenario in which the interaction is switched on for some time t, after which the state evolves freely:

$$|\Psi\rangle_t = \mathcal{T}\exp\left[\frac{1}{i\hbar}\int_0^t H(\tau)d\tau\right]|\Psi\rangle_0 . \tag{2.4-48}$$

In this case the time-ordering operator $\mathcal{T}$ can be neglected since the Hamiltonian in (2.4-47) does not have a time dependence. Taylor expanding (2.4-48) applied to an initial state with zero photons in the signal and idler modes one gets:

$$\begin{aligned}|\Psi\rangle_t &= \exp\left[\frac{1}{i\hbar}\int_0^t H(\tau)d\tau\right]|0,0\rangle_{s,i} \\ &= \left[1+\frac{1}{i\hbar}\int_0^t H(\tau)d\tau+\frac{1}{2!}\left(\frac{1}{i\hbar}\int_0^t H(\tau)d\tau\right)^2+\cdots\right]|0,0\rangle_{s,i} \\ &= \left[1+\frac{t}{i\hbar}H+\frac{1}{2!}\left(\frac{t}{i\hbar}H\right)^2+\cdots\right]|0,0\rangle_{s,i}.\end{aligned} \tag{2.4-49}$$

SPDC is usually operated in a regime where the interaction strength, governed by the nonlinear coefficient of the material, the optical power of the pump light and the total interaction time, is very weak. Therefore the Taylor series can be truncated to the first two terms:

$$\begin{aligned}|\Psi\rangle_t &= \exp\left[\frac{1}{i\hbar}\int_0^t H(\tau)d\tau\right]|0,0\rangle_{s,i} \approx \left[1+\frac{t}{i\hbar}H\right]|0,0\rangle_{s,i} \\ &= C_0|0,0\rangle_{s,i}+C_1|1,1\rangle_{s,i}.\end{aligned} \tag{2.4-50}$$

In the regime where this approximation is valid $C_0 \gg C_1$, which means that the above state contains mostly vacuum, and as described in Section 2.2.6, the local description of each mode corresponds to a weak thermal state. However, due to the strong photon-number correlation arising from the pair creation process, detection of a photon in one of the modes projects the other mode onto the Fock state $|1\rangle$. Even though the probability to project one of the modes onto this Fock state is typically low, the ability to herald the creation of a single-photon state in the second mode can be used to overcome the

probabilistic creation process by employing feedforward operations [211, 212, 213]. This can be done by, for example, simultaneously pumping processes in several independent modes, such that the probability that at least one of them produces a photon pair is high. The effective single-photon-generation probability of SPDC sources can therefore be increased.

While the toy-model Hamiltonian in (2.4-47) gives a qualitative description of the SPDC process and is sufficient for understanding how it can be used to generate single-photon states, many applications of SPDC require a more quantitative description of the process. Such a description begins with the interaction Hamiltonian of the process, which can be derived from a classical starting point. More concretely, by finding an expression for the total energy of the electromagnetic field inside a dielectric.

The energy of an electric field can be expressed many different ways. One way is in terms of the potential of the field, which by definition is the amount of energy, per unit charge, required to add electric charge at a given point. Consider an electric field with a potential $\phi(\vec{r})$, to which a small amount of charge $d\rho(\vec{r})$ is added. The total change in the electric energy W_E is then [214]

$$dW_E = \int \phi(\vec{r})\, d\rho(\vec{r})\, d^3r, \tag{2.4-51}$$

where the integration is taken over all space. Using the relation $\nabla \cdot \mathbf{D} = \rho$ from (2.4-6) this can be rewritten as

$$\begin{aligned} dW_E &= \int \phi \nabla \cdot d\mathbf{D} d^3r \\ &= \int \nabla \cdot (\phi d\mathbf{D}) d^3r - \int \nabla\phi \cdot d\mathbf{D} d^3r. \end{aligned} \tag{2.4-52}$$

The first integral can be rewritten as a surface integral using Gauss' theorem:

$$dW_E = \int \phi d\mathbf{D} \cdot d\mathbf{S} - \int \nabla\phi \cdot d\mathbf{D} d^3r. \tag{2.4-53}$$

When the integration is taken over all space the surface integral vanishes under the condition that the potential goes to zero at infinity, which is the case when considering a medium with a finite extent [215]. By substituting in $\mathbf{E} = -\nabla\phi$ the change in the total energy can be written as

$$dW_E = \int \mathbf{E} \cdot d\mathbf{D} d^3r. \tag{2.4-54}$$

The total energy can then be found by continuously increasing the strength of the displacement field from $\mathbf{D} = 0$ to its final value $\mathbf{D}$, and integrating the change in energy during this process:

$$W_E = \int d^3r \int_0^D \mathbf{E} \cdot d\mathbf{D}. \tag{2.4-55}$$

A similar derivation can be performed for the magnetic contribution to the total energy:

$$W_M = \int d^3r \int_0^B \mathbf{H} \cdot d\mathbf{B} = \frac{1}{2\mu_0} \int \mathbf{B} \cdot \mathbf{B} d^3r, \tag{2.4-56}$$

where the second step assumes a non-magnetic material. Integrating (2.4-55) is less straightforward since we assume that the medium is nonlinear, however it can be performed by, as before, Taylor expanding the nonlinear response. First, note that the integrand in (2.4-55) can be rewritten as [216]

$$\mathbf{E}\cdot d\mathbf{D} = \mathbf{E}\cdot\left(\frac{d\mathbf{D}}{d\mathbf{E}}\right)d\mathbf{E}, \tag{2.4-57}$$

where

$$\left(\frac{d\mathbf{D}}{d\mathbf{E}}\right)_{ij} = \frac{\partial D_i}{\partial E_j} \tag{2.4-58}$$

is a Jacobian. Recalling that:

$$\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}, \tag{2.4-59}$$

The integrand becomes

$$\mathbf{E}\cdot d\mathbf{D} = \epsilon_0\mathbf{E}\cdot d\mathbf{E} + \mathbf{E}\cdot\left(\frac{d\mathbf{P}}{d\mathbf{E}}\right)d\mathbf{E}, \tag{2.4-60}$$

where, as before

$$P_i = \epsilon_0\chi^{(1)}_{ij}E_j + \epsilon_0\chi^{(2)}_{ijk}E_jE_k + \epsilon_0\chi^{(3)}_{ijkl}E_jE_kE_l + \dots \tag{2.4-61}$$

and where repeated indices are summed over. The first-order contribution from the polarization density is simple to evaluate since

$$\left(\frac{d\mathbf{P}^{(1)}}{d\mathbf{E}}\right)_{ij} = \frac{\partial}{\partial E_j}\left(\epsilon_0\chi^{(1)}_{ik}E_k\right) = \epsilon_0\chi^{(1)}_{ij}, \tag{2.4-62}$$

and hence

$$\mathbf{E}\cdot\left(\frac{d\mathbf{P}^{(1)}}{d\mathbf{E}}\right)d\mathbf{E} = \epsilon_0E_i\chi^{(1)}_{ij}dE_j, \tag{2.4-63}$$

For the second-order term the Jacobian can be expressed as

$$\begin{aligned}\left(\frac{d\mathbf{P}^{(2)}}{d\mathbf{E}}\right)_{ij} &= \frac{\partial}{\partial E_j}\left(\epsilon_0\chi^{(2)}_{ikl}E_kE_l\right)\\ &= \epsilon_0\chi^{(2)}_{ikl}\left(2\delta_{jk}\delta_{jl}E_j + (1-\delta_{jk})\delta_{jl}E_k + \delta_{jk}(1-\delta_{jl})E_l\right).\end{aligned} \tag{2.4-64}$$

In the above expression the terms with two delta functions cancel. Under the assumption that the medium is lossless, nondispersive and uniform the susceptibility tensors are real and symmetric [217], and this can be used to further simplify the expression above by grouping the remaining terms through a permutation of the indices. This assumption is justified when the fields are detuned far away from the resonances of the medium [207]. After re-labelling the indices one finds

$$\left(\frac{d\mathbf{P}^{(2)}}{d\mathbf{E}}\right)_{ij} = 2\epsilon_0\chi^{(2)}_{ijk}E_j \tag{2.4-65}$$

$$\mathbf{E}\cdot\left(\frac{d\mathbf{P}^{(2)}}{d\mathbf{E}}\right)d\mathbf{E} = 2\epsilon_0E_i\chi^{(2)}_{ijk}E_jdE_k. \tag{2.4-66}$$

Although not necessary for SPDC, an expression for the n-th order term can be found by noting that

$$\frac{\partial}{\partial E_j}\left(\chi^{(n)}_{I_{n+1}}\prod_{i\in I_n}E_i\right)=\chi^{(n)}_{I_{n+1}}\sum_{i\in I_n}\delta_{ij}\prod_{k\neq i}E_k, \tag{2.4-67}$$

where I_n is the set of the n last indices of the susceptibility tensor. Using the permutation symmetry of the indices once again all the products on the right-hand side can be made equal and

$$\chi^{(n)}_{I_{n+1}}\sum_{i\in I_n}\delta_{ij}\prod_{k\neq i}E_k=n\chi^{(n)}_{I_{n+1}}\prod_{j\neq k}E_k, \tag{2.4-68}$$

where $l\in I_n$ can be chosen freely. Consequently

$$\mathbf{E}\cdot\left(\frac{d\mathbf{P}^{(n)}}{d\mathbf{E}}\right)d\mathbf{E}=\epsilon_0 nE_i\chi^{(n)}_{I_{n+1}}dE_l\prod_{k\neq i,l}E_k. \tag{2.4-69}$$

To find the total energy it remains to evaluate the integrals

$$\int\mathbf{E}\cdot\left(\frac{d\mathbf{P}^{(n)}}{d\mathbf{E}}\right)d\mathbf{E}. \tag{2.4-70}$$

These integrals should be understood as path integrals over the trajectory along which $\mathbf{E}$ is continuously varied. Let this path be parameterised by the variable τ, and consider the integral of the first-order term:

$$\begin{aligned}\epsilon_0\int E_i\chi^{(1)}_{ij}dE_j&=\epsilon_0\int_0^t E_i(\tau)\chi^{(1)}_{ij}\frac{dE_j(\tau)}{d\tau}d\tau\\&=\epsilon_0E_i(\tau)\chi^{(1)}_{ij}E_j(\tau)\Big|_{\tau=0}^{\tau=t}-\epsilon_0\int_0^\tau\frac{dE_i(\tau)}{d\tau}\chi^{(1)}_{ij}E_j(\tau)d\tau.\end{aligned} \tag{2.4-71}$$

Taking $E_i(0)=0$ and $E_j(t)=E_j$, and using the permutation symmetry on the right-hand integral gives

$$\epsilon_0\int_0^t E_i(\tau)\chi^{(1)}_{ij}\frac{dE_j(\tau)}{d\tau}d\tau=\epsilon_0\frac{1}{2}E_i\chi^{(1)}_{ij}E_j. \tag{2.4-72}$$

The integral of the first right-hand term in (2.4-60) also follows from this by taking the susceptibility tensor to be the identity tensor. The integral above can alternatively be evaluated directly by making use of the permutation symmetry of the susceptibility tensor:

$$\begin{aligned}\epsilon_0\int E_i\chi^{(1)}_{ij}dE_j&=\epsilon_0\frac{1}{2}\left(\int E_i\chi^{(1)}_{ij}dE_j+\int E_j\chi^{(1)}_{ji}dE_i\right)\\&=\epsilon_0\frac{1}{2}\int d\big[\chi_{ij}E_iE_j\big]\\&=\epsilon_0\frac{1}{2}\chi_{ij}E_iE_j,\end{aligned} \tag{2.4-73}$$

where the second step used the product rule for differentials. This approach simplifies the evaluation of the second-order term as well:

$$\begin{aligned} 2\epsilon_0 \int E_i \chi^{(2)}_{ijk} E_j dE_k &= \epsilon_0 \frac{2}{3} \left(\int E_i \chi^{(2)}_{ijk} E_j dE_k + \int E_i \chi^{(2)}_{ikl} E_k dE_j + \int E_j \chi^{(2)}_{jki} E_k dE_i \right) \\ &= \epsilon_0 \frac{2}{3} \int d\big[E_i \chi^{(2)}_{ijk} E_j dE_k \big] \\ &= \epsilon_0 \frac{2}{3} E_i \chi^{(2)}_{ijk} E_j E_k. \end{aligned} \tag{2.4-74}$$

More generally:

$$\int \mathbf{E} \cdot \left(\frac{d\mathbf{P}^{(n)}}{d\mathbf{E}} \right) d\mathbf{E} = \epsilon_0 \frac{n}{n+1} \chi^{(n)}_{I_{n+1}} \prod_{i \in I_{n+1}} E_i. \tag{2.4-75}$$

The total energy is given by the sum of the electric and magnetic contributions:

$$H = W_E + W_M, \tag{2.4-76}$$

which can finally be expressed as

$$H = \frac{\epsilon_0}{2} \int \Big(c^2 \mathbf{B}^2 + \mathbf{E}^2 + \chi^{(1)} \mathbf{E}^2 + 2 \sum_{n \geq 2} \frac{n}{n+1} \chi^{(n)} \mathbf{E}^{n+1} \Big) d^3 r. \tag{2.4-77}$$

The interaction Hamiltonian for the SPDC process will be given by the second-order term in (2.4-77):

$$H^{(2)} = \epsilon_0 \frac{2}{3} \int \chi^{(2)}_{ijk} E_i E_j E_k d^3 r. \tag{2.4-78}$$

Note that while the Hamiltonian was derived for electrostatic fields, the field energy is the same for time-varying fields. More concretely, one can arrive at the same expression for the energy density by using Lorentz's force law to express the rate of work done on the free charges in a volume V [215]:

$$\begin{aligned} \frac{dW}{dt} = \int_V \mathbf{E} \cdot \mathbf{J} d^3 r &= \int_V \left(\frac{1}{\mu_0} \mathbf{E} \cdot (\nabla \times \mathbf{B}) - \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \right) d^3 r \\ &= - \int_V \left(\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \frac{1}{\mu_0} \mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} \right) d^3 r - \frac{1}{\mu_0} \int_V \nabla \cdot (\mathbf{E} \times \mathbf{B}) d^3 r, \end{aligned} \tag{2.4-79}$$

where the first step used (2.4-9) for a non-magnetic medium, and the second step used (2.4-8) together with the vector-calculus identity $\nabla \cdot (\mathbf{E} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{B})$. By rewriting the term as a surface integral using Gauss' theorem one recovers Poynting's theorem. The first integral in the bottom r.h.s. can be identified as the time derivative of the energy density, and using the same methods as for the electrostatic energy density it can be brought back to a familiar form. For the electric-field term we have

$$\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} = \epsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} + \mathbf{E} \cdot \left(\frac{d\mathbf{P}}{d\mathbf{E}} \right) \frac{\partial \mathbf{E}}{\partial t}. \tag{2.4-80}$$

Evaluating, as an example, the second-order term gives

$$\mathbf{E}\cdot\left(\frac{d\mathbf{P}^{(2)}}{d\mathbf{E}}\right)\frac{\partial\mathbf{E}}{\partial t}=2\epsilon_0 E_i\chi^{(2)}_{ijk}E_j\frac{\partial E_k}{\partial t}=\epsilon_0\frac{2}{3}\frac{\partial}{\partial t}\chi^{(2)}_{ijk}E_iE_jE_k. \tag{2.4-81}$$

Performing the same steps for all terms in the Taylor expansion of $\mathbf{P}$ and moving the time derivative outside the integral recovers (2.4-77).

To obtain the quantum description the Hamiltonian needs to be quantized, which amounts to replacing the classical electric-field terms with field operators. The quantization of the electromagnetic field inside a nonlinear dielectric is a non-trivial problem, and a common approach is to simply use the operators obtained by quantizing the free electric field [104, 210, 218, 219]. This approach is generally justified when the nonlinearities in question are small, and there are not multiple interacting processes. While the remainder of the discussion on SPDC will use of the free electromagnetic field, some of the issues with this approach are explored in greater detail in Appendix C.

The free field we wish to quantize obeys Maxwell's equations (2.4-6)-(2.4-9) with vanishing free charge and current densities: $\rho=0$, $\mathbf{J}=0$. To perform the quantization it is convenient to first introduce the vector potential $\mathbf{A}$, through which the magnetic field can be expressed:

$$\mathbf{B}=\nabla\times\mathbf{A}. \tag{2.4-82}$$

Using (2.4-8) we can also express the electric field in terms of potentials:

$$\nabla\times\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t}=-\frac{\partial}{\partial t}\nabla\times\mathbf{A}=-\nabla\times\frac{\partial\mathbf{A}}{\partial t}. \tag{2.4-83}$$

Using the distributive property of the curl we find

$$\nabla\times\left(\mathbf{E}+\frac{\partial\mathbf{A}}{\partial t}\right)=0, \tag{2.4-84}$$

and, since the curl of a gradient is always zero, we can identify the scalar potential ϕ as

$$-\nabla\phi=\mathbf{E}+\frac{\partial\mathbf{A}}{\partial t}. \tag{2.4-85}$$

The electric field can then be written as

$$\mathbf{E}=-\nabla\phi-\frac{\partial\mathbf{A}}{\partial t}. \tag{2.4-86}$$

Defined in terms of these potentials the electric and magnetic fields satisfy (2.4-7) and (2.4-8). The potentials themselves are constrained by the remaining two Maxwell equations (2.4-6) and (2.4-9). Since we are considering the fields in vacuum we have $\mathbf{D}=\epsilon_0\mathbf{E}$ and $\mu_0\mathbf{H}=\mathbf{B}$. Inserting the expression for the electric field in terms of the potentials introduced above into (2.4-6) gives

$$\begin{aligned}\nabla\cdot\mathbf{E}&=\nabla\cdot\left(-\nabla\phi-\frac{\partial\mathbf{A}}{\partial t}\right)=0\\ \iff\qquad -\nabla^2\phi-\nabla\cdot\frac{\partial\mathbf{A}}{\partial t}&=0,\end{aligned} \tag{2.4-87}$$

and similarly for the magnetic field and (2.4-9):

$$\begin{aligned} \nabla \times (\nabla \times \mathbf{A}) &= -\mu_0\epsilon_0 \frac{\partial}{\partial t}\nabla\phi - \mu_0\epsilon_0 \frac{\partial^2}{\partial t^2}\mathbf{A} \\ \iff \quad \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A} &+ \frac{1}{c^2}\frac{\partial}{\partial t}\nabla\phi + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{A} = 0, \end{aligned} \tag{2.4-88}$$

where $\sqrt{\epsilon_0\mu_0} = 1/c$ was used. These two equations don't uniquely determine the potentials however, and if $\mathbf{A}$ and ϕ are a solution then one can show by direct insertion that

$$\mathbf{A}' = \mathbf{A} + \nabla\psi \tag{2.4-89}$$

$$\phi' = \phi - \frac{\partial\psi}{\partial t}, \tag{2.4-90}$$

where $\psi(\mathbf{r}, t)$ is some function, is also a solution. This indefiniteness in the potentials is known as gauge freedom [104]. It turns out that it is always possible to choose the gauge function ψ such that

$$\nabla \cdot \mathbf{A} = 0, \quad \phi = 0, \tag{2.4-91}$$

and this is known as the Coulomb gauge [214].[4] In this gauge (2.4-87) vanishes, and (2.4-88) reduces to the wave equation

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{A} = 0. \tag{2.4-92}$$

To proceed with the quantization one considers a finite volume of space outside of which the electromagnetic field is vanishing. More specifically, let this region be a cube with side lengths L on which periodic boundary conditions are imposed. This volume of space is referred to as the quantization volume, and the periodic boundary conditions, like those found in an optical cavity, give rise to a discrete set of modes. The basis functions for this cavity are of the form

$$u_{\mathbf{k}}(\mathbf{r}) = \hat{e}_{\mathbf{k},\alpha} e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.4-93}$$

where

$$k_i = \frac{2\pi m_i}{L}, \quad m_i \in \mathbb{Z}, \tag{2.4-94}$$

and $\hat{e}_{\mathbf{k},\alpha}$ is a unit vector; the subscript α will later index the polarizations of the field. The basis functions are orthogonal and can be used to express any time-independent function inside the quantization volume, and together with the time-dependent amplitudes $A_{\mathbf{k},\alpha}(t)$ they can be used to decompose the vector potential $\mathbf{A}$:

$$\mathbf{A}(\mathbf{r}, t) = \sum_{\mathbf{k}} \sum_{\alpha=1,2} \hat{e}_{\mathbf{k},\alpha} A_{\mathbf{k},\alpha}(t) e^{i\mathbf{k}\cdot\mathbf{r}} + \hat{e}_{\mathbf{k},\alpha} A^*_{\mathbf{k},\alpha}(t) e^{-i\mathbf{k}\cdot\mathbf{r}}. \tag{2.4-95}$$

[4] The Coulomb gauge more generally corresponds to the condition $-\nabla^2 = \rho/\epsilon_0$, but in the absence of free charges the trivial potential is a solution to this equation.

The expansion above ensures that the vector potential is real valued, and is possible since the wave equation can be solved by separation of variables. Because we chose the Coulomb gauge the potential must satisfy $\nabla \cdot \mathbf{A} = 0$, and this condition can be met by choosing the polarization vectors $\hat{e}_{\mathbf{k},\alpha}$ to be perpendicular to $\mathbf{k}$:

$$\nabla \cdot \hat{e}_{\mathbf{k},\alpha} e^{i\mathbf{k}\cdot\mathbf{r}} = \mathbf{k} \cdot \hat{e}_{\mathbf{k},\alpha} e^{i\mathbf{k}\cdot\mathbf{r}} = 0. \tag{2.4-96}$$

The two polarization components will also be chosen to be orthogonal:

$$\hat{e}_{\mathbf{k},\alpha} \hat{e}_{\mathbf{k},\beta} = \delta_{\alpha,\beta}, \tag{2.4-97}$$

where $\delta_{\alpha,\beta}$ is the Kronecker delta. Since the modes in the expansion are orthogonal they independently satisfy (2.4-92):

$$\begin{aligned} &\frac{1}{c^2}\frac{\partial}{\partial t^2} A_{\mathbf{k},\alpha}(t) e^{i\mathbf{k}\cdot\mathbf{r}} - \nabla^2 A_{\mathbf{k},\alpha}(t) e^{i\mathbf{k}\cdot\mathbf{r}} = 0 \\ \iff \qquad &\frac{1}{c^2}\frac{\partial}{\partial t^2} A_{\mathbf{k},\alpha}(t) + k^2 A_{\mathbf{k},\alpha}(t) = 0, \end{aligned} \tag{2.4-98}$$

where $k = |\mathbf{k}|$. Using the relation $\omega_k = ck$ the solutions to this equation can be written

$$A_{\mathbf{k},\alpha}(t) = A_{\mathbf{k},\alpha} e^{-i\omega_k t}, \quad A^*_{\mathbf{k},\alpha}(t) = A^*_{\mathbf{k},\alpha} e^{i\omega_k t} \tag{2.4-99}$$

and the full vector potential becomes

$$\mathbf{A}(\mathbf{r}, t) = \sum_{\mathbf{k},\alpha} \hat{e}_{\mathbf{k},\alpha} A_{\mathbf{k},\alpha} e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_k t} + \hat{e}_{\mathbf{k},\alpha} A^*_{\mathbf{k},\alpha} e^{-i\mathbf{k}\cdot\mathbf{r} + i\omega_k t}. \tag{2.4-100}$$

We now want to use this solution to find expressions for the electric and magnetic fields. Using $\nabla\phi = 0$ in (2.4-86) the electric field is

$$\begin{aligned} \mathbf{E} &= -\frac{\partial \mathbf{A}}{\partial t} \\ &= \sum_{\mathbf{k},\alpha} i\omega_k \hat{e}_{\mathbf{k},\alpha} \Big[A_{\mathbf{k},\alpha} e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_k t} - A^*_{\mathbf{k},\alpha} e^{-i\mathbf{k}\cdot\mathbf{r} + i\omega_k t} \Big]. \end{aligned} \tag{2.4-101}$$

The magnetic field is found using its definition in terms of the vector potential (2.4-82):

$$\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A} \\ &= \sum_{\mathbf{k},\alpha} i\mathbf{k} \times \hat{e}_{\mathbf{k},\alpha} \Big[A_{\mathbf{k},\alpha} e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_k t} - A^*_{\mathbf{k},\alpha} e^{-i\mathbf{k}\cdot\mathbf{r} + i\omega_k t} \Big]. \end{aligned} \tag{2.4-102}$$

For the free electric field the integral (2.4-55) is simple to evaluate, and the total energy inside the box is

$$W_L = \frac{\epsilon_0}{2} \int_V \big(c^2 \mathbf{B}^2 + \mathbf{E}^2 \big) d^3 r. \tag{2.4-103}$$

The expressions for the electric and magnetic fields inserted into (2.4-103) result in a set of integrals on the form:

$$\int_V e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} d^3 r = V \delta_{\mathbf{k},\mathbf{k}'}, \tag{2.4-104}$$

which can be used to cancel all cross terms between different momenta. After some manipulation [104], one then finds the total energy as

$$W_L = \epsilon_0 \sum_{\mathbf{k},\alpha} \omega_k^2 V (A_{\mathbf{k},\alpha} A^*_{\mathbf{k},\alpha} + A^*_{\mathbf{k},\alpha} A_{\mathbf{k},\alpha}), \tag{2.4-105}$$

written in this unsimplified form to facilitate a comparison with the Hamiltonian for a set of quantum harmonic oscillators:

$$H = \sum_{\mathbf{k},\alpha} \hbar\omega_k \Big(\hat{a}^\dagger_{\mathbf{k},\alpha} \hat{a}_{\mathbf{k},\alpha} + \frac{1}{2}\Big) = \frac{1}{2} \sum_{\mathbf{k},\alpha} \hbar\omega_k (\hat{a}_{\mathbf{k},\alpha} \hat{a}^\dagger_{\mathbf{k},\alpha} + \hat{a}^\dagger_{\mathbf{k},\alpha} \hat{a}_{\mathbf{k},\alpha}). \tag{2.4-106}$$

Noting that the raising and lowering operators of the harmonic oscillator obey the bosonic commutation relations, this comparison motivates promoting the amplitudes of the vector potential to operators:

$$A_{\mathbf{k},\alpha} \to \sqrt{\frac{\hbar}{2\epsilon_0 \omega_k V}} \hat{a}_{\mathbf{k},\alpha}, \qquad A^*_{\mathbf{k},\alpha} \to \sqrt{\frac{\hbar}{2\epsilon_0 \omega_k V}} \hat{a}^\dagger_{\mathbf{k},\alpha}. \tag{2.4-107}$$

The operator for the electric field can then be found the same way as in the classical case (2.4-101):

$$\mathbf{E}(\mathbf{r}, t) = i \sum_{\mathbf{k},\alpha} \sqrt{\frac{\hbar\omega_k}{2\epsilon_0 V}} \hat{e}_{\mathbf{k},\alpha} \Big(\hat{a}_{\mathbf{k},\alpha} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_k t)} - \hat{a}^\dagger_{\mathbf{k},\alpha} e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega_k t)} \Big). \tag{2.4-108}$$

Finally, we need to account for the fact this description will be used with the Hamiltonian for fields inside a dielectric. A comparison between (2.4-77) and (2.4-103) together with the relation $1 + \chi^{(1)} = n^2$ motivates the operator [110]

$$\mathbf{E}_\alpha(\mathbf{r}, t) = i \sum_{\mathbf{k}} \sqrt{\frac{\hbar\omega_k}{2\epsilon_0 n^2 V}} \hat{e}_{\mathbf{k},\alpha} \Big(\hat{a}_{\mathbf{k},\alpha} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_k t)} - \hat{a}^\dagger_{\mathbf{k},\alpha} e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega_k t)} \Big). \tag{2.4-109}$$

where n is the refractive index of the dielectric. Taking the limit of the quantization volume being infinitely long along one dimension and only considering linearly polarized plane waves along this direction allows the electric field operator to be expressed as [104, 220, 221]:

$$E_\alpha(z, t) = i \int_0^\infty d\omega \sqrt{\frac{\hbar\omega}{4\pi\epsilon_0 n^2 c A}} \Big(\hat{a}_\alpha(\omega) e^{i(kz - \omega_k t)} - \hat{a}^\dagger_\alpha(\omega) e^{-i(kz - \omega_k t)} \Big), \tag{2.4-110}$$

where A is the quantization cross section area and we labelled the modes in terms of their frequency instead of their wavenumber. Note that the seeming discrepancy in the units of (2.4-110) comes from the fact that the dimension of $\hat{a}_\alpha(\omega)$ changes due to the substitution

$$\hat{a}_{\omega,\alpha} \to \sqrt{\frac{2\pi c}{L}} \hat{a}_\alpha(\omega) \tag{2.4-111}$$

when replacing the sum with an integral [104, 110], which is motivated by maintaining the standard form of the bosonic commutation relations [222, 223] (see Appendix C.1). It is often more convenient to decompose the electric-field operator in terms of operators that only contain the positive (negative) frequency components:

$$E_\alpha(z,t) = E_\alpha^+(z,t) + E_\alpha^-(z,t), \tag{2.4-112}$$

where

$$\begin{aligned} E_\alpha^+(z,t) &= \pm i \int_0^\infty d\omega \mathcal{E} \hat{a}_\alpha(\omega) e^{i(kz-\omega_k t)}, \\ E_\alpha^-(z,t) &= \pm i \int_0^\infty d\omega \mathcal{E} \hat{a}_\alpha^\dagger(\omega) e^{-i(kz-\omega_k t)}, \end{aligned} \tag{2.4-113}$$

and

$$\mathcal{E} = \sqrt{\frac{\hbar\omega}{4\pi\epsilon_0 n^2 c A}}. \tag{2.4-114}$$

Substituting (2.4-113) into the Hamiltonian 2.4-78 gives

$$H = \frac{2\epsilon_0}{3} \int \chi_{ijk}^{(2)} (E_i^+ + E_i^-)(E_j^+ + E_j^-)(E_k^+ + E_k^-) d^3r. \tag{2.4-115}$$

Here the implicit sum only runs over two Cartesian components, since the problem has been reduced to collinear propagation along one direction with two polarization components. Expanding the product of the electric-field operators gives a total of eight different terms that describe different nonlinear interactions. Which of these processes actually occur is governed by the energy and momentum conservation conditions (2.4-32) and (2.4-33). In practice, more than one process (and its time reversal) rarely fulfils these conditions at any one time, except in situations explicitly engineered to facilitate this. Therefore it is sufficient to only consider one term and its Hermitian conjugate, and neglect the terms that are not of interest. In the case of SPDC the relevant terms are:

$$H_{\text{spdc}} = \frac{2\epsilon_0 d_{\text{eff}}}{3} \int (E_p^+ E_s^- E_i^- + E_p^- E_s^+ E_i^+) d^3r, \tag{2.4-116}$$

where d_{eff} is the effective nonlinear susceptibility for the process [207, 208],[5] and the indices label the pump, signal and idler field respectively. The first term in the Hamiltonian above describes the down-conversion process, and the second term its time reversal (sum-frequency generation). Given that the pump field is strong enough to be treated classically, the corresponding field operators can be replaced by classical fields:

$$\begin{aligned} E_p^+ &\to A_p \int_0^\infty \alpha(\omega_p) e^{i(kz-\omega_p t)} d\omega_p. \\ E_p^- &\to A_p \int_0^\infty \alpha^*(\omega_p) e^{-i(kz-\omega_p t)} d\omega_p, \end{aligned} \tag{2.4-117}$$

[5]The effective nonlinear susceptibility d_{eff} has to be calculated for a specific process, since it depends not only on the medium in question, but also on the polarizations of the interacting fields.

where A_p is the pump amplitude, and $\alpha(\omega_p)$ is called the pump envelope function, and describes the spectral amplitude of the pump light. This is commonly a Gaussian or similar function. Introducing the shorthand notation

$$\Delta k = k_p - k_s - k_i, \qquad \Delta\omega = \omega_p - \omega_s - \omega_i, \tag{2.4-118}$$

the Hamiltonian can be written as

$$H_{\text{spdc}} = -\frac{2\epsilon_0 d_{\text{eff}}}{3}\int d^3r \iiint_0^\infty A_p \mathcal{E}_s \mathcal{E}_i \alpha(\omega_p) e^{i(-\Delta\omega t + \Delta k z)} \hat{a}_s^\dagger \hat{a}_i^\dagger d\omega_p d\omega_s d\omega_i + \text{c.c.}, \tag{2.4-119}$$

and the state generated by the SPDC process can, as before, be expressed as

$$\begin{aligned} |\Psi\rangle &= \exp\left[-\frac{i}{\hbar}\int_0^t H_{\text{spdc}}(t)dt\right]|0,0\rangle_{s,i} \\ &\approx \left(1 - \frac{i}{\hbar}\int_0^t H_{\text{spdc}}(t)dt\right)|0,0\rangle_{s,i}, \end{aligned} \tag{2.4-120}$$

Here the effects of time ordering have been ignored, even though the Hamiltonian does not necessarily commute with itself at different times. This can be justified as long as the nonlinear interaction is relatively weak, and the generated photons do not co-propagate with the pump for an extended amount of time [224, 225]. If the process is pumped by a pulsed laser the integration limits can be taken to infinity, since the pump amplitude is vanishing at long times. Therefore the integral over time reduces to

$$\int_{-\infty}^{+\infty} e^{-i\Delta\omega t} dt = 2\pi\delta(\Delta\omega), \tag{2.4-121}$$

which is simply the energy conservation condition. Together with the fact that the sum-frequency-generation term vanishes when acting on vacuum, the state reads:

$$|\Psi\rangle = |0,0\rangle + \frac{i4\pi\epsilon_0 d_{\text{eff}}}{3\hbar} A_p \mathcal{E}_s \mathcal{E}_i \int d^3r \iint_0^\infty \alpha(\omega_s + \omega_i) e^{i\Delta k z} \hat{a}_s^\dagger \hat{a}_i^\dagger d\omega_s d\omega_i |0,0\rangle, \tag{2.4-122}$$

with $\Delta k = \Delta k(\omega_i, \omega_s) = k(\omega_s + \omega_i) - k(\omega_s) - k(\omega_i)$. This step also made use of the assumption that the bandwidth of the pump field is sufficiently small that $\mathcal{E}_i(\omega_i)$ and $\mathcal{E}_s(\omega_s)$ are slowly varying functions in the frequency range of interest, allowing them to be moved outside the integral. Since the Hamiltonian does not have a radial dependence the integral over the transverse quantization area is trivial, and only the integral along the propagation direction needs to be evaluated. The effect of a nonlinear crystal of finite length L can be modelled by multiplying the nonlinear Hamiltonian with an envelope function $g(z)$ that is zero outside the interval $[0, L]$:

$$\int_{-\infty}^{+\infty} e^{i\Delta k(\omega_s,\omega_s)z} g(z) dz = \int_0^L e^{i\Delta k(\omega_s,\omega_i)z} g(z) dz = \phi(\omega_s, \omega_i). \tag{2.4-123}$$

The function $\phi(\omega_s, \omega_i)$ is called the phase-matching function. Substituting this function in simplifies the expression for the state to

$$|\Psi\rangle = |0,0\rangle + C \iint_0^\infty \alpha(\omega_s + \omega_i)\phi(\omega_s, \omega_i) d\omega_s d\omega_i |0,0\rangle, \tag{2.4-124}$$

where $C = i4\pi\epsilon_0 d_{\text{eff}} A A_p \mathcal{E}_s \mathcal{E}_i / 3\hbar$. The state in (2.4-124) is the general form of a two-photon state generated by down-conversion, where the spectral properties depend on the pump envelope function and the phase-matching function. This will be explored in more depth in the next section.

2.4.3 Phase matching

Phase-matching is the process of satisfying the momentum conservation condition (2.4-33) in a nonlinear optical process. The name comes from the fact that fields with different propagation constants gradually dephase as they propagate. In the case of down-conversion, this can mean that the two-photon amplitude generated at one point of a nonlinear crystal is out of phase with the amplitude generated at another point, causing them to destructively interfere, and this interference effect is simply another way to understand the momentum conservation condition. Since any nonlinear medium exhibits at least some degree of absorption, the Kramers-Kronig relations imply that it will necessarily be dispersive as well [106], and this means that in general

$$k(\omega_p) \neq k(\omega_s) + k(\omega_i). \tag{2.4-125}$$

The most straightforward way of satisfying the momentum conservation condition is called birefringent phase matching, in which one exploits the fact that a uniaxial crystal has two different propagation constants for the ordinary and extraordinary axes of the crystal. In such a crystal an ordinarily polarized ray always sees the refractive index n_o, but an extraordinarily polarized ray sees the refractive index n_θ, which varies continuously between the ordinary and extraordinary refractive indices n_o and n_e, and which is a function of the angle θ between the ray and the optical axis of the crystal. This behaviour can be quantified by considering the so-called index ellipsoid

$$\frac{x_1^2}{n_o^2} + \frac{x_2^2}{n_o^2} + \frac{x_3^2}{n_e^2} = 1, \tag{2.4-126}$$

where x_i are coordinates the crystal axes. This ellipsoid, illustrated in Fig. 2.19, can be used to find the refractive index for a ray propagating through the crystal. More specifically, for a ray propagating along the vector $\hat{u} = (x_1, x_2, x_3)$ the refractive index is given by the distance from the origin to the surface of the ellipsoid along a direction orthogonal to $\hat{u}$ defined by the polarization. This is equivalent to the magnitude of a vector orthogonal to $\hat{u}$ with components that satisfy (2.4-126). To find the refractive index experienced by a ray propagating along $\hat{u}$, consider a circle with radius n_θ that lies in the plane spanned by $\hat{u}$ and the optical axis $\hat{x}_3$:

$$(x_2', x_3) = n_\theta(\cos\theta, \sin\theta), \tag{2.4-127}$$

where x_2' is the coordinate along the axis that lies in the plane spanned by $\hat{x}_1$ and $\hat{x}_2$. As illustrated in Fig. 2.20, the radius n_θ is the one for which this circle intersects the index ellipsoid at a point θ degrees from the $\hat{x}_3$ axis. The intersection can be found by choosing $x_2 = x_2'$ and inserting (2.4-127) into the equation for the ellipsoid (2.4-126):

$$\frac{n_\theta^2 \cos^2\theta}{n_o^2} + \frac{n_\theta^2 \sin^2\theta}{n_e^2} = 1, \tag{2.4-128}$$

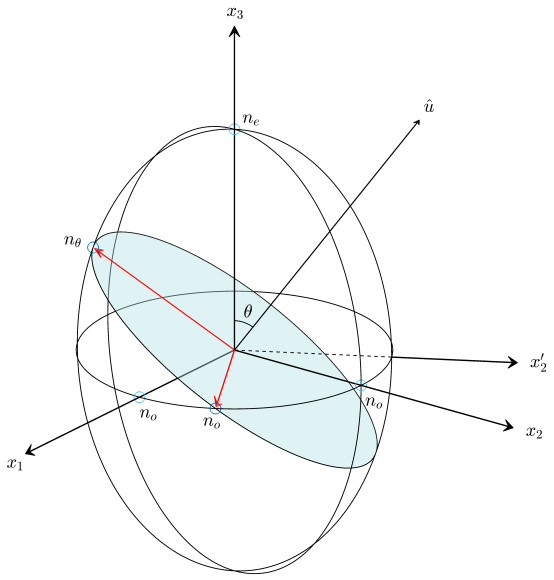


Figure 2.19: **Index ellipsoid.** In a birefringent crystal the refractive index depends on both the polarization and the propagation direction of the light. This dependence can be illustrated as an ellipsoid, with the refractive index being the distance from the origin to the surface. Here this is shown for a uniaxial crystal, with the propagation direction indicated by the vector $\hat{u}$ which lies at an angle θ to the optical axis, and in the plane spanned by x_3 and x_2'. The refractive indices of the ordinary and extraordinary polarized rays, n_o and n_θ respectively, are indicated by the red arrows. When $\hat{u}$ is parallel to the optical axis $n_\theta = n_o$, and when $\hat{u}$ is orthogonal to the optical axis $n_\theta = n_e$.

which can be inverted to give

$$\frac{1}{n_\theta^2} = \frac{\cos^2\theta}{n_o^2} + \frac{\sin^2\theta}{n_e^2}. \tag{2.4-129}$$

To find the refractive index n_θ one typically solves this equation numerically.

The ability to tune the refractive index by choosing the angle θ can be used to satisfy the phase-matching condition, the exact form of which depends on the polarizations of the fields in the process. In terms of the polarizations there are two kinds of processes enabled by birefringent phase matching [226]:

$$e \to o + o, \qquad o \to e + e \tag{2.4-130}$$

$$e \to e + o, \qquad o \to e + o \tag{2.4-131}$$

The first kind of process, in which the signal and idler have the same polarization but are orthogonal to the pump, is called a type-I process, while the second kind of process, where the signal and idler photons have orthogonal polarizations, is called a type-II process. Consider, as an example, a collinear type-I process. For this process the phase-matching condition can be written:

$$\frac{n_\theta(\lambda_p)}{\lambda_p} = \frac{n_o(\lambda_s)}{\lambda_s} + \frac{n_o(\lambda_i)}{\lambda_i}, \tag{2.4-132}$$

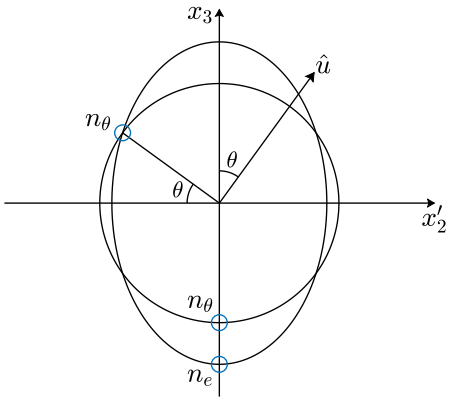


Figure 2.20: **Extraordinary refractive index.** The extraordinary refractive index n_θ for a ray travelling along $\hat{u}$ can be found as the radius of the circle that intersects the index ellipsoid at the point $(n_\theta \cos\theta, n_\theta \sin\theta)$. The ellipse in the figure is a cut of the index ellipsoid in the plane spanned by $\hat{u}$ and the axis x_3.

which in the case where the signal and idler are degenerate reduces to

$$n_\theta(\lambda_p) = n_o(\lambda_s). \tag{2.4-133}$$

An example of phase matching for this process is shown in Fig. 2.21. By considering non-collinear geometries one can also achieve phase matching for a wider set of parameters, as several of the refractive indices in the process can be tuned independently.

Birefringent phase matching is usually achieved in a uniform bulk crystal, with a constant nonlinearity. For such crystal the envelope function $g(z)$ in the phase-matching function $\phi(\omega_s, \omega_i)$ from (2.4-123) is a simple rectangular function, and the phase-matching function can be found explicitly:

$$\begin{aligned}
\phi(\omega_s, \omega_i) = \int_0^L e^{i\Delta k z} dz &= \frac{e^{i\Delta k L} - 1}{i\Delta k} \\
&= \frac{2}{\Delta k} e^{i\Delta k L/2} \frac{e^{i\Delta k L/2} - e^{-i\Delta k L/2}}{2i} \\
&= \frac{2}{\Delta k} e^{i\Delta k L/2} \sin\left(\frac{\Delta k L}{2}\right) \\
&= L e^{i\Delta k L/2} \mathrm{sinc}\left(\frac{\Delta k L}{2}\right).
\end{aligned} \tag{2.4-134}$$

The sinusoidal dependence on the crystal length L means that signal intensity is not a monotonic function of the interaction length, and for a crystal of length $L = 2\pi/\Delta k$ the probability to generate a pair of photons will be exactly zero. The length of the crystal also affects the bandwidth of the generated light, as it will set the width of the sinc function in terms of Δk. Due to dispersion, this in turn implies a limit on the spectral bandwidth.

It is often desirable to have a phase-matching function that grows monotonically with the length of the crystal, since this allows for a higher interaction strength in an

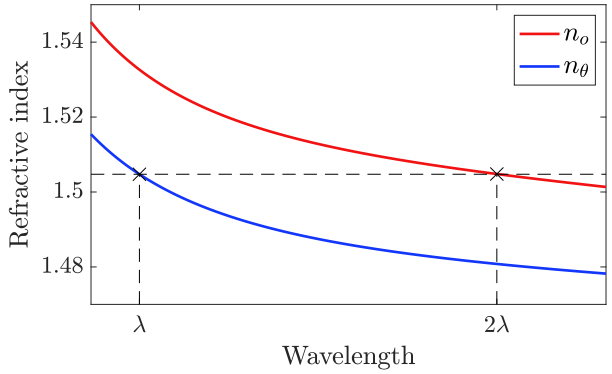


Figure 2.21: **Birefringent phase matching.** By tuning the angle θ between the light and the optical axis of the nonlinear crystal momentum conservation can often be achieved. The figure shows a degenerate, collinear SPDC / SFG process in KDP, where $\lambda_s = \lambda_i = 694\,\text{nm}$ and $\lambda_p = 347\,\text{nm}$. This process is phase matched for $\theta = 50.5°$.

otherwise very weak nonlinear process. This can be achieved by an appropriate choice of envelope function $g(z)$. More concretely, by adding a modulation to the rectangular envelope. Such a modulation amounts to changing the magnitude or sign of the effective nonlinear coefficient d_{eff}. The utility of such a modulation can be understood visually, by considering the integrand in (2.4-123) as a vector in the complex plane, called a phasor. Without any modulation the addition of infinitesimally small phasors that the integral represents will lead to the tracing out of a circle in the complex plane, and this is why the phase-matching amplitude periodically returns to zero. A modulation can remedy this by redirecting the phasors as they begin to decrease the phase-matching amplitude. This is illustrated in Fig. 2.22.

The perhaps mathematically most simple modulation is a periodic one using a complex exponential:

$$g(z) = \text{rect}\left(\frac{z - L/2}{L}\right) e^{-i2\pi z/\Lambda} = r(z)f(z), \tag{2.4-135}$$

where $r(z) = \text{rect}((z - L/2)/L)$ is a rectangular function that is zero outside the interval $[0, L]$ and $f(z) = e^{-i2\pi z/\Lambda}$. Here $1/\Lambda$ is the spatial frequency of the modulation. The phase-matching function for this choice of envelope is, as before, given by its Fourier transform:

$$\begin{aligned} \phi(\omega_s, \omega_i) &= \int_{-\infty}^{+\infty} r(z)f(z)e^{i\Delta kz}dz \\ &= \int_0^L e^{i(\Delta k - 2\pi/\Lambda)z}dz. \end{aligned} \tag{2.4-136}$$

Direct comparison with (2.4-134) shows that

$$\phi(\omega_s, \omega_i) = L\exp\left[i\frac{L}{2}\left(\Delta k - \frac{2\pi}{\Lambda}\right)\right]\text{sinc}\left[\frac{L}{2}\left(\Delta k - \frac{2\pi}{\Lambda}\right)\right]. \tag{2.4-137}$$

The same result can be found by taking convolution of the Fourier transformed functions:

$$\phi(\omega_s, \omega_i) = \tilde{r}(\Delta k) * \tilde{f}(\Delta k), \tag{2.4-138}$$

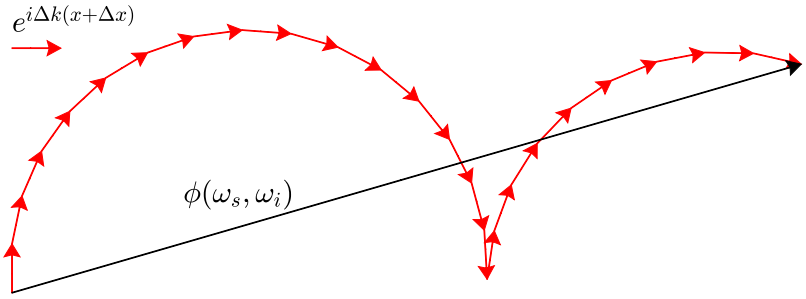


Figure 2.22: **Quasi-phase-matching.** The figure shows how the phase-matching function $\phi(\omega_s, \omega_i)$ grows throughout the length of a quasi-phase-matched crystal with a square-wave modulation. The red arrows are called phasors, and are vectors that indicate the relative phase between the down-conversion process at different points in the crystal. In this discrete picture they are given by $f(z)e^{i\Delta k(x+\Delta x)}$. The phase-matching amplitude, indicated by the black arrow, is given as the sum of the phasors. When the momentum mismatch Δk is nonzero the phasors trace out a circle that gradually brings the phase-matching amplitude back to zero. In a QPM crystal this is avoided by periodically alternating the sign of the nonlinear coefficient, which rotates the phasors by 180°. This leads to a phase-matching amplitude whose magnitude grows monotonically throughout the crystal.

where the tilde denotes the Fourier transformed function, and using the shift property of the delta function:

$$h(x) * \delta(x - y) = h(x - y), \tag{2.4-139}$$

since

$$\tilde{f}(\Delta k) = \delta\Big(\Delta k - \frac{2\pi}{\Lambda}\Big). \tag{2.4-140}$$

If the spatial frequency Λ is chosen to be

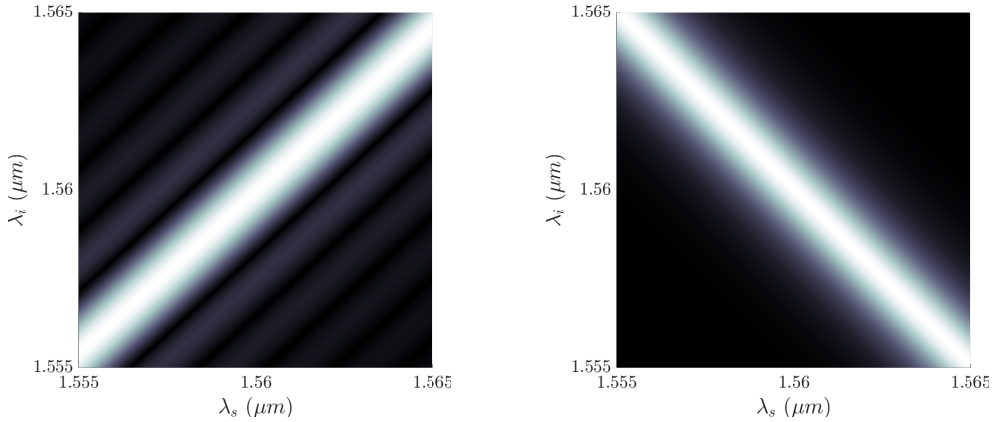
$$\frac{1}{\Lambda} = \frac{\Delta k}{2\pi}, \tag{2.4-141}$$

the sinc term in (2.4-137) vanishes and the phase-matching amplitude grows linearly with the crystal length. This way of cancelling the momentum mismatch by way of a periodic structure is called quasi-phase-matching (QPM) [206, 227]. The idea of QPM dates back further than birefringent phase matching, but took longer to be realised due to the challenge of making the required structures. In practice it is hard to engineer a continuous modulation of the effective nonlinearity in a crystal, and one therefore uses a square-wave modulation instead, for which precision manufacturing techniques exist. Assuming that the square wave is an odd function with unity amplitude it can be written as the exponential Fourier series:

$$\mathrm{sw}(z) = \sum_{m=\mathrm{odd}} \frac{2i}{\pi m} e^{-i2\pi m z/\Lambda}. \tag{2.4-142}$$

Performing the same steps as for the continuous modulation the phase-matching function becomes:

$$\phi(\omega_s, \omega_i) = \frac{2iL}{\pi m} \sum_{m=\mathrm{odd}} \exp\left[i\frac{L}{2}\left(\Delta k - \frac{2\pi m}{\Lambda}\right)\right] \mathrm{sinc}\left[\frac{L}{2}\left(\Delta k - \frac{2\pi m}{\Lambda}\right)\right], \tag{2.4-143}$$



(a) $|\phi(\omega_s, \omega_i)|$ (b) $|\alpha(\omega_s, \omega_i)|$

Figure 2.23: **Phase-matching function and pump envelope.** Figure (a) shows an example of a phase-matching function for a quasi-phase-matched crystal. The process in question is degenerate type-II SPDC in periodically-poled KTP at $\lambda_s = 1560\,\text{nm}$, in a crystal with a poling period of $\lambda = 46.06\,\mu\text{m}$ and a temperature of $T = 52\,°\text{C}$. The phase-matching function exhibits a distinctive decaying oscillatory behaviour originating from the sinc-function. In figure (b) the pump envelope function, taken to be a hyperbolic secant with a pulse duration of 2.5 ps, is shown. This function is constant along the lines $\lambda_i - \lambda_p = \lambda$, representing points with the same energy, and has the hyperbolic secant shape along the transverse lines given by $\lambda_i = \lambda_s$.

and the phase-matching condition is

$$k_p - k_s - k_i - \frac{2\pi m}{\Lambda} = 0. \tag{2.4-144}$$

Generally it is only possible to satisfy this equation for one choice of m, and since the higher-order terms in the Fourier series are suppressed one typically chooses Λ so that the process is phase matched for $m = 1$. Crystals with this kind of square wave modulation in the effective nonlinearity are called periodically poled crystals. The use of QPM allows for phase matching in many more scenarios than birefringent phase matching would, and can also enable so-called type-0 processes in which the pump, signal and idler all have the same polarization:

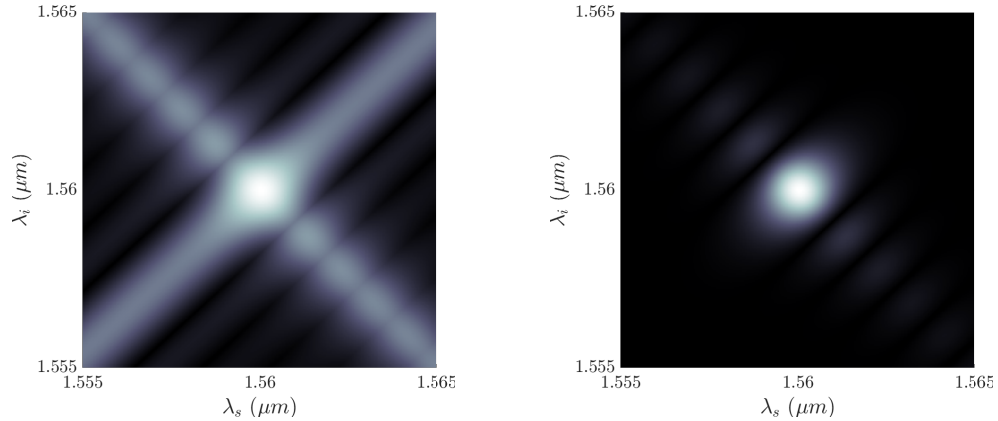
$$o \to o + o, \qquad e \to e + e. \tag{2.4-145}$$

These processes can often have higher nonlinear response, leading to very bright down-conversion sources [228, 229, 230].

The joint spectral amplitude (JSA) of the two-photon state generated by the down-conversion process is given by the product of the pump envelope function and the phase-matching function:

$$\text{JSA} = \alpha(\omega_s, \omega_i)\phi(\omega_s, \omega_i). \tag{2.4-146}$$

The absolute square of this function is known as the joint-spectral intensity, or JSI: $\text{JSI} = |\text{JSA}|^2$. Engineering the two-photon spectrum is a critical part of the design of SPDC sources. An example of QPM phase-matching and pump envelope functions



(a) $|\alpha(\omega_s, \omega_i)| + |\phi(\omega_s, \omega_i)|$ (b) $|JSA| = |\alpha(\omega_s, \omega_i)\phi(\omega_s, \omega_i)|$

Figure 2.24: **Phase matching and joint-spectral amplitude.** The two figures display the same process as Fig. 2.23. Figure (a) shows the sum of the pump envelope and phase-matching functions. To create a separable joint-spectral amplitude, and therefore uncorrelated two-photon spectrum, these functions should intersect each other at 90° [209]. In (b) the joint-spectral amplitude is shown. This function is the product of the pump envelope function and the phase-matching function from Fig. 2.23. The process achieves a nearly circular central two-photon spectrum, however the side lobes originating from the higher-order maxima of the sinc functions give rise to spectral correlations between the signal and idler photons.

is shown in Fig. 2.23. For sources intended for multi-photon experiments the ideal two-photon spectrum is one that is factorizable into two single-photon spectra. This condition implies that the two photons are separable in frequency, and conversely a two-photon spectrum that cannot be factored means that the photons have some degree of entanglement in frequency. If two photons from two independent frequency-correlated sources interfere then the local description of both photons is that they are in spectrally mixed states, and this decreases the visibility of the Hong-Ou-Mandel interference. More concretely, the visbility can be expressed in terms of the spectral purities P_a and P_b of the two photons as [231, 232]

$$\mathcal{V} = \frac{P_a + P_b - ||\rho_a - \rho_b||_{\mathrm{F}}^2}{2}, \tag{2.4-147}$$

where $||\cdot||_{\mathrm{F}}^2$ is the Frobenius norm and $||\rho_a - \rho_b||_{\mathrm{F}}^2$ is a measure of the indistinguishablility of the two photons. The visibility above corresponds to the following definition in terms of observed photon counts N:

$$\mathcal{V} = \frac{N_{\max} - N_{\min}}{N_{\max}}. \tag{2.4-148}$$

When using a periodic modulation the spectral purity is limited to around $P \approx 0.82$ for photons with nearly indistinguishable marginal spectra, due to fact that the sinc function gives rise correlated side lobes in the two-photon spectrum [233]. This is illustrated in Fig. 2.24. For states with these spectral side lobes the reduced interference visibility between two SPDC sources has an intuitive explanation, since the two-photon spetrum

can be approximated by a central uncorrelated peak, and a maximally entangled two-frequency state:

$$|\Psi\rangle = \alpha\,|\omega_0,\omega_0\rangle + \beta\frac{|\omega_1,\omega_2\rangle + |\omega_2,\omega_1\rangle}{\sqrt{2}}. \tag{2.4-149}$$

Hong-Ou-Mandel interference between two different sources can therefore, when post-selecting on the frequency modes in the side lobes, be understood as entanglement swapping in frequency, and since entanglement swapping is probabilistic the photons will display imperfect bunching. However, this means that if two photons from independent spectrally correlated sources bunch, then the remaining two photons will be projected onto a frequency entangled state and will exhibit perfect interference [234, 235]. Unfortunately this phenomenon is hard to observe due to terms arising from double-pair emission from either SPDC source, which can also give rise to four-fold coincidence events that cannot be distinguished from cases where one pair was generated by each source. HOM interference between frequency entangled photons will be discussed in more detail in Section 2.4.5.

The $\sim 82\,\%$ visibility bound of periodically poled crystals can be overcome by using more complicated nonlinearity profiles, and sources of this type have demonstrated near unity spectral purity [236, 237, 238, 239]. The downside is that these structures are more challenging to manufacture, and additionally exhibit a lower phase-matching amplitude overall, leading to lower source brightness [240].

2.4.4 SPDC sources

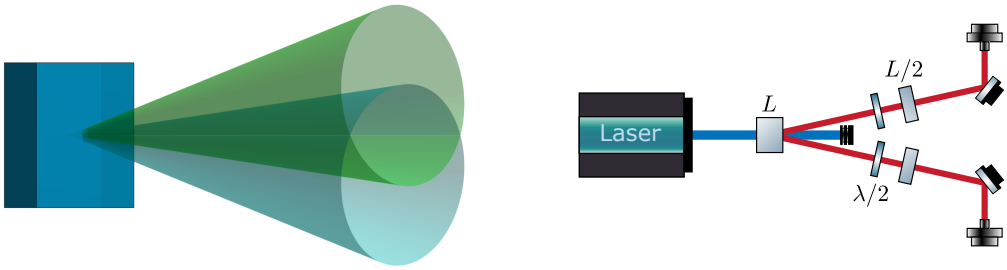
The design of SPDC sources varies depending on the intended application of the single photons. For example, as mentioned in the previous section, sources of single photons intended for complex multi-photon tasks are required to generate spectrally uncorrelated photon pairs, while in other contexts frequency correlated photons can be considered a resource [241, 242]. In many applications the heralded single-photon purity $g^{(2)}(0)$ is an important figure of merit, despite the fact that measurements of heralded $g^{(2)}(0)$ do not fully quantify the second-order coherence of the heralded state, but is mainly a measure of the correlations between the signal and idler photons [243, 244]. In SPDC sources $g^{(2)}(0)$ is increased when the probability of creating more than one photon pair increases, and the Taylor expansion in (2.4-49) can no longer be truncated after the first-order term. A factor in the source design is the lack of dependence of the $g^{(2)}(0)$ on loss. In a linear optical system loss can be modelled by a beamsplitter with transmission and reflection coefficients $\mathcal{T}$ and $\mathcal{R}$ respectively [104]:

$$\hat{a}^\dagger \mapsto \mathcal{T}\hat{a}^\dagger + \mathcal{R}\hat{b}^\dagger, \tag{2.4-150}$$

and tracing out the loss mode b. In this model it's easy to see that loss does not affect the $g^{(2)}(0)$:

$$g_a^{(2)}(0) = \frac{\langle \hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a}\rangle}{\langle \hat{a}^\dagger\hat{a}\rangle^2} \mapsto \frac{\langle \mathcal{T}^*\hat{a}^\dagger\mathcal{T}^*\hat{a}^\dagger\mathcal{T}\hat{a}\mathcal{T}\hat{a}\rangle}{\langle \mathcal{T}^{*2}\hat{a}^\dagger\mathcal{T}^2\hat{a}\rangle^2} = \frac{|\mathcal{T}|^4\langle \hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a}\rangle}{|\mathcal{T}|^4\langle \hat{a}^\dagger\hat{a}\rangle^2}. \tag{2.4-151}$$

The $g^{(2)}(0)$ therefore only depends on the SPDC process itself, and one cannot compensate high losses by pumping the process harder. As consequence, the efficiency



(a) Emission cones (b) Top-down view

Figure 2.25: **Type-II BBO source.** (a) In a non-collinear type-II BBO source the single photon emission occurs along two cones, each one containing photons with a definite polarization, either H or V. At the intersection points of the two cones photons of both polarizations are created. By collecting only the emission from these two intersection points the polarization in both spatial modes becomes undetermined, but is anti-correlated since there is always one H and one V photon produced. The result is the entangled state $|\Psi\rangle = (|H,V\rangle + e^{i\varphi}|V,H\rangle)\sqrt{2}$. (b) shows a top-down view of a typical source, which in addition to the main BBO-crystal of length L contains two compensation crystals with half the length of the first one (see main text and Fig. 2.26).

of the single-photon collection in the source is of high importance. Optimizing the collection efficiency of SPDC sources is a non-trivial problem that has been studied extensively [245, 246, 247, 248, 249]. Interestingly, there is often a trade-off between the absolute brightness of a photon source and the correlated coupling efficiency, or heralding efficiency, meaning the conditional probability to collect say the idler photon given that the signal photon was collected [250, 251]. This effect can be qualitatively understood in terms of momentum correlations: a collimated pump beam has a well defined momentum, and consequently the down-converted photons exhibit strong momentum correlations. A collection geometry that attempts to mode match the pump will similarly only collect photons with well defined momenta, and this acts as a filter on the momentum distribution of the photons. However, the momentum correlations mean that if one photon is transmitted through this filter the partner photon will be too, and thus the coupling is highly correlated. In the case of a focused pump the situation is the opposite: the momentum of the pump photons is less well defined, erasing the momentum correlations in the signal-idler pair, and while a similar focusing geometry for the collection optics will allow for a broad momentum bandwith, and therefore higher photon rates overall, the coupling for the signal and idler will be largely uncorrelated.

SDPC sources can broadly speaking be classified in terms of their spectral properties: degenerate / non-degenerate, their geometry: collinear / non-collinear, their type: type-0 / type-I / type-II, and whether or not they generate photon pairs that are entangled in some degree of freedom. One of the most commonly used SPDC sources is a non-collinear degenerate type-II polarization entangling source using a beta-Barium Borate (BBO) crystal, first proposed and demonstrated in Ref. [252], illustrated in Fig. 2.25. The phase-matching condition leads to single-photon emission along two cones, one containing H-polarized photons and the other one V-polarized ones. By tilting the optical axis of the crystal with respect to the pump beam these cones can be made to overlap. In the two overlap regions photons of both polarizations are generated, and by

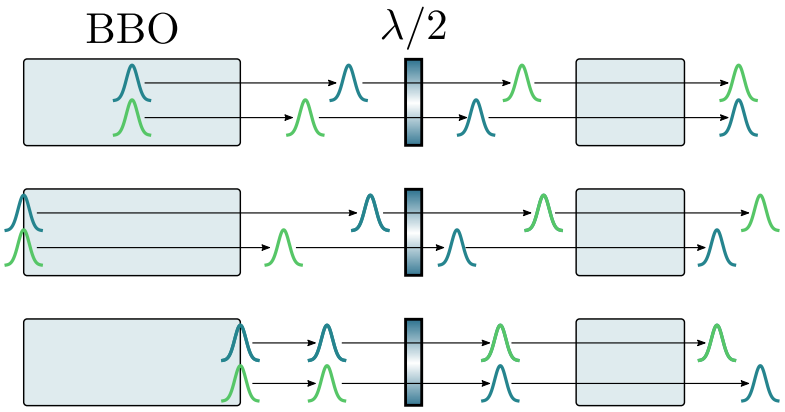


Figure 2.26: **Temporal walk-off compensation.** In a birefringent crystal photons with different polarizations propagate at different speeds and their wavepackets gradually separate in time. In SPDC the single photons are created coherently over the entire length of the nonlinear crystal, and the total amount of temporal walk-off will be different for photon amplitudes corresponding to down-conversion at different points in the crystal. The figure shows how this can be corrected with the use of a second birefringent crystal half the length of the first one, but that is otherwise identical. The top row of the figure shows photons created in the middle of the crystal, with the colors indicating two orthogonal polarizations. After propagating through half the crystal, a half-wave plate exchanges the two polarizations, and the second crystal applies exactly the same temporal delay to the initially undelayed photon. The second and third rows show pairs of photons created symmetrically around the centre of the crystal. For these wavepackets the second crystal does not fully cancel the walk-off, however it still erases the information about the photon polarization in the arrival time of the photons, as in every time bin there is an equal probability amplitude for each polarization.

balancing the coupling of both polarizations a $|\Phi^{\pm}\rangle$ state can be obtained directly from the source.

Inside the BBO crystal the extraordinarily polarized photons in the pump or the signal/idler experience both transverse and temporal walk-off. The transverse walk-off negatively affects the simultaneous coupling of both polarizations, an effect which can be partially corrected with a second crystal with half the length of the main crystal. The transverse walk-off also affects the transverse mode of the down-converted photons, and gives rise to an asymmetric broadening of the angular photon distributions [253]. The use of elliptic pump modes can help compensate for the mode distortion caused by the walk-off [254]. More problematic, however, is the temporal walk-off that occurs inside the crystal, since this has the effect of labelling the photons by their arrival time:

$$|H,t_0\rangle\,|V,t_0\rangle \mapsto |H,t_1\rangle\,|V,t_2\rangle\,, \tag{2.4-152}$$

and this destroys the entanglement that would otherwise be generated by the source:

$$\frac{|H,t_0\rangle\,|H,t_{,0}\rangle + |V,t_0\rangle\,|V,t_0\rangle}{\sqrt{2}} \mapsto |\Psi\rangle = \frac{|H,t_1\rangle\,|H,t_{,1}\rangle + |V,t_2\rangle\,|V,t_2\rangle}{\sqrt{2}}. \tag{2.4-153}$$

Or, more precisely, the state is entangled in polarization and time, and if $\Delta t = t_1 - t_2$ is too short to be resolved by the detectors, the temporal modes are traced over in

detection and the polarization state becomes

$$\rho_{\text{pol}} = \text{Tr}_t\big[\,|\Psi\rangle\langle\Psi|\,\big] = \frac{|H,H\rangle\langle H,H| + |V,V\rangle\langle V,V|}{2}, \tag{2.4-154}$$

which is not a pure state. In practice the single photons are created coherently over the entire length of the crystal and are therefore in a superposition of many temporal modes. Somewhat surprisingly the temporal walk-off can, in contrast to the transverse walk-off, be exactly cancelled by a second crystal with half the length of the main one [252, 255]. This process is illustrated in Fig. 2.26. Photons created in the middle of the crystal experience half a crystal's worth of temporal walk-off, and the second crystal applies the same walk-off to the initially undelayed photon. Pairs of photon-pair amplitudes created symmetrically around the center of the crystal do not have their walk-off fully cancelled, however they still interfere pairwise. This intuitive argument is slightly misleading because it does not explicitly consider the delay of the pump, something which becomes clear when making the argument more concrete. Consider photon pairs created at the beginning as well as at the end of the crystal. The pair created at the beginning sees one polarization delayed by 2τ:

$$|H,0\rangle_{x=0} \mapsto |H,0\rangle_{x=0}\,, \qquad |V,0\rangle_{x=0} \mapsto |V,2\tau\rangle_{x=0}\,, \tag{2.4-155}$$

while a pair created at the end of the crystal does not acquire any relative delay between its polarization components, however it does have an initial delay given by the pump delay through the crystal τ_p:

$$|H,\tau_p\rangle_{x=L}\,, \qquad |V,\tau_p\rangle_{x=L}\,. \tag{2.4-156}$$

The second crystal delays the opposite polarization component by τ:

$$\begin{aligned} |H,0\rangle_{x=0} &\mapsto |\psi_1\rangle = |H,\tau\rangle_{x=0}\,, & |V,2\tau\rangle_{x=0} &\mapsto |\psi_3\rangle = |V,2\tau\rangle_{x=0}\,, \\ |H,\tau_p\rangle_{x=L} &\mapsto |\psi_2\rangle = |H,\tau_p+\tau\rangle_{x=L}\,, & |V,\tau_p\rangle_{x=L} &\mapsto |\psi_4\rangle = |V,\tau_p\rangle_{x=L}\,. \end{aligned} \tag{2.4-157}$$

The pairs of amplitudes that should, according to the argument, interfere in the diagonal basis are $|\psi_1\rangle$ and $|\psi_4\rangle$, as well as $|\psi_2\rangle$ and $|\psi_3\rangle$. However, it is clear that this only happens if $\tau = \tau_p$, meaning that pump delay through the main crystal needs to be equal to the single-photon delay through the half-length crystal. For a degenerate source, meaning $\lambda_p = 2\lambda_s = 2\lambda_i$, this turns out to actually be the case, since the phase-matching condition implies:

$$\frac{n_p}{\lambda_p} = \frac{n_s}{\lambda_s} + \frac{n_i}{\lambda_i} \iff n_p = \frac{n_s + n_i}{2}, \tag{2.4-158}$$

which means that, for weak birefrigence, the pump propagates as the center of mass of the two-photon wavepacket and experiences the average delay, which is indeed τ. This argument assumes dispersion is small enough that the group velocity of the photons equals the phase velocity.

Type-II sources of this ilk have been iterated on many times, and continue to see use in state-of-the-art experiments [139, 256]. One notable derivative is the so-called beam-like source in which the phase-matching condition is engineered such that the two

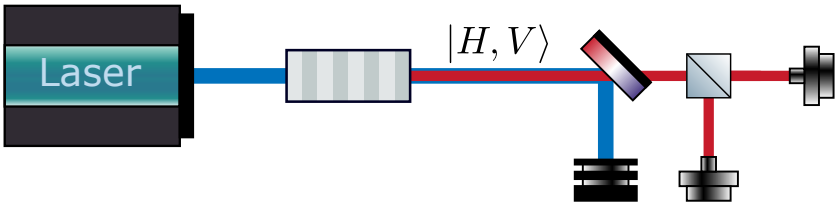


Figure 2.27: **Probabilistically entangling type-II source.** A QPM collinear type-II source generating the state $|H,V\rangle$ can also produce post-selected entanglement with the help of a single balanced beamsplitter. The state after the BS is $|\Psi\rangle = 1/2(\hat{a}_H^\dagger \hat{b}_V^\dagger + \hat{a}_V^\dagger \hat{b}_H^\dagger + \hat{a}_H^\dagger \hat{a}_V^\dagger + \hat{b}_H^\dagger \hat{b}_V^\dagger)\,|0\rangle$. Post-selecting on one photon exiting in each port of the BS therefore generates a $|\Psi^+\rangle$ Bell state with 50 % probability. While less efficient than directly entangling sources, the simplicity of this scheme can allow for high fidelity entanglement.

emission cones collapse down to a single line [257, 258]. This leads the single-photons being emitted into a Gaussian mode, which can significantly increase the collection efficiency of the source [259]. The process is still non-collinear, and since the signal and idler emissions do not spatially overlap in this configuration the source does not directly generate entanglement, however this kind of source can still be made to generate polarization entanglement by combining two crystals in a so-called sandwich configuration [260].

As mentioned in Section 2.4.3, down-conversion sources using birefringent phase matching in bulk crystals have many limitations in terms of brightness and spectral properties, since for a given crystal there is a fixed set of processes and geometries that are phase matched. The use of quasi-phase matching, in contrast, offers much greater flexibility in the source design. Sources based on QPM are generally collinear, although exceptions exist [261]. Similarly to the beamlike sources, collinear sources do not directly generate polarization entanglement. The simplest way to generate entangled photon pairs in such a source is to do it post-selectively, by placing a single beamsplitter after the source [262]. This type of geometry is shown in Fig. 2.27.

More common, however, is to place the nonlinear crystal inside a Sagnac interferometer, and superpose two propagation directions of the pump through the crystal. This superposition of down-conversion processes can then be used to directly produce a two-qubit entangled state, and the inherent stability of the Sagnac geometry ensures that the entangled state has a well defined phase. The use of this geometry also removes the need for walk-off compensation. The first proposed source of this kind used a Sagnac interferometer with a regular balanced beamsplitter [263]. This source configuration generates entanglement post-selectively when the two photons exit in different ports of the beamsplitter, which, in rather anti-climactic fashion, happens with the same 50 % probability as the source in Fig. 2.27, thereby negating the entire point of the interferometer. Later designs remedied this by using a polarizing beamsplitter for the Sagnac interferometer [264, 265]. This configuration is illustrated in Fig. 2.28, and was used for all the experiments in this thesis.

Sagnac sources can also be used in configurations with more than one crystal. This can for example facilitate frequency-degenerate polarization entanglement in type-0 sources [266], or non-degenerate polarization entanglement without the need for walk-off compensation [267]. The SPDC sources discussed above all generate discrete variable

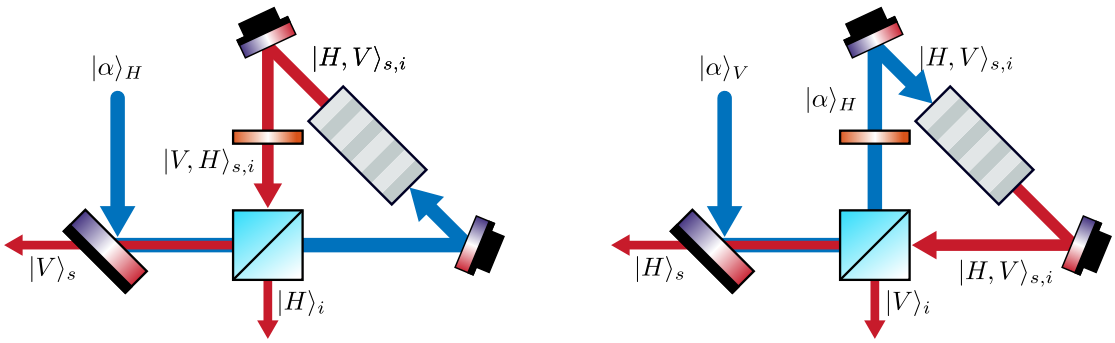


Figure 2.28: **Type-II polarization entangling Sagnac source.** A nonlinear crystal phase matched for type-II collinear SPDC can be made to produce polarization entangled states by placing it inside a Sagnac interferometer based on a polarizing beamsplitter. The two figures show the photon generation process for two different, orthogonal pump polarizations. (a) The H-polarized pump is transmitted through the PBS, and produces an $|H,V\rangle_{s,i}$ photon pair. A HWP then flips the polarization of both photons, after which they are separated on the PBS. Finally, the dichroic mirror separates the signal photon from the pump beam. (b) The V-polarized pump is reflected by the PBS, and is converted to horizontal polarization by the HWP to ensure phase matching. An $|H,V\rangle_{s,i}$ photon pair is then produced, which as before is separated into two spatial modes by the PBS, but for this pump direction the polarization of the two single photons after the PBS is the opposite of that in (a). By pumping the source with a diagonally polarized pump $|\alpha\rangle_H + |\alpha\rangle_V$ the two diagrams above are coherently superposed, giving the output state $|\Psi^-\rangle = (|H,V\rangle_{s,i} - |V,H\rangle_{s,i})/\sqrt{2}$. The inherent phase stability of the Sagnac geometry ensures the coherence of the process, and in practice the phase in the entangled state above depends on the phase in the pump superposition and is therefore tunable. Note that the temporal walk-off between the two polarizations inside the crystal is cancelled by the half-wave plate, since it ensures that the signal and idler always end up in the same ports, independently of the pumping direction.

entanglement between the two down-converted photons. This entanglement generation is enabled by the fact that the process is coherent, and two different emission processes can be superposed. Since entanglement is a highly useful resource in quantum information, and one which cannot be deterministically generated using linear optics alone, the ability to directly generate entangled states is one of the key advantages of SPDC over most current solid-state-emitter-based sources. A good question to ask then is if SPDC sources can be used to efficiently generate entanglement between more than two photons, through some clever optical arrangement. It's quite easy to understand that this is in fact not possible, since any terms containing more than two photons will be of the form:

$$H^n |0\rangle = \left[\text{poly}_2(\hat{a}^\dagger, \ldots, \hat{b}^\dagger)\right]^n |0\rangle , \qquad n \geq 2, \tag{2.4-159}$$

which directly implies that these states can be factored. This problem can also be understood as arising from the fact that the emission of two independent photon pairs is uncorrelated. Direct generation of higher-dimensional qubit states therefore requires higher-order nonlinear processes. For example, a superposition of three different third-

order processes can be used to generate maximally entangled three-qubit states:

$$H = \hat{a}_H^\dagger \hat{b}_H^\dagger \hat{c}_V^\dagger + \hat{a}_H^\dagger \hat{b}_V^\dagger \hat{c}_H^\dagger + \hat{a}_V^\dagger \hat{b}_H^\dagger \hat{c}_H^\dagger + \text{c.c.} \mapsto \frac{|H,H,V\rangle + |H,V,H\rangle + |V,H,H\rangle}{\sqrt{3}}. \tag{2.4-160}$$

While such processes have been investigated theoretically [268, 269, 270, 271, 272, 273], they remain experimentally challenging and have not been observed. This is largely due to the difficulty of achieving phase matching for four different fields simultaneously, and at a larger frequency separation than in second-order SPDC. Additionally, the much smaller third-order nonlinear susceptibility also poses a challenge, however this can to some extent be compensated for by using higher pump power, or potentially by using more exotic materials [274].

An alternative to third-order nonlinearities is cascaded down-conversion. In this type of process a down-converted photon is sent to a second nonlinear crystal in which it undergoes down-conversion, generating a three photon state. If the state from the first-stage down-conversion is for example a Bell-state:

$$\frac{\hat{a}_H^\dagger \hat{b}_H^\dagger + \hat{a}_V^\dagger \hat{b}_V^\dagger}{\sqrt{2}} |0\rangle, \tag{2.4-161}$$

then acting on this state with the SPDC-Hamiltonian:

$$H = \hat{c}_H^\dagger \hat{b}_H^\dagger \hat{b}_H + \hat{c}_V^\dagger \hat{b}_V^\dagger \hat{b}_V + \text{c.c.}, \tag{2.4-162}$$

directly generates the three-qubit GHZ-state

$$\frac{|H,H,H\rangle + |V,V,V\rangle}{\sqrt{2}}. \tag{2.4-163}$$

These types of sources have been experimentally demonstrated [275, 276], however due to the almost vanishing probability of a given photon undergoing down-conversion the achievable count rates are prohibitively low. It should furthermore be noted that cascaded down-conversion is fundamentally different from a true third-order process, and cannot be used to directly generate maximally entangled three-qubit states such as the W-state in (2.4-160).

Similar to how linear-optical gates can be used to probabilistically entangle photons, it is possible to construct source geometries that produce heralded entangled states [277], for example Bell states [278, 279]. The drawback of these schemes is that they are resource heavy in terms of the total amount of needed photons. For instance, to herald the generation of a single Bell state a total of six photons is needed. Generation of larger entangled states then requires operation in a regime where the gain is sufficiently strong for the probability of unwanted higher-order emission terms to be comparable to, or greater than the generation probability for the desired process order, leading to states that do not well approximate single-photon states. Finally, SPDC sources can also be made to directly generate post-selected (unheralded) entanglement. This is done by arranging multiple sources in such a way that the emission from one source passes through a second source, and ultimately erases the information about which source

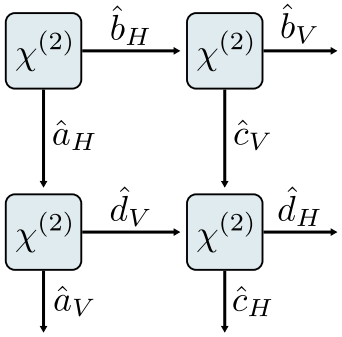


Figure 2.29: **Entanglement by path identity.** Four independent SPDC sources are arranged in such a way that the photons generated by the top-left source propagates through the sources on the anti-diagonal, and one photon from both of these two sources in turn propagates through the source in the bottom right. This arrangement ensures that there are only two possible emission patterns that generate one photon in four modes a, b, c, d, and these are that either the two sources on the diagonal generate one pair of photons each, or the two sources on the anti-diagonal generate photon pairs. The first emission pattern gives the state $|H, H, H, H\rangle$ while the second one generates $|V, V, V, V\rangle$. Coherent pumping of all four sources therefore generates the four-qubit GHZ state when post-selecting on detection of one photon in each mode.

fired. The most trivial example of this is a collinear source consisting of two crystals where one emits for example the state $|H, V\rangle$ and the second one $|V, H\rangle$, and if the emission modes of the two sources are identical the state $(|H, V\rangle + e^{i\varphi}|V, H\rangle)/\sqrt{2}$ is produced. This particular example is of course equivalent to the entangling sources already discussed, however the concept, dubbed entanglement by path identity [280, 281, 282], can be extended to more complex scenarios. An example of this is shown in Fig. 2.29, where an arrangement of four crystals can be made to generate a four-qubit GHZ state, post-selected on exactly one photon being each spatial mode. That this is in fact the case is easy to understand visually. Mathematically, it comes from the fact that when squaring the Hamiltonian[6]

$$H_{GHZ} = \hat{a}_H^\dagger \hat{b}_H^\dagger + \hat{c}_H^\dagger \hat{d}_H^\dagger + \hat{a}_V^\dagger \hat{d}_V^\dagger + \hat{c}_V^\dagger \hat{b}_V^\dagger, \tag{2.4-164}$$

the only terms containing one operator in each spatial mode are the products of the first (last) two terms:

$$\hat{a}_H^\dagger \hat{b}_H^\dagger \hat{c}_H^\dagger \hat{d}_H^\dagger + \hat{a}_V^\dagger \hat{d}_V^\dagger \hat{c}_V^\dagger \hat{b}_V^\dagger. \tag{2.4-165}$$

The conventional way to generate a four-qubit GHZ state using linear optics is to take two independent Bell-pair sources, and interfere one photon from each source on a polarizing beamsplitter. A consequence of (2.2-79) is that the unitary transformation describing the PBS can be applied on the Hamiltonian itself. The Hamiltonian for the two Bell-pair sources is

$$H = \hat{a}_H^\dagger \hat{b}_H^\dagger + \hat{a}_V^\dagger \hat{b}_V^\dagger + \hat{c}_H^\dagger \hat{d}_H^\dagger + \hat{c}_V^\dagger \hat{d}_V^\dagger. \tag{2.4-166}$$

[6]The conjugate terms describing SFG have been omitted.

A PBS acting on the modes b and d implements the transformation:

$$\begin{aligned} \hat{b}_H^\dagger &\mapsto \hat{b}_H^\dagger, \qquad \hat{b}_V^\dagger \mapsto \hat{d}_V^\dagger \\ \hat{d}_H^\dagger &\mapsto \hat{d}_H^\dagger, \qquad \hat{d}_V^\dagger \mapsto \hat{b}_V^\dagger, \end{aligned} \tag{2.4-167}$$

which means that the transformed Hamiltonian is

$$H \mapsto \hat{a}_H^\dagger \hat{b}_H^\dagger + \hat{a}_V^\dagger \hat{d}_V^\dagger + \hat{c}_H^\dagger \hat{d}_H^\dagger + \hat{c}_V^\dagger \hat{b}_V^\dagger = H_{GHZ}. \tag{2.4-168}$$

The source using 'entanglement by path identity' is therefore formally equivalent to two standard SPDC sources combined with a simple linear-optical transformation. It can, however, be argued that the entanglement-by-path-identity source is conceptually different, and the theoretical simplicity of these types of source designs has allowed the study of sources capable of generating high-dimensional entangled states [283, 284, 285], that would otherwise be difficult to engineer. While these SDPC sources see photons from one crystal propagating through a second one, there is no actual interference between the down-conversion processes. Such interference is possible though, and the coherent cancellation of the emission between two different has been observed [286]. Unfortunately, interference between photon amplitudes from different SPDC sources cannot be used to improve the efficiency of entanglement generation. This is because while such nonlinear interferometers can give cancellations in the Hamiltonian, the resulting Hamiltonian is still that of a $\chi^{(2)}$ process.

2.4.5 HOM interference with SPDC photons

The discussion of Hong-Ou-Mandel interference in Section 2.2.9 used a single-mode description of the light, specifically a spectrally single-mode description. In practice single photons are never monochromatic, and a more realistic model of the phenomenon can be obtained by taking into account the multi-mode structure of the light. This leads to photons that are localized in space, and the amount of photon bunching will depend on the relative time-delay between the two wavepackets that undergo interference. A spectrally pure single-photon wavepacket can be written as

$$|1\rangle_{a,\phi} = \int_0^\infty d\omega \phi(\omega) \hat{a}^\dagger(\omega) |0\rangle_a . \tag{2.4-169}$$

In the frequency domain a time-delay τ of such a wavepacket is described by a frequency-dependent phase shift:

$$\hat{a}^\dagger(\omega) \mapsto e^{-i\omega\tau} \hat{a}^\dagger(\omega). \tag{2.4-170}$$

A relative time-delay between the photons in a spectrally non-separable two-photon wavepacket generated by SPDC can be described the same way:

$$\begin{aligned} &\iint_0^\infty d\omega_1 d\omega_2 f(\omega_1, \omega_2) \hat{a}^\dagger(\omega_1) \hat{b}^\dagger(\omega_2) |0\rangle_{a,b} \\ \mapsto &\iint_0^\infty d\omega_1 d\omega_2 f(\omega_1, \omega_2) \hat{a}^\dagger(\omega_1) \hat{b}^\dagger(\omega_2) e^{-i\omega_2\tau} |0\rangle_{a,b} , \end{aligned} \tag{2.4-171}$$

where $f(\omega_1, \omega_2) = \alpha(\omega_1 + \omega_2)\phi(\omega_1, \omega_2)$ is the joint spectral amplitude, which, since the state is normalized, is also normalized to one:

$$\iint_0^\infty d\omega_1 d\omega_2 |f(\omega_1, \omega_2)|^2 = 1. \tag{2.4-172}$$

In a Hong-Ou-Mandel scenario the state in (2.4-171) is the input state to a balanced beamsplitter acting on the modes a and b. Assuming that this beamsplitter is frequency independent, the state after the beamsplitter can be written as

$$\begin{aligned} |\Psi\rangle &= \iint_0^\infty d\omega_1 d\omega_2 f(\omega_1, \omega_2) \big(U_{\mathrm{BS}} \hat{a}^\dagger(\omega_1)\big) \big(U_{\mathrm{BS}} \hat{b}^\dagger(\omega_2)\big) e^{-i\omega_2 \tau} |0\rangle_{a,b} \\ &= \frac{1}{2} \iint_0^\infty d\omega_1 d\omega_2 f(\omega_1, \omega_2) e^{-i\omega_2 \tau} \\ &\times \left[\hat{a}^\dagger(\omega_1)\hat{a}^\dagger(\omega_2) - \hat{a}^\dagger(\omega_1)\hat{b}^\dagger(\omega_2) + \hat{b}^\dagger(\omega_1)\hat{a}^\dagger(\omega_2) - \hat{b}^\dagger(\omega_1)\hat{b}^\dagger(\omega_2)\right] |0\rangle_{a,b}. \end{aligned} \tag{2.4-173}$$

From this state the coincidence probability, meaning the probability to detect one photon in each mode, needs to be calculated. Following [287], this can be done by taking the expectation value of a frequency-independent projector onto the state $|1,1\rangle_{a,b}$. This projector is simply the tensor product of the projectors

$$P_a = \int_0^\infty d\omega_a \hat{a}^\dagger(\omega_a) |0\rangle\langle 0|_a \hat{a}(\omega_a), \qquad P_b = \int_0^\infty d\omega_b \hat{b}^\dagger(\omega_b) |0\rangle\langle 0|_b \hat{b}(\omega_b). \tag{2.4-174}$$

The coincidence probability is then

$$\begin{aligned} P_{\mathrm{cc}} &= \langle \Psi | P_a \otimes P_b | \Psi \rangle \\ &= \langle \Psi | \iint_0^\infty d\omega_a d\omega_b \hat{a}^\dagger(\omega_a)\hat{b}^\dagger(\omega_b) |0\rangle\langle 0|_{a,b} \hat{a}(\omega_a)\hat{b}(\omega_b) | \Psi \rangle \\ &= \frac{1}{4} \iiiint_0^\infty d\omega_1' d\omega_2' d\omega_a d\omega_b d\omega_1 d\omega_2 f^*(\omega_1', \omega_2') e^{i\omega_2' \tau} f(\omega_1, \omega_2) e^{-i\omega_2 \tau} \\ &\times \langle 0|_{a,b} \left(\hat{a}(\omega_1')\hat{a}(\omega_2') - \hat{a}(\omega_1')\hat{b}(\omega_2') + \hat{b}(\omega_1')\hat{a}(\omega_2') - \hat{b}(\omega_1')\hat{b}(\omega_2')\right) \\ &\times \hat{a}^\dagger(\omega_a)\hat{b}^\dagger(\omega_b) |0\rangle\langle 0|_{a,b} \hat{a}(\omega_a)\hat{b}(\omega_b) \\ &\times \left[\hat{a}^\dagger(\omega_1)\hat{a}^\dagger(\omega_2) - \hat{a}^\dagger(\omega_1)\hat{b}^\dagger(\omega_2) + \hat{b}^\dagger(\omega_1)\hat{a}^\dagger(\omega_2) - \hat{b}^\dagger(\omega_1)\hat{b}^\dagger(\omega_2)\right] |0\rangle_{a,b}. \end{aligned} \tag{2.4-175}$$

This expression can be simplified by noting that terms on the form

$$\langle 0 | \hat{a}(\omega_i)\hat{a}(\omega_j)\hat{a}^\dagger(\omega_k)\hat{b}^\dagger(\omega_l) | 0 \rangle, \tag{2.4-176}$$

with three operators acting on one mode, vanish, since these expressions turn into inner products between orthogonal Fock states. All the remaining terms contain one creation and one annihilation operator per mode, and these evaluate to products of delta functions:

$$\langle 0 | \hat{a}(\omega_i)\hat{b}(\omega_j)\hat{a}^\dagger(\omega_k)\hat{b}^\dagger(\omega_l) | 0 \rangle = \delta(\omega_i - \omega_k)\delta(\omega_j - \omega_l). \tag{2.4-177}$$

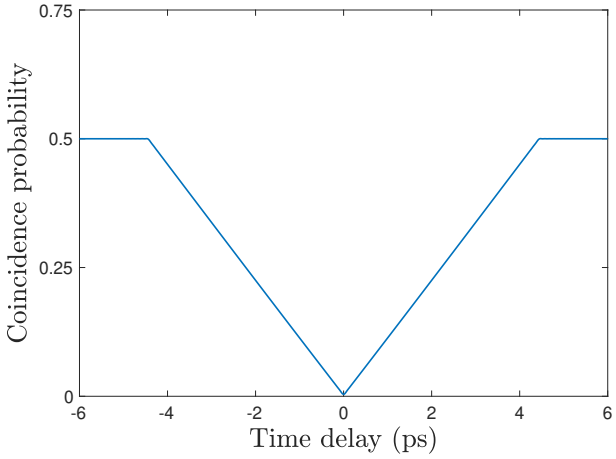


Figure 2.30: **Hong-Ou-Mandel dip.** The probability for two photons to anti-bunch on a beamsplitter as a function of their relative delay. This curve is known as a Hong-Ou-Mandel dip. The triangular shape of the curve is caused by the sinc function in the phase-matching amplitude. The dip was calculated for the joint-spectral amplitude pictured in Fig. 2.24b and has a visibility slightly over 0.99.

Substituting this in, the coincidence probability can be written

$$\begin{aligned} P_{\mathrm{cc}} = \frac{1}{4}\iiiiiint_0^\infty & d\omega_1' d\omega_2' d\omega_a d\omega_b d\omega_1 d\omega_2 f^*(\omega_1', \omega_2') f(\omega_1, \omega_2) e^{i(\omega_2' - \omega_2)\tau} \\ & \times \big[\delta(\omega_2' - \omega_a)\delta(\omega_1' - \omega_b) - \delta(\omega_1' - \omega_a)\delta(\omega_2' - \omega_b)\big] \\ & \times \big[\delta(\omega_2 - \omega_a)\delta(\omega_1 - \omega_b) - \delta(\omega_1 - \omega_a)\delta(\omega_2 - \omega_b)\big]. \end{aligned} \tag{2.4-178}$$

Using the relation

$$\int \delta(x_1 - x)\delta(x_2 - x)dx = \delta(x_1 - x_2) \tag{2.4-179}$$

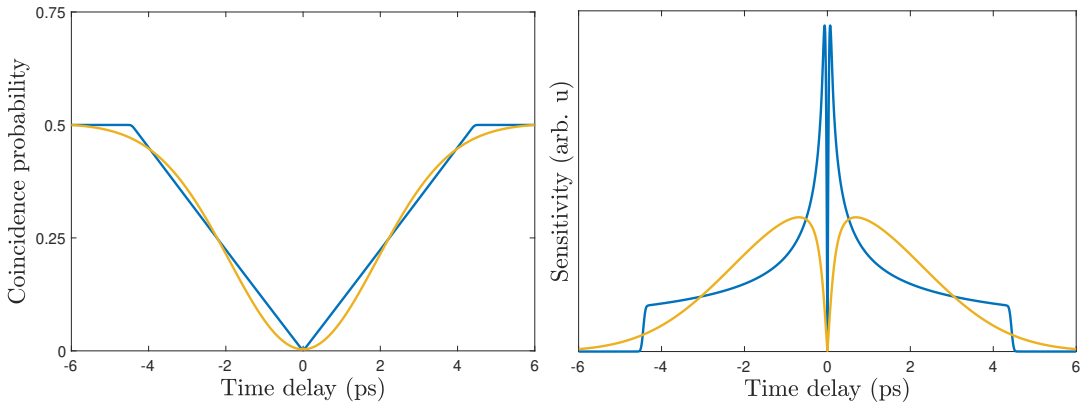
to integrate over ω_a and ω_b the coincidence probability reduces to

$$\begin{aligned} P_{\mathrm{cc}} = \frac{1}{2}\iiiint_0^\infty & d\omega_1' d\omega_2' d\omega_1 d\omega_2 f^*(\omega_1', \omega_2') f(\omega_1, \omega_2) e^{i(\omega_2' - \omega_2)\tau} \\ & \times \big[\delta(\omega_1' - \omega_1)\delta(\omega_2' - \omega_2) - \delta(\omega_2' - \omega_1)\delta(\omega_1' - \omega_2)\big]. \end{aligned} \tag{2.4-180}$$

After integrating over the primed frequencies the coincidence probability can finally be written as

$$\begin{aligned} P_{\mathrm{cc}} &= \frac{1}{2}\iint_0^\infty d\omega_1 d\omega_2 \big(|f(\omega_1, \omega_2)|^2 - f(\omega_1, \omega_2) f^*(\omega_2, \omega_1) e^{i(\omega_1 - \omega_2)\tau}\big) \\ &= \frac{1}{2} - \frac{1}{2}\iint_0^\infty d\omega_1 d\omega_2 f(\omega_1, \omega_2) f^*(\omega_2, \omega_1) e^{i(\omega_1 - \omega_2)\tau}, \end{aligned} \tag{2.4-181}$$

where the last step used the normalization condition from (2.4-172). For joint spectral amplitudes that cannot be factored as $f(\omega_s, \omega_i) = f_s(\omega_s) f_i(\omega_i)$, meaning for frequency entangled states, the integrals in (2.4-181) need to be evaluated numerically. Doing this for the JSA pictured in Fig. 2.24 gives the HOM-dip shown in Fig. 2.30.



(a) HOM dips (b) HOM interference sensitivity

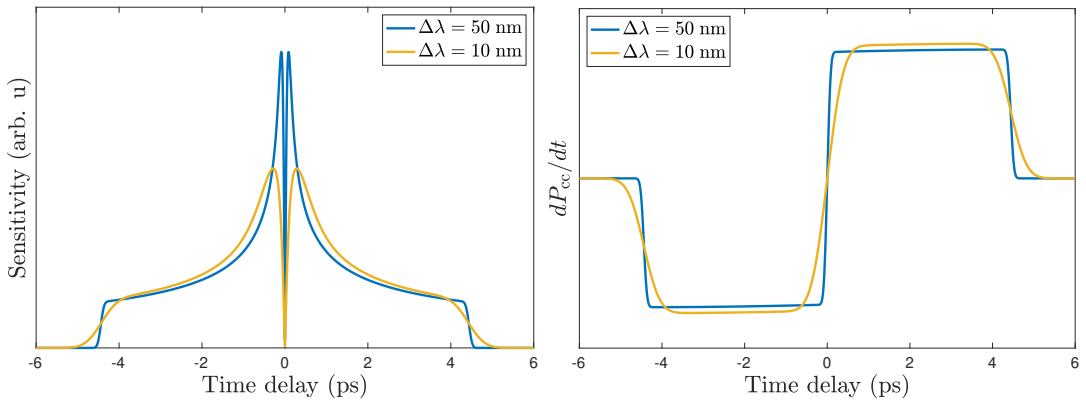
Figure 2.31: **HOM dip sensitivity.** Hong-Ou-Mandel interference can be used to measure short time delays. The sensitivity of such measurements depends on the particular shape of the HOM dip. (a) shows two different HOM dips arising from different photon spectra. The blue curve shows a triangular HOM dip while the yellow curve shows a Gaussian dip with the same FWHM. (b) illustrates the measurement sensitivity the two different dips achieve along different points of the respective curves. Due to the almost fully linear slope, the triangular HOM dip can achieve a higher sensitivity close to the minimum of the dip. Note that the blue curves have been calculated using a joint-spectral amplitude that is filtered by a 50 nm Gaussian filter to remove the Gibbs oscillations that would otherwise be present.

As can be seen, the indistinguishability at zero time delay is high even though the two-photon spectrum is entangled. The triangular shape of the HOM dip is caused by the sinc term in the phase-matching function. Loosely speaking, the HOM-dip is similar to a convolution of the single-photon wavepackets, and since the Fourier transform of a sinc spectral function is a rectangle function in time, the convolution between the wavepackets produces a triangle function. This shape of the HOM dip is interesting from a quantum metrology perspective. Hong-Ou-Mandel interferometry has recently attracted attention for its ability to measure small time delays [288], and has also been proposed as a way to probe the interface between quantum mechanics and general relativity [289]. One advantage of HOM interference over regular phase interference is the increased dynamic range, without the need for counting interference fringes. More significantly, HOM interference can exhibit group-velocity dispersion cancellation [226, 290], an effect that is particularly pernicious in optical coherence tomography [291]. The aforementioned dispersion cancellation properties of HOM interference can be used to overcome these limitations in what is known as quantum optical coherence tomography [292, 293, 294], and has been used to image 3D biological samples [295].

In HOM metrology, such as in Ref. [288], the point of highest sensitivity is determined by a trade-off between the slope of the HOM dip and the background photon rate. A simple estimate of the sensitivity is given by the ratio:

$$\left| \frac{d\langle N_{\mathrm{cc}} \rangle}{d\tau} \Big/ \sqrt{\langle N_{\mathrm{cc}} \rangle} \right|, \tag{2.4-182}$$

where τ is the time delay between the photons, $\langle N_{\mathrm{cc}} \rangle$ is the expected number of detected



(a) HOM sensitivity (b) HOM slope

Figure 2.32: **Spectral filtering in HOM interference.** The triangularly shaped HOM dips in Fig. 2.30 and Fig. 2.31a approach a triangular shape in the limit of infinite spectral bandwidth. In practice, however, various effects limit the achievable bandwidth and these need to be taken into account. The two figures above show the effect of spectral filtering on the sensitivity of the triangular HOM dip as well as its slope for two different Gaussian filter bandwidths. In (a) it can be seen that a filter bandwidth of $\Delta\lambda = 10\,\text{nm}$ has a significant effect on the sensitivity, even though the single-photon spectral bandwidth, defined as the FWHM of the marginal distributions of the joint-spectral intensity, is only about $\Delta\lambda_s = \Delta\lambda_i = 1.2\,\text{nm}$. (b) shows the derivative of the two HOM dips with respect to the time delay between the two photons. The spectral filtering removes the higher Fourier components, smooths out the square wave and thereby reduces the sensitivity.

coincidence events, which is directly proportional to the coincidence probability P_{cc}, and $\sqrt{\langle N_{\text{cc}}\rangle}$ is the uncertainty arising from the photon counting noise. The intuition behind this is that an absolute change in signal, or photon rate, corresponds to a bigger relative change in signal when the initial signal is small.

However, due to the fact that the coincidence probability cannot cross zero the slope of the signal necessarily decreases as the probability approaches zero. The triangularly shaped HOM dip is therefore interesting because it exhibits a linear slope almost down to zero coincidence probability. A comparison between the sensitivity for a triangularly shaped HOM dip and a typical Gaussian-shaped one with the same visibility and full width at half maximum (FWHM) is shown in Fig. 2.31.

In theory the triangularly shaped HOM dip can achieve much higher sensitivity close to the bottom of the dip, due to the fact that the slope does not vanish. In practice however, the sharpness of the discontinuities in the triangular HOM dip is set by the spectral bandwidth, and the highest Fourier components. This is illustrated in Fig. 2.32, where the sensitivity and derivative of the dip is shown for two different spectral bandwidths. The sensitivities therefore have to take the bandwidth into account, and spectrally Gaussian photons with an equivalent bandwidth could likely achieve a higher sensitivity, albeit with smaller dynamic range. A second problem with the triangular HOM dip is the fact that the optimal measurement point being so close to the minimum of the dip necessitates a very high HOM dip visibility. In practice one of the main contributors to a degraded HOM interference visibility is multiphoton events

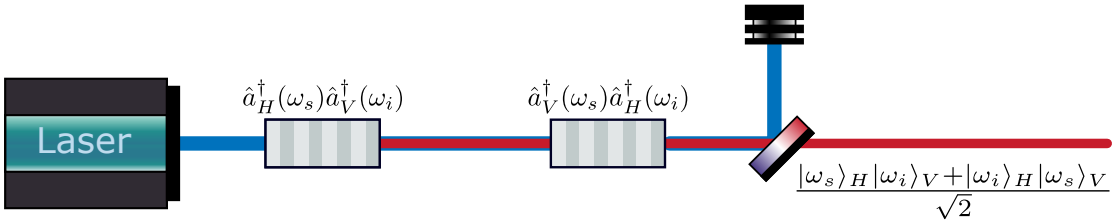


Figure 2.33: **Frequency entangling SPDC source.** A source consisting of two identical nonlinear crystals phase matched for non-degenerate type-II SPDC can be used to generate frequency entangled photon pairs by orienting the two crystals at 90° from each other and pumping the two crystals with a polarization that lies between the two crystals' optical axes. The polarization-frequency entanglement can be mapped to path-frequency entanglement by means of a polarizing beamsplitter (not pictured), and the phase in the entangled state can be controlled by the relative phase of the two polarization components of the pump light, with respect to the basis defined by the two crystals.

from the SPDC process. This therefore means that the pump power of the source needs to be kept low, which in turn reduces the detected photon rate. Precision metrology using single photons suffers from an intrinsically poor square-root noise scaling due to the Poissonian counting statistics of single-photon detection, and the total number of detected photons is one of the most important factors determining the achievable sensitivity.

While the triangularly shaped HOM dip discussed above results from a non-separable two-photon spectrum, this is in fact not a necessary condition, and a similarly shaped HOM dip can be produced by two spectrally pure photons as long as their individual spectra are sinc shaped. However, there are features in Hong-Ou-Mandel interference that only emerge as a consequence of two-photon frequency entanglement. In particular, a HOM dip between two photons in a frequency-entangled state will be modulated by a beat note whose oscillation frequency depends on the frequency difference between the two modes in the entangled state [296]. Such frequency entanglement can, unlike entanglement in many other degrees of freedom, be produced in a fairly straightforward manner.

Consider the collinear SPDC source pictured in Fig. 2.33 consisting of two crystals producing the frequency non-degenerate states:

$$|\Phi_1\rangle = |H\rangle_{\omega_s}|V\rangle_{\omega_i} \tag{2.4-183}$$

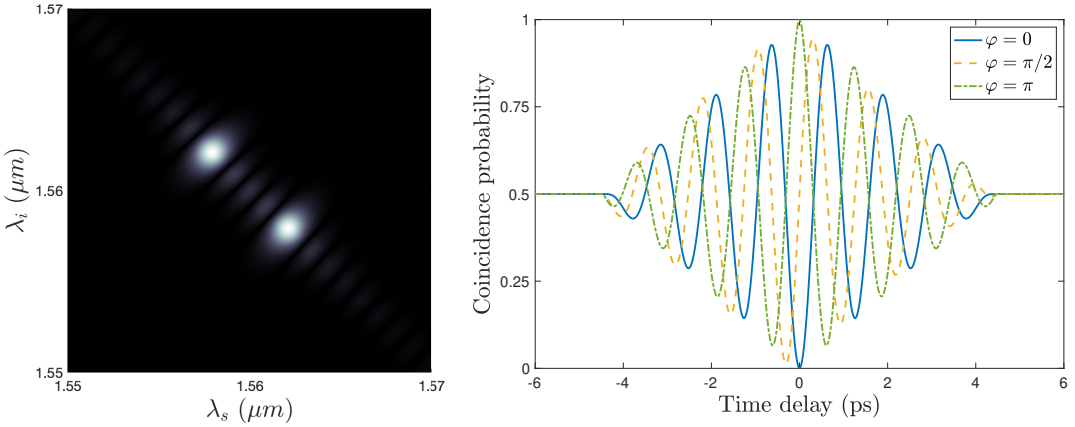
$$|\Phi_2\rangle = |V\rangle_{\omega_s}|H\rangle_{\omega_i}, \tag{2.4-184}$$

where $\omega_s \neq \omega_i$ are the centre frequencies of the signal and idler. Coherently pumping both crystals produces a superposition of both states

$$|\Psi\rangle = \frac{|H\rangle_{\omega_s}|V\rangle_{\omega_i} + e^{i\varphi}|V\rangle_{\omega_s}|H\rangle_{\omega_i}}{\sqrt{2}}, \tag{2.4-185}$$

where φ is the relative phase between the two polarization components of the pump field. The state can equivalently be written

$$|\Psi\rangle = \frac{|\omega_s\rangle_H|\omega_i\rangle_V + e^{i\varphi}|\omega_i\rangle_H|\omega_s\rangle_V}{\sqrt{2}}, \tag{2.4-186}$$



(a) $|f(\omega_1, \omega_2) + f(\omega_2, \omega_1)|$ (b) HOM beat notes

Figure 2.34: **Frequency entanglement in HOM interference.** (a) Two non-degenerate joint-spectral amplitudes can be added together by coherently pumping two different nonlinear crystals. This leads to a highly entangled two-photon spectrum. The non-degeneracy can be tuned either through the crystal temperature, or by detuning the pump light. In the figure above the ppKTP crystal is at 86.5 °C, while the other parameters are the same as in Fig. 2.24b. (b) The frequency entangled two-photon spectrum gives rise to a beat note in the HOM dip. The relative phase φ between the two down-conversion processes determines the phase of this beat note.

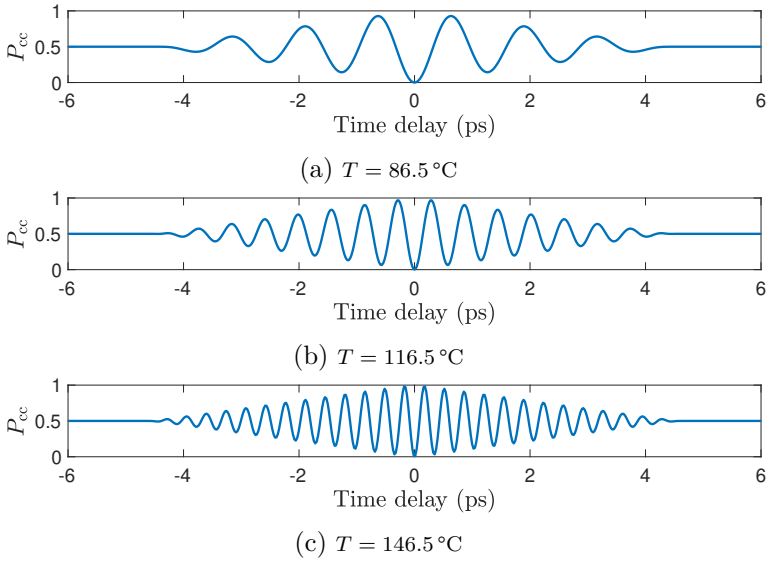
to highlight the frequency entanglement. For multi-mode states the result is similar. Assuming the two crystals are oriented at 90° to one another, but are otherwise identical, the state after the second crystal is

$$\begin{aligned} |\Psi\rangle = &\frac{1}{\sqrt{2}} \iint_0^\infty d\omega_s d\omega_i f(\omega_s, \omega_i) \hat{a}_H^\dagger(\omega_s) \hat{a}_V^\dagger(\omega_i) |0\rangle \\ &+ \frac{e^{i\varphi}}{\sqrt{2}} \iint_0^\infty d\omega_s d\omega_i f(\omega_s, \omega_i) \hat{a}_H^\dagger(\omega_i) \hat{a}_V^\dagger(\omega_s) |0\rangle . \end{aligned} \tag{2.4-187}$$

By exchanging the indices in the second integral the state can be written as

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \iint_0^\infty d\omega_s d\omega_i \big(f(\omega_s, \omega_i) + e^{i\varphi} f(\omega_i, \omega_s)\big) \hat{a}_H^\dagger(\omega_s) \hat{a}_V^\dagger(\omega_i) |0\rangle , \tag{2.4-188}$$

and $f_s(\omega_s, \omega_i) = f(\omega_s, \omega_i) + e^{i\varphi} f(\omega_i, \omega_s)$ becomes an effective joint-spectral amplitude that can be plugged into (2.4-181). An example of such a JSA and the corresponding HOM dip is shown in Fig. 2.34. This type of source has been used to perform temperature sensing [297], and has the potential to significantly outperform unmodulated HOM interference such as the one in [288]. In particular, since the amount of non-degeneracy in the down-conversion process can be continuously tuned, either by varying the crystal temperature, or offsetting the pump wavelength, the frequency of the bi-photon beat note can be tuned to produce an optimal trade-off between the dynamic range and the measurement precision. This tuning is illustrated in Fig. 2.35.



(a) $T = 86.5\,°\text{C}$

(b) $T = 116.5\,°\text{C}$

(c) $T = 146.5\,°\text{C}$

Figure 2.35: **Temperature tuning of HOM beat note.** The phase-matching amplitude in SPDC depends on the temperature of the nonlinear crystal, due to both thermal expansion of the crystal as well as the thermo-optic effect. This can be used to control the frequency detuning between the signal and idler, and therefore in turn the frequency of the Hong-Ou-Mandel beat note. (a) shows a HOM dip with the same parameters as in Fig. 2.34b, which corresponds to a wavelength detuning $\Delta\lambda = 6.5\,\text{nm}$. (b) shows the HOM dip for a crystal at 116.5 °C with a detuning of $\Delta\lambda = 14.2\,\text{nm}$. Finally, the HOM dip in (c) is plotted for a crystal with a temperature of $T = 146.5\,°\text{C}$ and a corresponding detuning of $\Delta\lambda = 23.7\,\text{nm}$.

2.5 Integrated photonics

The previous sections described the generation and manipulation of quantum light using bulk optical components, such as mirrors, lenses and birefringent crystals. Optical devices constructed out of these components are generally resource inefficient, in that they typically have a large footprint, are time consuming to construct and operate, and have a high price tag. In applications that aim to scale up the size of the quantum systems being manipulated, it is therefore necessary to miniaturize the optical devices used to generate and control the light. Such miniaturized devices that incorporate multiple optical elements are known as integrated photonic devices, and have applications outside the realm of quantum information processing as well, particularly in optical communication systems [298, 299] and display technologies [300, 301].

Integrated photonic devices rely on waveguides, which are material structures with boundary conditions engineered such that the solutions to Maxwell's equations are propagating modes confined within the structure. By far the most common example of this is silica optical fibers [302], which form the backbone of the global optical communications network. Such fibers exhibit extremely low loss per unit length, down to as little as $0.16\,\text{dB}\,\text{km}^{-1}$. Another widely used material is $LiNbO_3$ (Lithium Niobate) [303]. This material is transparent in the telecom C-band (1535 - 1565 nm) and can exhibit waveguide loss as low as $0.5\,\text{dB}\,\text{cm}^{-1}$ [304, 305], while offering a high $\chi^{(2)}$ response. The latter property has led to its widespread use in electro-optic modulators for classical

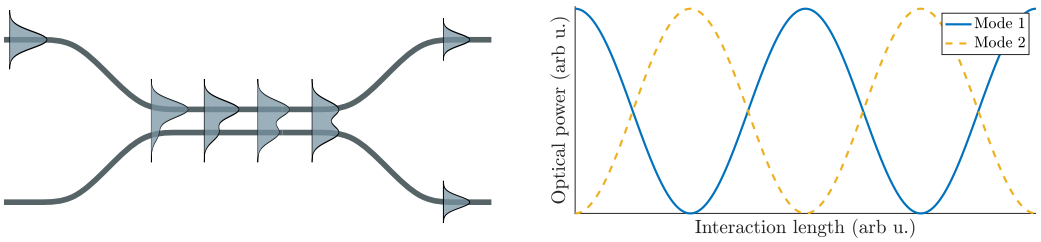
information transfer [306], with integrated devices achieving modulation frequencies of up to 100 GHz at low voltages [307, 308]. The nonlinear response of the material also makes it useful in optical frequency conversion [309], which is further facilitated by the ability to employ QPM structures [310]. $LiNbO_3$ can also be used to construct integrated lasers [311], ring-laser gyroscopes [312] and various other devices. One of the main drawbacks of $LiNbO_3$ is the comparatively low refractive-index contrast, resulting in a weak confinement of the light, thereby necessitating larger structures with longer bend radii [313]. In recent years there have been efforts to overcome this problem by embedding the Lithium Niobate waveguides in other structures [314, 315].

In addition to Lithium Niobate, other widely studied material platforms include GaAs (Gallium Arsenide) [316, 317, 318], SN (Silicon Nitride) [319, 320], Silicon-on-Insulator (SOI) [321, 322, 323], as well as so-called femtosecond laser direct written waveguides in various glasses and polymers [324, 325, 326]. Out of these, the SOI platform is the one which by far the most research effort has been dedicated towards. The main reason for this is that SOI photonic devices are compatible with the complementary metal-oxide-semiconductor (CMOS) process [327], and can therefore leverage the huge investments and technological progress this process has seen. Like $LiNbO_3$, Silicon is transparent for telecom wavelengths and can exhibit propagation losses as low as $0.3\,\mathrm{dB\,cm^{-1}}$ [328]. Due to its very high index contrast, SOI waveguides can be made very small and typically have a cross-section of $220 \times 450\,\mathrm{nm^2}$ [329, 330], enabling bend radii as small as 2 µm [331]. One of the main drawbacks of the SOI platform is that silicon has a centro-symmetric crystal structure, which in turn means that the material only has a vanishingly small second-order nonlinear response [332]. Direct electro-optic modulation of light is therefore infeasible in silicon [333, 334], and other methods such as thermo-optic phase shifters [335, 336] or the plasma dispersion effect and free carrier absorption [337, 338, 339] have to be used instead. Additionally, the low 1.1 eV bandgap of Silicon can lead to two-photon absorption in the near-infrared wavelength range, which places limits on the optical power that can be used in nonlinear processes [340, 341].

2.5.1 Integrated quantum photonics

For all the same reasons that integrated photonics revolutionised classical optics and vastly expanded the potential of optical information processing, integrated quantum photonics holds the promise of moving quantum photonic information processing from tabletop experiments to large scale, commercially relevant applications. As a result, the field has attracted considerable research attention over the past 20 years [342, 343, 344, 345, 346, 347], and has enabled many experiments that would have been infeasible using bulk optics, such as studies of complex quantum walks [348, 349, 350] and the generation of very large entangled states [351].

There are, broadly speaking, four engineering goals for integrated quantum photonic circuits: (1) the generation of quantum light through nonlinear processes, (2) the manipulation of quantum states of light using linear optical elements, (3) high efficiency detection of quantum light, and (4) the realisation of photon-photon interactions mediated by a nonlinear element. Although this final goal is the most challenging one, it has seen experimental progress in the last few years [352, 353, 354, 355]. Progress towards the other three goals is comparatively far along, and the largest challenge is



(a) Mode coupling. (b) Power transfer.

Figure 2.36: **Directional coupler.** (a) A schematic view of a directional coupler, with light travelling from left to right. The transverse mode of the waveguide extends into the cladding, and if the separation between two different waveguides becomes small enough the corresponding modes begin to couple. This manifests as a power transfer between the two waveguide modes. (b) The power transfer has a periodic behaviour that depends on the mode overlap between the waveguides and the interaction length. By choosing these parameters properly, a beamsplitter with the desired splitting ratio can be realised.

to integrate all these technologies in a single device. Single photons are for example routinely made in waveguides using both spontaneous three-wave [356] and four-wave mixing [357], and in more recent years integration of solid-state emitters such as quantum dots [201, 358] and various color centres have been successfully demonstrated [359, 360]. Meanwhile, in the realm of continuous-variable quantum information the strong mode confinement of waveguides has enabled the demonstration of extremely broadband and highly squeezed light in single-pass configurations [361]. High efficiency single-photon detection has also approached a mature stage. In some material platforms, photo-detection can be performed using single-photon avalanche diodes [362], however both single-photon superconducting nanowire detectors [363] and transition-edge sensors [364] have been demonstrated integrated in waveguide structures. Finally, various techniques exist for performing linear optical operations in integrated photonics. While there have been demonstrations of control over the polarization degree of freedom of photons in waveguides [365, 366, 367] this approach is generally not pursued, since most integrated photonic platforms exhibit birefringence that is too strong and furthermore hard to control dynamically. Instead most integrated photonic quantum information processing schemes rely on a so-called dual-rail encoding, where the computational basis states of a qubit or qudit are encoded in two or more different waveguides. This approach will be discussed in the next section.

2.5.2 Manipulating path encoded states

The two basic primitives for controlling the path of photon through a waveguide network are phase shifters and directional couplers. As discussed in the previous section, phase shifters can be made using a multitude of physical principles [334], such as the thermo-optic effect [335], Pockels effect [368] or various carrier injection or depletion methods. The latter tends to be problematic in quantum application due to the loss that it introduces, however attempts have been made to mitigate these effects by combining carrier injection with thermo-optic phase shifters in the same device [369].

Directional couplers, being completely static devices, are easier to manufacture.

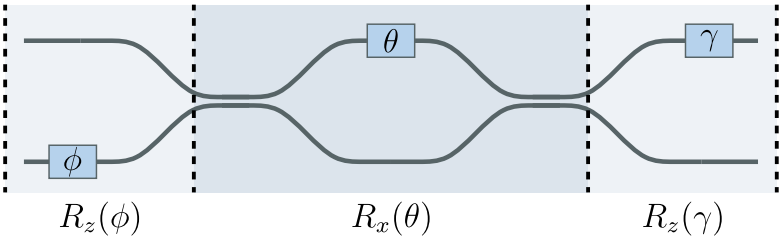


Figure 2.37: **Waveguide Mach-Zehnder interferometer.** In a waveguide a Mach-Zehnder interferometer (MZI) can be constructed using two directional couplers as beamsplitters. An internal tunable phase shifter enables the MZI to realise any R_x rotation. Combined with a phase shifter in one of the input modes and one of the output modes, both implementing R_z rotations, any $U \in \mathrm{SU}(2)$ can be realised on the path degree of freedom of a single photon. The two external phase shifters are drawn in opposite physical modes here, due to the fact that the identity operation of the MZI, $R_x(0)$, performs a mode swap.

Since the evanescent electric field of a waveguide mode extends outside the waveguide itself and into the cladding material, two different waveguide modes can be made to interact if the waveguides are sufficiently close to each other. The interaction between two such waveguides along the z-direction can be modelled by a pair of coupled differential equations [106, 370]:

$$\frac{dA_1}{dt} = -i\mathscr{C}_{21}e^{i\Delta\beta z}A_2(z) \tag{2.5-1}$$

$$\frac{dA_2}{dt} = -i\mathscr{C}_{12}e^{-i\Delta\beta z}A_1(z), \tag{2.5-2}$$

where A_k is the amplitude of the corresponding waveguide mode, $\Delta\beta = \beta_1 - \beta_2$ is difference between the propagation constants of the two modes, and $\mathscr{C}_{21}$ and $\mathscr{C}_{12}$ are the coupling constants between the two waveguides that are given by overlap integrals of the two modes. Assuming that the coupling between the modes is symmetric, i.e. $\mathscr{C}_{21} = \mathscr{C}_{12} = \mathscr{C}$, that the modes have the same propagation constant, and that there is initially no light in the second waveguide, the system of differential equations has the solution

$$P_1(z) = P_1(0)\cos^2\left(\mathscr{C}z\right) \tag{2.5-3}$$

$$P_2(z) = P_1(0)\sin^2\left(\mathscr{C}z\right), \tag{2.5-4}$$

where $P_k = |A_k|^2$ is the optical power in the corresponding waveguide mode. By choosing the length of the interaction region a beamsplitter with arbitrary reflectivity can be realised. This is illustrated in Fig. 2.36.

Two directional couplers with a balanced splitting ratio can be combined to create a Mach-Zehnder interferometer (MZI), and placing a phase shifter in one of the arms of the MZI lets the interferometer work as a reconfigurable beamsplitter. More concretely, the balanced directional couplers realise Hadamard operations, and the phase shifter

effects an R_z rotation. The action of the MZI can therefore be written

$$\begin{aligned} U_{\text{MZI}} &= U_{\text{Had}} R_z(\phi) U_{\text{Had}} \\ &= \frac{X+Z}{\sqrt{2}} \left(\cos\frac{\phi}{2}\mathbb{1} - i\sin\frac{\phi}{2} Z \right) \frac{X+Z}{\sqrt{2}} \\ &= \cos\frac{\phi}{2}\mathbb{1} - i\sin\frac{\phi}{2} X \\ &= R_x(\phi). \end{aligned} \tag{2.5-5}$$

Putting a phase shifter before and after the MZI, as shown in Fig. 2.37, therefore allows any $U \in \text{SU}(2)$ to be realised on the path degree of freedom:

$$R_z(\theta) R_x(\phi) R_z(\gamma), \tag{2.5-6}$$

as this is an Euler angle decomposition. As was shown in Ref. [371], any higher dimensional unitary acting on the path degree of freedom can be decomposed in terms of such 2×2 unitaries, and this decomposition is known as the Reck encoding. Surprisingly, only a single phase shifter external to each MZI is required. Similarly to the method for decomposing qubit unitaries using wave plates presented in Section 2.3.1, the Reck decomposition can be found by creating the inverse of an $N \times N$ unitary $U(N)$ step by step, by multiplying it from the right with unitaries of the form

$$T_{jk}(\theta, \phi) = R_z(\phi) R_x(\theta), \tag{2.5-7}$$

acting on the modes j and k, and whose physical interpretation is an MZI followed by a phase shifter. The decomposition proceeds by first using $N-1$ unitaries to null all the elements in the last row of U, except the element on the diagonal:

$$U(N) T_{N,N-1} T_{N,N-2} \cdots T_{N,1} = \begin{bmatrix} U(N-1) & 0 \\ 0 & e^{i\varphi} \end{bmatrix}. \tag{2.5-8}$$

Note that the order in which the elements in the row are nulled is arbitrary. In the next step the same procedure is repeated for $U(N-1)$, using $N-2$ multiplications from the right. The total number of 2-mode unitaries T_{jk} required in the decomposition of an $N \times N$ unitary can be found as

$$\sum_{k=2}^{N} k - 1 = \sum_{k=1}^{N-1} k = \frac{N(N-1)}{2}. \tag{2.5-9}$$

After the matrix $U(N)$ has been diagonalized the identity matrix can be obtained through multiplication with a diagonal matrix D, that corresponds to a single phase shifter in every mode:

$$\mathbb{1} = U(N) T_{N,N-1} T_{N,N-2} \cdots T_{2,1} D. \tag{2.5-10}$$

The desired unitary is found by multiplying this expression by the inverse of the decomposition:

$$U(N) = D^\dagger T_{2,1}^\dagger \cdots T_{N,N-2}^\dagger T_{N,N-1}^\dagger, \tag{2.5-11}$$

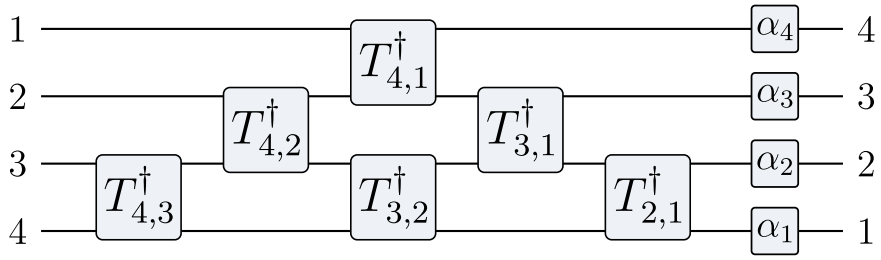


Figure 2.38: **Reck encoding.** Visual representation of the Reck decomposition into operations on spatial modes for a 4×4 unitary. The large boxes represent MZIs as in Fig. 2.37, but with only the first external phase shifter, and the α_k boxes represent single-mode phase shifts. Since the identity operation of an MZI exchanges the two spatial modes, the mode labels get permuted throughout the circuit and this naturally facilitates what would otherwise be operations between non-adjacent modes.

where

$$T_{jk}(\theta, \phi)^\dagger = R_x(-\theta) R_z(-\phi), \tag{2.5-12}$$

or in other words an MZI with a phase shifter acting on the input, and with both phases inverted. A schematic representation of the Reck decomposition for a 4×4 unitary is shown in Fig. 2.38.

Since each of the $N(N-1)/2$ two-mode unitaries has two free parameters the total number of free parameters in the decomposition is

$$N(N-1) + N = N^2, \tag{2.5-13}$$

where the second term on the left-hand side comes from the last N phase shifts. This means that the decomposition is optimal, since an arbitrary unitary $U(N)$ also has exactly N^2 free real parameters.

As an example of how the decomposition is performed, consider the three-dimensional quantum Fourier transform:

$$\mathscr{F}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \\ 1 & \frac{-1+i\sqrt{3}}{2} & \frac{-1-i\sqrt{3}}{2} \\ 1 & \frac{-1-i\sqrt{3}}{2} & \frac{-1+i\sqrt{3}}{2} \end{bmatrix}. \tag{2.5-14}$$

To null the matrix element $\left(\mathscr{F}_3\right)_{3,1}$ the rotation angle $\phi_{3,1}$ should be chosen such that:

$$\arg\left[\left(\mathscr{F}_3\right)_{3,3}\right] - \arg\left[\left(\mathscr{F}_3\right)_{3,1}\right] = \frac{\pi}{2}. \tag{2.5-15}$$

This comes from the fact that the diagonal and off-diagonal elements in

$$R_x(2\theta) = \begin{bmatrix} \cos\theta & -i\sin\theta \\ -i\sin\theta & \cos\theta \end{bmatrix}, \tag{2.5-16}$$

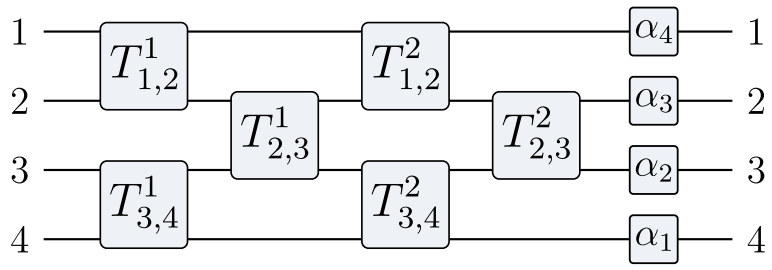


Figure 2.39: **Clements encoding.** The figure shows a representation of the Clements encoding scheme for a 4×4 unitary. While this scheme uses the same number of gates as the Reck scheme, it is more symmetric and space efficient, and the shorter overall propagation distance reduces the total loss. Note that in contrast to Fig. 2.38 there are no implicit mode swaps, since the decomposition only uses unitaries that couple adjacent modes.

always have a relative phase of $\pm\pi/2$. The condition (2.5-15) can be satisfied by choosing

$$\phi_{3,1} = \mathrm{mod}\Big(\arg\Big[\big(\mathscr{F}_3\big)_{3,3}\Big] - \arg\Big[\big(\mathscr{F}_3\big)_{3,1}\Big], 2\pi\Big) - \frac{\pi}{2} = \frac{\pi}{6}. \tag{2.5-17}$$

This choice of phase guarantees that the R_x rotation can null the matrix element $\big(\mathscr{F}_3\big)_{3,1}$. The rotation angle $\theta_{3,1}$ can be calculated as

$$\theta_{3,1} = -2\mathrm{atan}\left|\frac{\big(\mathscr{F}_3\big)_{3,1}}{\big(\mathscr{F}_3\big)_{3,3}}\right| = -\frac{\pi}{2}. \tag{2.5-18}$$

Direct calculation then shows that $T_{3,1}$ indeed nulls the correct element of $\mathscr{F}_3$

$$\mathscr{F}_3 T_{3,1}(\theta_{3,1}, \phi_{3,1}) = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & g & h \end{bmatrix}. \tag{2.5-19}$$

The same steps as the ones above can be repeated to calculate the remaining angles, giving

$$\begin{aligned} \phi_{3,2} &= \frac{3\pi}{4}, \qquad & \theta_{3,2} &= -2\mathrm{atan}\frac{1}{\sqrt{2}} \\ \phi_{2,1} &= \frac{5\pi}{8}, \qquad & \theta_{2,1} &= -\frac{\pi}{2}. \end{aligned} \tag{2.5-20}$$

Note that after the second step:

$$\mathscr{F}_3 T_{3,1}(\theta_{3,1}, \phi_{3,1}) T_{3,2}(\theta_{3,2}, \phi_{3,2}) = \begin{bmatrix} a' & b' & 0 \\ d' & e' & 0 \\ 0 & 0 & h' \end{bmatrix}, \tag{2.5-21}$$

the last column of $\mathscr{F}_3$ is nulled as well. This is a consequence of the fact that $\mathscr{F}_3$ is a unitary matrix, and is why it is only necessary to null the elements in each row.

Due to its asymmetric, triangular shape, shown in Fig. 2.38, the Reck encoding is not optimal in terms of footprint and propagation loss. In Ref. [372] a more symmetric, but

otherwise equivalent encoding, known as the Clements encoding, was proposed. This decomposition is illustrated in Fig. 2.39. To perform the decomposition one proceeds similarly to the Reck method, however, instead of only multiplying from the right to null a given row of the desired matrix, one instead nulls the sub-diagonals and alternates multiplying from the left and right after each sub-diagonal.

3
Engineering reciprocity in polarization optics

This chapter examines how polarization rotators, or gadgets, behave when the propagation direction of the light through them is reversed. After characterising the most commonly used polarization rotators, a set of missing devices is identified. Recipes for the realisation of these devices are then introduced and analysed, concluding with a discussion on their potential applications.

3.1 Introduction

The polarization of light is one of the most commonly used degrees of freedom for encoding quantum information in single photons, and control over the polarization is of great importance in many fields of optics, both classical and quantum [106,373,374,375]. There are many types of optical geometries, such as Michelson and Sagnac interferometers, in which light retraces its own path. The light will therefore propagate through any polarization optics placed in these optical geometries in two different propagation directions, and it is natural to ask how the polarization transformations in the two directions are related. Polarization devices tend to be divided into two distinct categories: devices which are invariant under the reversal of propagation direction are said to be reciprocal, and devices which do not exhibit such an invariance are said to be nonreciprocal [106]. This distinction, however, will turn out to be insufficiently precise.

The most well known example of a nonreciprocal polarization device is the Faraday rotator. Such a device rotates linear polarization using the magneto-optic effect. This effect breaks Lorenz reciprocity, and this is the origin of the optical nonreicprocity [376]. Faraday rotators can be used to construct a device known as an optical isolator [377], which only permits light to travel through it in one direction. This device is illustrated in Fig. 3.1, and it can be seen that the diagonally polarized state does not return to the original vertically polarized state under reversal of the propagation direction. Linear polarization retarders, such as wave plates and liquid crystal retarders, are commonly referred to as being reciprocal. However, there is a simple optical device, shown in Fig. 3.2, in which a quarter-wave plate behaves exactly the same as Faraday rotator [378]. This seeming inconsistency can be explained in many different ways; fundamentally, it boils down to under exactly which transformation one requires a polarization optic to be invariant, and in which reference frame this invariance is viewed. A Faraday rotator, for example, breaks time reversal invariance, while linear polarization retarders do not. However, in the context of experimental physics this is not an operationally useful

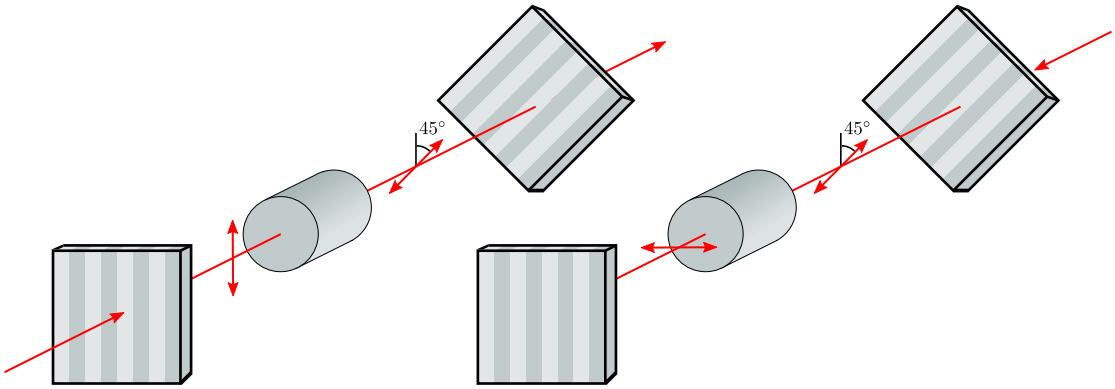


Figure 3.1: **Optical isolator.** An optical isolator consists of a Faraday rotator (grey cylinder), sandwiched by two polarizers: one aligned vertically, transmitting V-polarized light, and one aligned at 45° to the vertical axis, transmitting +-polarized light. In the forwards propagation direction the vertically polarized light is rotated clockwise by the Faraday rotator to + and is transmitted through the second polarizer. In the backwards propagation direction (right) light aligned with the second polarizer is still rotated clockwise, this time to H, and is then blocked by the vertically aligned polarizer. This is due to the nonreciprocity of the Faraday rotator.

distinction, since time reversal does not take place in the laboratory. In the next section we will therefore outline a more operationally useful point of view from which to discuss polarization operations in two propagation directions, and how they are related.

3.2 Preliminaries

Before moving onto the description of the various polarization devices that will be discussed, we will first introduce some notation and useful identities. The convention for the polarization states in this section will be the same as (2.3-3). Rotations about the principal axes of the Bloch sphere will be written as

$$X(\theta) = R_x(\theta) = \exp\left(-\frac{i}{2}\theta\sigma_x\right) \tag{3.2-1}$$

$$Y(\theta) = R_y(\theta) = \exp\left(-\frac{i}{2}\theta\sigma_y\right) \tag{3.2-2}$$

$$Z(\theta) = R_z(\theta) = \exp\left(-\frac{i}{2}\theta\sigma_z\right). \tag{3.2-3}$$

We recall that a rotation about some arbitrary axis $\vec{n}$ can be written

$$R_{\vec{n}}(\theta) = \cos\left(\frac{\theta}{2}\right) I - i \sin\left(\frac{\theta}{2}\right) \vec{n}\cdot\vec{\sigma}. \tag{3.2-4}$$

From this expression it is clear that

$$R_{\vec{n}}(2\pi) = -I. \tag{3.2-5}$$

It is worth pointing out the trivial fact that sequential rotations about the same axis are additive:

$$R_{\vec{n}}(\alpha)R_{\vec{n}}(\beta) = R_{\vec{n}}(\alpha+\beta), \tag{3.2-6}$$

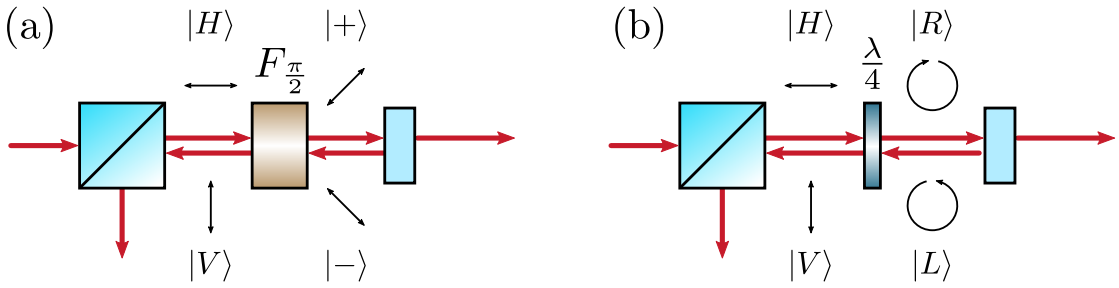


Figure 3.2: **Nonreciprocal elements.** The two figures show two different ways to build a device that functions similarly to the optical isolator in Fig. 3.1, blocking any backreflections in the transmitted port of the polarizing beamsplitter. In (a) the H-polarized light transmitted by the PBS is turned into '+'-polarized light by a Faraday rotator. Any light backreflected from an optic in the beam path becomes '−'-polarized in the backwards propagating frame, and is rotated to V by the Faraday rotator, after which it is rejected by the PBS. In (b) the Faraday rotator is replaced by a quarter-wave plate at 45°, which turns the H-polarized light into right-handed circularly polarized light. In the backwards moving frame this turns into L-polarized light, which the QWP transforms into V, since the angle of the wave plate is −45° in this frame. This is an example of a supposedly reciprocal element, a QWP, being used to realise a nonreciprocal device.

as we will make frequent use of this fact. We now prove a useful permutation identity for Bloch-sphere rotations. Let j, k, l be some permutation of the axes x, y, z. Then

$$\begin{aligned} R_j\left(\pm\frac{\pi}{2}\right) R_k(\theta) &= \frac{1}{\sqrt{2}}\left[I \mp i\sigma_j\right]\left[\cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}\sigma_k\right] \\ &= \frac{1}{\sqrt{2}}\left[\cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}\sigma_k \mp i\cos\frac{\theta}{2}\sigma_j \mp \sin\frac{\theta}{2}\sigma_j\sigma_k\right] \\ &= \frac{1}{\sqrt{2}}\left[\cos\frac{\theta}{2}I - \sin\frac{\theta}{2}\epsilon_{jkl}\sigma_l\sigma_j \mp \cos\frac{\theta}{2}\sigma_j \mp i\sin\frac{\theta}{2}\epsilon_{jkl}\sigma_l\right] \\ &= \left[\cos\frac{\theta}{2}I \mp i\sin\frac{\theta}{2}\epsilon_{jkl}\sigma_l\right]\frac{1}{\sqrt{2}}\left[I \mp i\sigma_j\right] \\ R_j\left(\pm\frac{\pi}{2}\right) R_k(\theta) &= R_l(\pm\epsilon_{jkl}\theta) R_j\left(\pm\frac{\pi}{2}\right), \end{aligned} \tag{3.2-7}$$

where ϵ_{jkl} is the Levi-Civita symbol. The inverse of this identity, for performing permutations in the reverse order, is

$$R_j(\theta) R_k\left(\pm\frac{\pi}{2}\right) = R_k\left(\pm\frac{\pi}{2}\right) R_l(\pm\epsilon_{jkl}\theta) \tag{3.2-8}$$

These identities also hold more generally for any three orthogonal axes j, k, l forming a right-handed coordinate system, since the transformed Pauli matrices representing those axes still obey the same commutation relationships. A corollary that follows directly from (3.2-7) is

$$\begin{aligned} R_j(\pm\pi) R_k(\theta) &= R_j\left(\pm\frac{\pi}{2}\right) R_l(\pm\epsilon_{jkl}\theta) R_j\left(\pm\frac{\pi}{2}\right) \\ &= R_l(\epsilon_{jlk}\epsilon_{jkl}\theta) R_j(\pm\pi) \qquad\qquad j \neq k \\ &= R_k(-\theta) R_j(\pm\pi). \end{aligned} \tag{3.2-9}$$

Using the relation above, one can also show a useful 'sandwich' identity:

$$\begin{aligned} R_j(\pm\theta) &= R_j(\pm\theta)R_k(\pm\pi)R_k(\mp\pi) \\ &= R_k(\pm\pi)R_j(\mp\theta)R_k(\mp\pi) \qquad j \neq k \\ &= R_k(\mp\pi)R_j(\mp\theta)R_k(\pm\pi). \end{aligned} \tag{3.2-10}$$

Additionally, an analogue of the cyclic relation $\sigma_j\sigma_k = \delta_{jk}I + i\epsilon_{jkl}\sigma_l$ for Pauli matrices also follows from (3.2-7) and (3.2-9):

$$\begin{aligned} R_j(\pm\pi)R_k(\pm\pi) &= R_j(\pm\pi/2)R_k(\pm\pi)R_j(\mp\pi/2) \\ &= R_l(\pm^2\epsilon_{jkl}\pi)R_j(\pm\pi/2)R_j(\mp\pi/2) \qquad j \neq k \\ &= R_l(\epsilon_{jkl}\pi), \end{aligned} \tag{3.2-11}$$

and similarly if R_j and R_k have opposite sign in their argument:

$$\begin{aligned} R_j(\pm\pi)R_k(\mp\pi) &= R_j(\pm\pi/2)R_k(\mp\pi)R_j(\pm\pi/2) \\ &= R_l(\pm\mp\epsilon_{jkl}\pi)R_j(\mp\pi/2)R_j(\pm\pi/2) \qquad j \neq k \\ &= R_l(-\epsilon_{jkl}\pi). \end{aligned} \tag{3.2-12}$$

Finally, we mention a result from [161], which is that a sequence of two half-wave plates and one quarter-wave plate can always be reduced to a sequence containing just one of each kind of wave plate:

$$Q(\alpha)H(\beta)H(\gamma) = Q(\alpha+\pi/2)H(\alpha-\beta+\gamma-\pi/2) \tag{3.2-13}$$

$$H(\alpha)H(\beta)Q(\gamma) = H(\alpha-\beta+\gamma-\pi/2)Q(\gamma+\pi/2). \tag{3.2-14}$$

Recursive application of these rules can reduce the number of HWPs in longer sequences to a single one.

3.2.1 Polarization gadgets

Since the operation realised by a polarization rotator is not necessarily the same in the two propagation directions it can be probed in, it is helpful to make a distinction between the physical device, which will be referred to as a gadget, and the mathematical operation effected by the gadget. A gadget $\mathbf{G}$ is defined as a sequence of polarization rotators $\mathbf{G} = \mathbf{G}_N \ldots \mathbf{G}_2\mathbf{G}_1$ probed in a certain direction. Gadgets will be written in boldface, to distinguish them from the SU(2) operations they realise and to emphasize that these operation are implementation dependent. In this chapter, the discussion will be restricted to half-wave plates $\mathbf{H}$, quarter-wave plates $\mathbf{Q}$ and Faraday rotators $\mathbf{F}$, as well as gadgets that are compositions of these elements.

To obtain the description of a gadget in the reverse propagation direction we introduce a counterpropagation operator Θ. This operator is order reversing and gives the action of all the sub-gadgets in the backwards propagation direction:

$$\begin{aligned} \Theta(\mathbf{G}) &= \Theta(\mathbf{G}_N \ldots \mathbf{G}_2\mathbf{G}_1) \\ &= \Theta(\mathbf{G}_1)\Theta(\mathbf{G}_2)\Theta(\mathbf{G}_N). \end{aligned} \tag{3.2-15}$$

A gadget that is a function of some parameter ϕ will be called symmetric in this parameter if

$$\Theta[\mathbf{G}^S(\phi)] = \mathbf{G}^S(\phi), \tag{3.2-16}$$

and anti-symmetric if

$$\Theta[\mathbf{G}^A(\phi)] = \mathbf{G}^A(-\phi). \tag{3.2-17}$$

3.2.2 Basic gadgets

Wave plates are examples of linear retarders, and in general a linear retarder with retardance η at an angle ϕ to the vertical axis corresponds to the gadget

$$\begin{aligned}\mathbf{W}_\eta(\phi) &= Y(2\phi)\mathbf{W}_\eta(0)Y(-2\phi)\\ &= Y(2\phi)Z(\eta)Y(-2\phi).\end{aligned} \tag{3.2-18}$$

Here the Y-rotations describe the rotation of the wave plate about the propagation direction of the light. On the Bloch sphere, this gadget corresponds to a rotation around an axis $\vec{n}$ by an angle η:

$$\begin{aligned}\mathbf{W}_\eta(\phi) &= (\cos\phi I - i\sin\phi Y)(\cos\frac{\eta}{2}I - i\sin\frac{\eta}{2}Z)(\cos\phi I + i\sin\phi Y)\\ &= \mathrm{c}_\phi\mathrm{c}_{\frac{\eta}{2}}\mathrm{c}_\phi I + i\mathrm{c}_\phi\mathrm{c}_{\frac{\eta}{2}}\mathrm{s}_\phi Y - i\mathrm{c}_\phi\mathrm{s}_{\frac{\eta}{2}}\mathrm{c}_\phi Z + \mathrm{c}_\phi\mathrm{s}_{\frac{\eta}{2}}\mathrm{s}_\phi ZY\\ &- i\mathrm{s}_\phi\mathrm{c}_{\frac{\eta}{2}}\mathrm{c}_\phi Y + \mathrm{s}_\phi\mathrm{c}_{\frac{\eta}{2}}\mathrm{s}_\phi YY - \mathrm{s}_\phi\mathrm{s}_{\frac{\eta}{2}}\mathrm{c}_\phi YZ - i\mathrm{s}_\phi\mathrm{s}_{\frac{\eta}{2}}\mathrm{s}_\phi YZY\\ &= (\mathrm{c}_\phi^2 + \mathrm{s}_\phi^2)\mathrm{c}_{\frac{\eta}{2}}I - 2i\mathrm{c}_\phi\mathrm{s}_\phi\mathrm{s}_{\frac{\eta}{2}}X - i(\mathrm{c}_\phi^2 - \mathrm{s}_\phi^2)\mathrm{s}_{\frac{\eta}{2}}Z\\ &= \cos\frac{\eta}{2}I - i\sin\frac{\eta}{2}\big(\sin 2\phi X + \cos 2\phi Z\big)\\ &= R_{\vec{n}}(\eta),\end{aligned} \tag{3.2-19}$$

where

$$\vec{n} = \begin{bmatrix}\sin 2\phi & 0 & \cos 2\phi\end{bmatrix}, \tag{3.2-20}$$

and $\mathrm{c}_{\{\cdot\}}$ and $\mathrm{s}_{\{\cdot\}}$ are short for $\cos(\cdot)$ and $\sin(\cdot)$, respectively. The factor of two in the angle ϕ comes from the difference between physical angles and Bloch-sphere angles. Using this notation for a wave-plate gadget, half-wave and quarter-wave plates can be written

$$\mathbf{H}(\phi) = \mathbf{W}_\pi(\phi), \tag{3.2-21}$$

$$\mathbf{Q}(\phi) = \mathbf{W}_{\frac{\pi}{2}}(\phi). \tag{3.2-22}$$

In the counterpropagating frame the angles with respect to the vertical axis undergo a sign change:

$$\Theta[\mathbf{W}_\eta(\phi)] = \mathbf{W}_\eta(-\phi) \tag{3.2-23}$$

$$\Theta[\mathbf{H}(\phi)] = \mathbf{H}(-\phi) \tag{3.2-24}$$

$$\Theta[\mathbf{Q}(\phi)] = \mathbf{Q}(-\phi). \tag{3.2-25}$$

This is simply a geometric consequence of any coordinate transformation to the counterpropagating frame that respects $\Theta[\mathbf{W}_\eta(0)] = \mathbf{W}_\eta(0)$. For Faraday rotators it is not

necessary to specify the angle to the vertical axis, since their action is independent of it. The same is true for any circularly birefringent gadget, since the circular polarization does not 'see' this angle, as there is nothing to reference it to. Faraday rotators will therefore only be denoted by their circular retardance ψ:

$$\mathbf{F}_\psi = R_y(\psi). \tag{3.2-26}$$

As mentioned in the introduction to this chapter, Faraday rotators violate Lorenz reciprocity, and quantitatively this manifests as a change in the sign of the retardance under counterpropagation:

$$\Theta[\mathbf{F}_\phi] = \mathbf{F}_{-\psi}. \tag{3.2-27}$$

Since, in practice Faraday, rotators are almost exclusively sold with a fixed circular retardance of $\psi = \pi/2$ these devices will be referred to using the shorthand notation $\mathbf{F}_\pm = \mathbf{F}_{\pm\frac{\pi}{2}}$.

3.2.3 Wave plate only gadgets

As discussed in Section 2.3.1 a combination of two quarter-wave plates and one half-wave plate can be used to realise any $U \in \mathrm{SU}(2)$, using what is known as a Simon–Mukunda gadget [161]. In this section, we will start by examining three simpler wave-plate gadgets implementing rotations about the main axes of the Bloch sphere.

Using a gadget consisting of two fixed quarter-wave plates and one tunable half-wave plate an $X(\phi)$ rotation can be realised. This gadget turns out to be anti-symmetric, and has the following form

$$\begin{aligned}
\mathbf{X}^A(\phi) &= \mathbf{Q}(\pi/2)\mathbf{H}(\phi/4)\mathbf{Q}(\pi/2) \\
&= \big[Y(\pi)Z(\pi/2)Y(-\pi)\big]\big[Y(\phi/2)Z(\pi)Y(-\phi/2)\big]\big[Y(\pi)Z(\pi/2)Y(-\pi)\big] \\
&= Z(-\pi/2)\big[Y(\phi/2)Z(\pi)Y(-\phi/2)\big]Z(-\pi/2) \\
&= X(\phi/2)X(\phi/2) \\
&= X(\phi),
\end{aligned} \tag{3.2-28}$$

where the second-to-last step used the permutation identity (3.2-7). Verifying that this gadget is in fact anti-symmetric is straightforward:

$$\begin{aligned}
\Theta\big[\mathbf{X}^A(\phi)\big] &= \Theta\big[\mathbf{Q}(\pi/2)\mathbf{H}(\phi/4)\mathbf{Q}(\pi/2)\big] \\
&= \mathbf{Q}(-\pi/2)\mathbf{H}(-\phi/4)\mathbf{Q}(-\pi/2) \\
&= \mathbf{Q}(\pi/2)\mathbf{H}(-\phi/4)\mathbf{Q}(\pi/2) \\
&= X(-\phi),
\end{aligned} \tag{3.2-29}$$

where the second step used the fact that

$$\begin{aligned}
\mathbf{Q}(-\pi/2) &= Y(-\pi)Z(\pi/2)Y(\pi) \\
&= Y(-2\pi)Y(\pi)Z(\pi/2)Y(-\pi)Y(2\pi) \\
&= Y(\pi)Z(\pi/2)Y(-\pi) \\
&= \mathbf{Q}(\pi/2),
\end{aligned} \tag{3.2-30}$$

or in other words the fact that the quarter-wave plate has a period of π in the rotation angle. Next, we consider a $Y(\phi)$ rotation. Such a rotation can be realised using a gadget consisting of two half-wave plates and is symmetric:

$$\begin{aligned}\mathbf{Y}^S(\phi) &= \mathbf{H}\Big(\frac{\phi+\alpha}{4}+\frac{\pi}{2}\Big)\mathbf{H}\Big(\frac{\alpha}{4}\Big)\\ &= Y\Big(\frac{\phi+\alpha}{2}+\pi\Big)Z(\pi)Y\Big(-\frac{\phi+\alpha}{2}-\pi\Big)Y\Big(\frac{\alpha}{2}\Big)Z(\pi)Y\Big(-\frac{\alpha}{2}\Big)\\ &= Y(\phi+\alpha+2\pi)Z(\pi)Y\Big(\frac{\alpha}{2}\Big)Z(\pi)Y\Big(-\frac{\alpha}{2}\Big)\\ &= -Y(\phi+\alpha)Z(2\pi)Y(-\alpha)\\ &= Y(\phi+\alpha)Y(-\alpha)\\ &= Y(\phi).\end{aligned} \tag{3.2-31}$$

It can be seen that the Y-rotation only depends on the relative angle between the two HWPs, which is expected since as mentioned before circularly polarized light, which correspond to the eigenstates of Y, does not have a well defined angle with respect to the wave plates. The gadget can therefore, without loss of generality, be written in the simpler form

$$\mathbf{Y}^S(\phi) = \mathbf{H}\Big(\frac{\phi}{4}+\frac{\pi}{2}\Big)\mathbf{H}(0). \tag{3.2-32}$$

The fact that the rotation angle only depends on the relative angle of the wave plates makes the symmetry of the gadget obvious, since

$$\begin{aligned}\Theta\big[\mathbf{Y}^S\big] &= \Theta\big[\mathbf{H}\Big(\frac{\phi}{4}+\frac{\pi}{2}\Big)\mathbf{H}(0)\big]\\ &= \mathbf{H}(0)\mathbf{H}\Big(-\frac{\phi}{4}-\frac{\pi}{2}\Big)\\ &= \mathbf{H}\Big(\frac{\phi}{4}+\frac{\pi}{2}\Big)\mathbf{H}(0),\end{aligned} \tag{3.2-33}$$

however the second step can also be explicitly verified:

$$\begin{aligned}\Theta\big[\mathbf{Y}^S\big] &= \mathbf{H}(0)\mathbf{H}\Big(-\frac{\phi}{4}-\frac{\pi}{2}\Big)\\ &= Z(\pi)\Big[-i\sin\Big(-\frac{\phi}{2}-\pi\Big)X - i\cos\Big(-\frac{\phi}{2}-\pi\Big)Z\Big]\\ &= -iZ\Big[-i\sin\frac{\phi}{2}X + i\cos\frac{\phi}{2}Z\Big]\\ &= \cos\frac{\phi}{2}I - i\sin\frac{\phi}{2}Y\\ &= Y(\phi).\end{aligned} \tag{3.2-34}$$

The last of the rotations about the main axes, a Z-rotation, can be realised by a symmetric gadget consisting of two fixed quarter-wave plates and one variable half-wave

plate:

$$\begin{aligned}\mathbf{Z}^S(\phi) &= \mathbf{Q}(\pi/4)\mathbf{H}(\phi/4-\pi/4)\mathbf{Q}(\pi/4)\\ &= \big[Y(\pi/2)Z(\pi/2)Y(-\pi/2)\big]\big[Y(\phi/2-\pi/2)Z(\pi)Y(-\phi/2+\pi/2)\big]\\ &\times \big[Y(\pi/2)Z(\pi/2)Y(-\pi/2)\big]\\ &= \big[Y(\pi/2)X(\pi/2)\big]\big[X(-\phi/2+\pi/2)Z(2\pi)X(-\phi/2+\pi/2)\big]\big[X(\pi/2)Y(-\pi/2)\big]\\ &= -Y(\pi/2)X(-\phi+2\pi)Y(-\pi/2)\\ &= Y(\pi/2)X(\phi)Y(-\pi/2)\\ &= Z(\phi)Y(\pi/2)Y(-\pi/2)\\ &= Z(\phi),\end{aligned} \tag{3.2-35}$$

where the second step made use of the permutation identity (3.2-7) twice. The symmetry of the gadget can be verified by noting that the Z-rotations of the wave plates in the calculation above ultimately reduce to a global π phase, and one can therefore skip directly to the third last line and simply substitute in negative angles resulting from the counterpropagation:

$$\begin{aligned}\Theta\big[\mathbf{Z}^S(\phi)\big] &= \mathbf{Q}(-\pi/4)\mathbf{H}(-\phi/4+\pi/4)\mathbf{Q}(-\pi/4)\\ &= Y(-\pi/2)X(-\phi)Y(\pi/2)\\ &= Z(\phi)Y(-\pi/2)Y(\pi/2)\\ &= Z(\phi).\end{aligned} \tag{3.2-36}$$

In addition to rotations about the three main axes described above, another useful gadget is the arbitrary linear retarder realising an arbitrary rotation about some axis lying in the linear plane:

$$U = \exp\Big[-\frac{i}{2}\alpha\big(\cos(\beta)Z+\sin(\beta)X\big)\Big] = Y(\beta)Z(\alpha)Y(-\beta). \tag{3.2-37}$$

Such a unitary can be realised by the gadget

$$\begin{aligned}\mathbf{L}(\alpha,\beta) &= \mathbf{Q}\Big(\frac{\beta}{2}-\frac{\pi}{4}\Big)\mathbf{H}\Big(\frac{\beta}{2}-\frac{\alpha}{4}+\frac{\pi}{4}\Big)\mathbf{Q}\Big(\frac{\beta}{2}-\frac{\pi}{4}\Big)\\ &= \mathbf{Q}(\phi/2)\mathbf{H}(\theta/2)\mathbf{Q}(\phi/2)\\ &= Y(\phi)Z(\pi/2)Y(-\phi)Y(\theta)Z(\pi)Y(-\theta)Y(\phi)Z(\pi/2)Y(-\phi)\\ &= Y(\phi)X(\phi)X(-\theta)Z(2\pi)X(-\theta)X(\phi)Y(-\phi)\\ &= -Y(\phi)X(2\phi-2\theta)Y(-\phi)\\ &= -Y(\beta-\pi/2)X(\alpha-2\pi)Y(-\beta+\pi/2)\\ &= Y(\beta)Z(\alpha)Y(-\beta),\end{aligned} \tag{3.2-38}$$

where $\phi = \beta-\pi/2$ and $\theta = \beta-\alpha/2+\pi/2$. This gadget does not have an overall

symmetry under $\Theta[\cdot]$, but is symmetric in α and anti-symmetric in β:

$$\begin{aligned}\Theta\big[\mathbf{L}(\alpha,\beta)\big] &= \mathbf{Q}\Big(-\frac{\beta}{2}+\frac{\pi}{4}\Big)\mathbf{H}\Big(-\frac{\beta}{2}+\frac{\alpha}{4}-\frac{\pi}{4}\Big)\mathbf{Q}\Big(-\frac{\beta}{2}+\frac{\pi}{4}\Big)\\ &= \mathbf{Q}(-\phi/2)\mathbf{H}(-\theta/2)\mathbf{Q}(-\phi/2)\\ &= -Y(-\phi)X(-2\phi+2\theta)Y(\phi)\\ &= Y(-\beta+\pi/2)X(-\alpha)Y(\beta-\pi/2)\\ &= Y(-\beta)Z(\alpha)Y(\beta)\\ &= \mathbf{L}(\alpha,-\beta).\end{aligned} \tag{3.2-39}$$

Finally, while the Simon–Mukunda gadget was proven to be universal through a constructive proof in Section 2.3.1, this property can also be shown using the types of decompositions above:

$$\begin{aligned}\mathbf{G}_{\mathrm{SM}} &= \mathbf{H}(\alpha/2)\mathbf{Q}(\beta/2)\mathbf{Q}(\gamma/2)\\ &= \big[Y(\alpha)Z(\pi)Y(-\alpha)\big]\big[Y(\beta)Z(\pi/2)Y(-\beta)\big]\big[Y(\gamma)Z(\pi/2)Y(-\gamma)\big]\\ &= Y(2\alpha-\beta)Z(2\pi)X(-\beta+\gamma)Y(-\gamma)\\ &= Y(2\alpha-\beta)X(-\beta+\gamma)Y(-\gamma).\end{aligned} \tag{3.2-40}$$

It's easy to verify that $\alpha' = 2\alpha-\beta$, $\beta' = -\beta+\gamma$ and $\gamma' = -\gamma$ are three independent Euler angles, and the gadget can therefore implement any $U \in \mathrm{SU}(2)$. This gadget does not have a symmetry in any of the parameters α, β or γ, but as will be explored next, its Euler-angle parametrisation does have one.

3.2.4 Linearly birefringent gadgets

In the previous section the symmetry properties of a few examples of wave-plate gadgets were calculated explicitly. However, for the special case of gadgets consisting only of a sequence of linear retarders, such as wave plates, there is a simpler way to find the action of the counterpropagation operator. Using (3.2-10) with $j = z$ the action of a single linear retarder in the backwards direction, given by (3.2-23), can be re-written as

$$\begin{aligned}\Theta[\mathbf{W}_\eta(\phi)] &= Y(-2\phi)Z(\eta)Y(2\phi)\\ &= Z(\pm\pi)Y(2\phi)Z(\mp\pi)Z(\pm\pi)Z(\eta)Z(\mp\pi)Z(\pm\pi)Y(-2\phi)Z(\mp\pi)\\ &= Z(\pm\pi)Y(2\phi)Z(\eta)Y(-2\phi)Z(\mp\pi)\\ &= Z(\pm\pi)\mathbf{W}_\eta(\phi)Z(\mp\pi).\end{aligned} \tag{3.2-41}$$

Note that any linear retarder is described by a symmetric matrix:

$$\begin{aligned}\mathbf{W}_\eta(\phi)^T &= Y(-2\phi)^T Z(\eta)^T Y(2\phi)^T\\ &= Y(2\phi)Z(\eta)Y(-2\phi)\\ &= \mathbf{W}_\eta(\phi),\end{aligned} \tag{3.2-42}$$

since

$$X(\phi)^T = X(\phi) \tag{3.2-43}$$
$$Y(\phi)^T = Y(-\phi) \tag{3.2-44}$$
$$Z(\phi)^T = Z(\phi). \tag{3.2-45}$$

The action of the counterpropagation operator on a sequence of linear retarders can therefore be expressed as

$$\begin{aligned}\Theta[\mathbf{G}_{\text{linear}}] &= \Theta[\mathbf{W}_{\eta_1}(\phi_1)\mathbf{W}_{\eta_2}(\phi_2)\dots\mathbf{W}_{\eta_N}(\phi_N)] \\ &= \mathbf{W}_{\eta_N}(-\phi_N)\dots\mathbf{W}_{\eta_2}(-\phi_2)\mathbf{W}_{\eta_1}(-\phi_1) \\ &= Z(\pm\pi)\mathbf{W}_{\eta_N}(\phi_N)\dots\mathbf{W}_{\eta_2}(\phi_2)\mathbf{W}_{\eta_1}(\phi_1)Z(\mp\pi) \\ &= Z(\pm\pi)[\mathbf{W}_{\eta_1}(\phi_1)\mathbf{W}_{\eta_2}(\phi_2)\dots\mathbf{W}_{\eta_N}(\phi_N)]^T Z(\mp\pi) \\ &= Z(\pm\pi)\mathbf{G}_{\text{linear}}^T Z(\mp\pi),\end{aligned} \tag{3.2-46}$$

where the sandwich identity (3.2-10) was used in second step. Using this relation one can see how a given parametrisation of a unitary $U \in \mathrm{SU}(2)$ changes under Θ. Take for example the Tait-Bryan decomposition of a rotation, which is defined as three successive rotations about all three main axes:

$$\mathbf{G}_{\text{linear}} = X(\alpha)Y(\beta)Z(\gamma). \tag{3.2-47}$$

In the counterpropagating direction this parametrisation becomes

$$\begin{aligned}\Theta[\mathbf{G}_{\text{linear}}] &= Z(\pm\pi)\mathbf{G}_{\text{linear}}^T Z(\mp\pi) \\ &= Z(\pm\pi)\big(X(\alpha)Y(\beta)Z(\gamma)\big)^T Z(\mp\pi) \\ &= Z(\pm\pi)Z(\gamma)^T Y(\beta)^T X(\alpha)^T Z(\mp\pi) \\ &= Z(\pm\pi)Z(\gamma)Y(-\beta)X(\alpha)Z(\mp\pi) \\ &= Z(\pm\pi)Z(\gamma)Z(\mp\pi)Y(\beta)X(-\alpha) \\ &= Z(\gamma)Y(\beta)X(-\alpha).\end{aligned} \tag{3.2-48}$$

Here (3.2-9) was used in the second-to-last step. What the equation above shows is that, in addition to the permutation of the rotations, the X-component in the Tait-Bryan decomposition experiences a sign change. A different way to parametrise a unitary is the axis-angle parametrisation (3.2-4). Expressing an arbitrary unitary this way, the parametrisation in the counterpropagating direction is

$$\begin{aligned}\Theta[\mathbf{G}_{\text{linear}}] &= Z(\pi)\Big[\cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}\big(n_x X + n_y Y + n_z Z\big)\Big]^T Z(-\pi) \\ &= \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}\big(-n_x X + n_y Y + n_z Z\big),\end{aligned} \tag{3.2-49}$$

and we see that the X-component acquires a sign change in this parametrisation as well. This result is quite natural since the changes to the parametrisations have to agree in the trivial case of rotations about one of the main axes. A consequence of (3.2-48) and

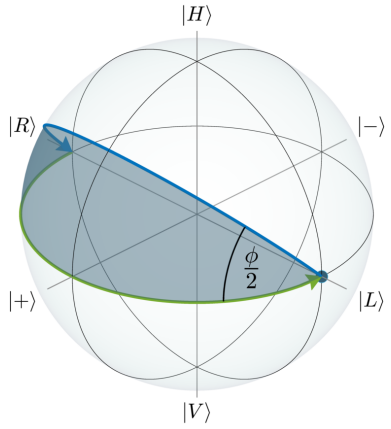


Figure 3.3: **Illustration of how the state $|L\rangle$ evolves under the action of the $\mathbf{Y}^S(\phi)$ gadget.** The rotation angle of the transformation in the figure is $\phi = \pi/2$, which is twice the physical angle between the two half-wave plates that constitute the gadget. Since the state evolves solely along geodesics, only geometric phase is acquired. In this figure, the state initially travels along the blue curve to $|R\rangle$, and then back to $|L\rangle$ along the green curve that lies on the equator. In the process the state acquires a geometric phase of $\phi_g = -\pi/4$, which is equal in magnitude to the angular width of the ungula represented by the shaded area. This shaded area also corresponds to the area enclosed by the trajectory of the state. The signed area of the wedge is $\Omega = -\pi/2 = 2\phi_g$, and the rotation generated by the gadget is $Y(-\Omega) = \exp\big(i\frac{\Omega}{2}Y\big)$. Adding an equal angle offset to both half-wave plates in the gadget has the effect of rotating the ungula, but does not affect the acquired phase.

(3.2-49) is that given a gadget consisting of linear retarders implementing some unitary U, the transformation of this unitary under counterpropagation is determined only by its parametrisation, and is independent of what sequence of linear retarders was used to actually implement U. This seemingly trivial property is not true for more general polarization gadgets.

3.2.5 The geometric phase

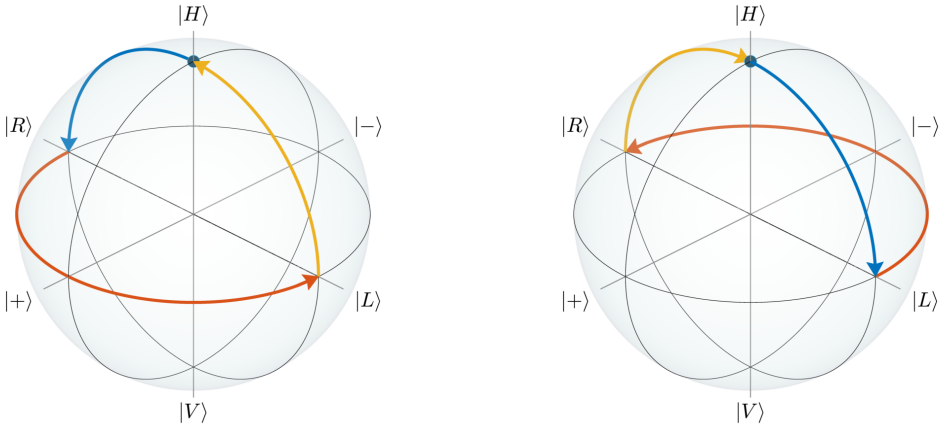
The rotations produced by all three of the gadgets described in Section 3.2.3 result from a phase shift that is geometric in origin, known as the Pancharatnam phase, which is one of the earliest examples of a Berry phase [379,380,381,382]. The observation made by Pancharatnam and Berry relevant to this discussion relates to the phase acquired by states on the Bloch sphere evolving in time in such a way that they trace out a closed path.[1] Such a closed trajectory is equivalent to the initial and final states $|\Psi_i\rangle$ and $|\Psi_f\rangle$ having unity overlap, i.e. $|\langle\Psi_i|\Psi_f\rangle| = 1$. Pancharatnam and Berry realised that the phase difference between these initial and final states

$$\phi = \arg\big[\langle\Psi_i|\Psi_f\rangle\big], \tag{3.2-50}$$

has a geometric contribution ϕ_g, which is given by half the solid angle subtended by the path C:

$$\phi_g = \frac{\Omega_C}{2}. \tag{3.2-51}$$

[1] If the path is open the geometric phase can be found by closing it with a geodesic.



(a) $\mathbf{Z}^S(\pi)$ (b) $\Theta[\mathbf{Z}^S(\pi)]$

Figure 3.4: **Evolution of the state $|H\rangle$ under the $\mathbf{Z}^S(\phi)$ gadget in the forwards (a) and backwards (b) propagation directions.** The trajectory of the state is essentially the same as for the $\mathbf{Y}^S$ gadget shown in Fig. 3.3, with the difference being the starting point, indicated by a blue dot, and total enclosed area. Even though the trajectories in the two propagation directions (a) and (b) are different, the area enclosed by the curves, and therefore the acquired geometric phase, is the same. Visually this can be seen by the fact that the two trajectories are related by a 180° rotation about the z-axis, and rotations preserve the signed area. In this example, it is easy to see that the area enclosed by the curve is $\Omega = -\pi$, and the corresponding unitary is $Z(\pi)$.

Since the Bloch sphere is a unit sphere the solid angle Ω_C is simply equal to the signed area enclosed by the curve C on the surface of the Bloch sphere. The total phase ϕ is the sum of the geometric phase and the dynamical phase ϕ_d [381, 383, 384]:

$$\phi = \phi_g + \phi_d. \tag{3.2-52}$$

The dynamical phase is given as an integral over the expectation value of the Hamiltonian:

$$\phi_d = -\frac{1}{\hbar}\int_0^t \langle H(t)\rangle dt. \tag{3.2-53}$$

In the case of a photon passing through a series of birefringent elements the Hamiltonian can be decomposed into a sequence of time-independent Hamiltonians H_k, and for these trivially time-dependent Hamiltonians the expectation value can be moved outside the integral:

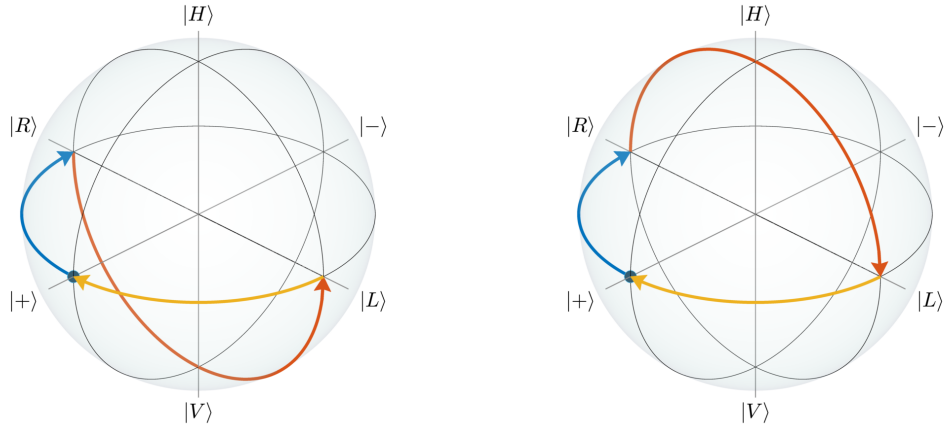
$$-\frac{1}{\hbar}\int_{t_{k-1}}^{t_k} \langle H_k\rangle dt = \left\langle -\frac{1}{\hbar}\int_{t_{k-1}}^{t_k} H_k dt \right\rangle. \tag{3.2-54}$$

In the eigenbasis of the Hamiltonian the time integral becomes

$$-\frac{1}{\hbar}\int_{t_{k-1}}^{t_k} H_k dt = \begin{bmatrix} -\delta/2 & 0 \\ 0 & \delta/2 \end{bmatrix}, \tag{3.2-55}$$

where δ is the total rotation angle on the Bloch sphere under the Hamiltonian. For an initial state

$$|\Psi\rangle = c_+ |+1\rangle + c_- |-1\rangle, \tag{3.2-56}$$



(a) $\mathbf{X}^A(\pi)$ (b) $\Theta[\mathbf{X}^A(\pi)]$

Figure 3.5: **Action of the $\mathbf{X}^A(\phi)$ gadget in the two propagation directions.** Unlike the $\mathbf{Z}^S(\phi)$ gadget shown in Fig. 3.4, the trajectories in the forwards (a) and backwards (b) propagation directions are not related by a rotation, but instead by a mirroring in the $X-Y$ plane. This mirroring changes the sign of the enclosed area, and therefore the geometric phase, and is one way to understand the nonreciprocal behaviour $\Theta\big[\mathbf{X}^A(\pi)\big] = X(-\pi)$ of the gadget.

written in the eigenbasis of the Hamiltonian the dynamical phase is then

$$\phi_d = -\frac{\delta}{2}\big(|c_+|^2 - |c_-|^2\big). \tag{3.2-57}$$

If the initial state is in a balanced superposition of the two eigenstates of the Hamiltonian then $|c_+|^2 - |c_-|^2 = 0$ and hence the dynamical phase vanishes:

$$\phi_d = 0. \tag{3.2-58}$$

States that are a balanced superposition of the eigenstates of the Hamiltonian have a 90° angle to these eigenstates on the Bloch sphere. Consequently the path they trace out as they evolve under the Hamiltonian is a geodesic. This leads to the following conclusion: any state that evolves along geodesics acquires zero dynamical phase. The eigenstates of the gadgets presented in the Section 3.2.3 evolve exclusively along geodesics, and the action of these gadgets, and their symmetry properties, can therefore be understood geometrically, and it was in fact through geometric arguments that they were first proposed [384].

Examples of the geodesic trajectories produced by the three gadgets for rotations about the main axes of the Bloch sphere, described in Section 3.2.3, are shown in Figs. 3.3-3.5. In Fig. 3.3 the trajectory traced out by the state $|L\rangle$, as well as the enclosed area, under the evolution of a $\mathbf{Y}^S$ gadget is shown. The area of a spherical wedge on the unit sphere is 2α, where α is the angle between the two 'cuts'. This angle is exactly $\phi - \pi/2$, where ϕ is the difference in angle between the two half-wave plates. As an example to illustrate this, if one HWP is at 0° and the other at 90° to the vertical axis, this corresponds to rotations about the positive and negative z-axis in the figure, which obviously does not enclose any area. Rotating both wave plates by the same

n_x	n_y	n_z	Operator
$+$	$+$	$+$	U
$+$	$+$	$-$	XU^TX
$+$	$-$	$+$	U^T
$+$	$-$	$-$	XUX
$-$	$+$	$+$	ZU^TZ
$-$	$+$	$-$	$YUY = U^*$
$-$	$-$	$+$	ZUZ
$-$	$-$	$-$	$U^\dagger$

Table 3.1: The table shows the different symmetries that a unitary operator U can have under counterpropagtion, under the condition that the magnitudes of the elements n_k of the vector $\vec{n}$, defining the axis of rotation on the Bloch sphere, remain unchanged. The possible symmetries are therefore all the ways to introduce a sign change, or not, in the different components n_k, and these are shown in the first three columns. The last column shows the unitary transformation resulting from the different combinations of sign changes.

amount would have the effect of rotating the trajectory about the y-axis, and would not affect the enclosed area.

In Fig. 3.4 the $\mathbf{Z}^S$ gadget is illustrated in both propagation directions. All three constituent rotations are different in the counterpropagating frame, however they still produce a trajectory that is simply a rotated version of the trajectory traced out in the forwards direction, and therefore the gadget is reciprocal. This is in contrast to the $\mathbf{X}^A$ gadget shown in Fig. 3.5. The trajectory traced out by the eigenstate $|+\rangle$ of the gadget resembles the one that is traced out by the $|H\rangle$ under the $\mathbf{Z}^S$ gadget, however under counterpropagation this trajectory becomes mirrored in the $X-Y$ plane, resulting in a sign change in the enclosed area.

3.3 General gadgets

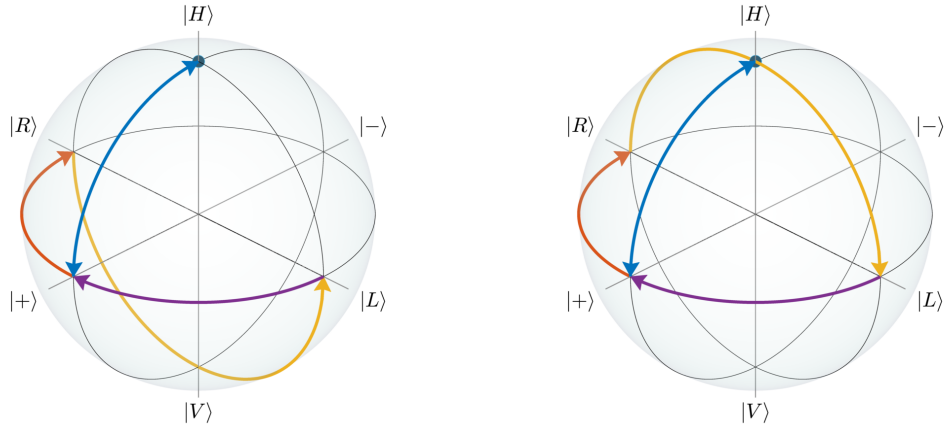
As discussed in sections 3.2.4 and 3.2.3 any qubit unitary implemented by a gadget consisting solely of linear retarders transform as

$$\begin{bmatrix}\theta & n_x & n_y & n_z\end{bmatrix} \mapsto \begin{bmatrix}\theta & -n_x & n_y & n_z\end{bmatrix} \tag{3.3-1}$$

under counterpropagation. However, there are many other similar transformations that preserve the magnitude of the n_k components of $\vec{n}$. These transformations, together with the unitary generated under the transformation in question, are summarised in Table 3.1. One might now ask if it is possible to create polarization gadgets that have some of these other symmetries under counterpropagation, and this question will be explored in the following sections.

3.3.1 Pauli gadgets with Faraday rotators

In Section 3.2.3 we explained how the gadgets $\mathbf{X}^A(\phi)$, $\mathbf{Y}^S(\phi)$ and $\mathbf{Z}^S(\phi)$ can be constructed using short sequences of wave plates. As a starting point for filling out Table 3.1 with possible gadgets, we will begin by constructing the hitherto missing



(a) $\mathbf{Z}^A(\phi)$ (b) $\Theta[\mathbf{Z}^A(\phi)]$

Figure 3.6: **Trajectory of $|H\rangle$ under the anti-symmetric gadget $\mathbf{Z}^A(\phi)$ in both propagation directions.** The rotation angle of the gadget in the figure is $\phi = \pi$. The blue path is traversed twice, and the order in which the states are visited is: $|H\rangle \mapsto |+\rangle \mapsto |R\rangle \mapsto |L\rangle \mapsto |+\rangle \mapsto |H\rangle$. Just as for $\mathbf{Z}^S(\phi)$ the trajectory is entirely geodesic and the phase is therefore geometric.

Pauli gadgets $\mathbf{X}^S(\phi)$, $\mathbf{Y}^A(\phi)$ and $\mathbf{Z}^A(\phi)$. First, it is worth commenting that since a Faraday rotator obeys

$$\Theta\big[\mathbf{F}_\psi\big] = \mathbf{F}_{-\psi}, \tag{3.3-2}$$

such a rotator with a tunable retardance would in fact be a realisation of $\mathbf{Y}^A(\phi)$. However, due to the fact that the vast majority of applications of Faraday rotators require a fixed circular retardance of $\psi = \pi/2$, as well as the fact that packaging a tunable magnet with magnetic field strengths of 1 T or more in a small tabletop optic is challenging, Faraday rotators are in practice sold with permanent rare-earth magnets [385], and consequently have a fixed retardance. In the remainder of the chapter we will therefore restrict our discussion to these $\mathbf{F}_\pm$ devices.

Out of the missing Pauli gadgets the simplest one to construct is $\mathbf{Z}^A(\phi)$. By using two Faraday rotators to 'sandwich' a $\mathbf{X}^A(\phi)$ gadget it is converted to a $Z(\phi)$ rotation with the same symmetry properties:

$$\mathbf{Z}^A(\phi) = \mathbf{F}_-\mathbf{X}^A(\phi)\mathbf{F}_+. \tag{3.3-3}$$

The action of the gadget in the forward direction can be verified using the permutation identity (3.2-7):

$$\begin{aligned}\mathbf{Z}^A(\phi) &= \mathbf{F}_-\mathbf{X}^A(\phi)\mathbf{F}_+ \\ &= Y(-\pi/2)X(\phi)Y(\pi/2) \\ &= Y(-\pi/2)Y(\pi/2)Z(\phi) \\ &= Z(\phi).\end{aligned} \tag{3.3-4}$$

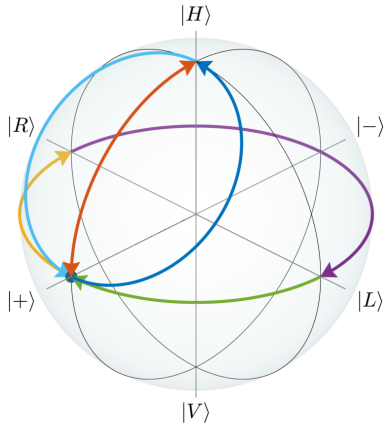


Figure 3.7: **Evolution of $|+\rangle$ under the gadget $\mathbf{X}^S(\phi)$ for $\phi = 0$.** Unlike the trajectory under the $\mathbf{X}^A(\phi)$ gadget shown in Fig. 3.5, the trajectory does not exclusively consist of geodesics, and the smaller circular trajectory $|0\rangle \mapsto |H\rangle \mapsto |+\rangle$ contributes dynamical phase. This cancels the 2π geometric phase from the great circle, and gives zero net phase. The state evolution is: $|+\rangle \mapsto |H\rangle \mapsto |+\rangle \mapsto |R\rangle \mapsto |L\rangle \mapsto |+\rangle \mapsto |H\rangle \mapsto |+\rangle$.

It remains to confirm that the gadget is anti-symmetric:

$$\begin{aligned}\Theta\big[\mathbf{Z}^A(\phi)\big] &= \Theta\big[\mathbf{F}_+\big]\Theta\big[\mathbf{X}^A(\phi)\big]\Theta\big[\mathbf{F}_-\big] \\ &= Y(-\pi/2)X(-\phi)Y(\pi/2) \\ &= Z(-\phi).\end{aligned} \tag{3.3-5}$$

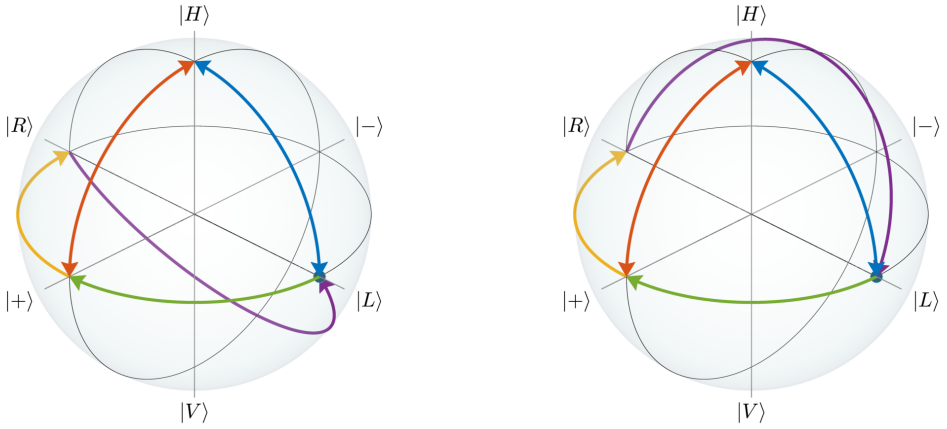
Using the $\mathbf{Z}^A(\phi)$ gadget as a building block, it then becomes possible to construct an $\mathbf{X}^S(\phi)$ gadget by sandwiching the $\mathbf{Z}^A(\phi)$ gadget with two half-wave plates:

$$\begin{aligned}\mathbf{X}^S(\phi) &= \mathbf{H}(\pi/8)\mathbf{Z}^A(\phi - 2\pi)\mathbf{H}(\pi/8) \\ &= Y(\pi/4)Z(\pi)Y(-\pi/4)Z(\phi - 2\pi)Y(\pi/4)Z(\pi)Y(-\pi/4) \\ &= Y(\pi/2)Z(\pi)Z(\phi - 2\pi)Z(\pi)Y(-\pi/2) \\ &= Y(\pi/2)Z(\phi)Y(-\pi/2) \\ &= X(\phi).\end{aligned} \tag{3.3-6}$$

The symmetry of the gadget can be verified using the same steps:

$$\begin{aligned}\Theta\big[\mathbf{X}^S(\phi)\big] &= \Theta\big[\mathbf{H}(\pi/8)\big]\Theta\big[\mathbf{Z}^A(\phi - 2\pi)\big]\Theta\big[\mathbf{H}(\pi/8)\big] \\ &= \mathbf{H}(-\pi/8)\mathbf{Z}^A(-\phi + 2\pi)\mathbf{H}(-\pi/8) \\ &= Y(-\pi/4)Z(\pi)Y(\pi/4)Z(-\phi + 2\pi)Y(-\pi/4)Z(\pi)Y(\pi/4) \\ &= Y(-\pi/2)Z(-\phi + 4\pi)Y(\pi/2) \\ &= Y(-\pi/2)Z(-\phi)Y(\pi/2) \\ &= X(\phi).\end{aligned} \tag{3.3-7}$$

The final gadget $\mathbf{Y}^S(\phi)$ can be constructed in a way similar to the $\mathbf{X}^S(\phi)$ gadget, but



(a) $\mathbf{Y}^A(\phi)$ (b) $\Theta[\mathbf{Y}^A(\phi)]$

Figure 3.8: **Trajectories for $|L\rangle$ under the anti-symmetric $\mathbf{Y}^A(\phi)$ gadget for $\phi = 4\pi/3$.** Like for $\mathbf{Y}^A(\phi)$ the trajectory consists entirely of geodesics. The state evolution through the gadget is: $|L\rangle \mapsto |H\rangle \mapsto |+\rangle \mapsto |R\rangle \mapsto |L\rangle \mapsto |+\rangle \mapsto |H\rangle \mapsto |L\rangle$. Since the path $|L\rangle \mapsto |H\rangle \mapsto |+\rangle$ is traversed twice in two different directions, the spherical triangle $|L\rangle\,|H\rangle\,|+\rangle$ does not contribute geometric phase.

with a quarter-wave plate sandwich instead:

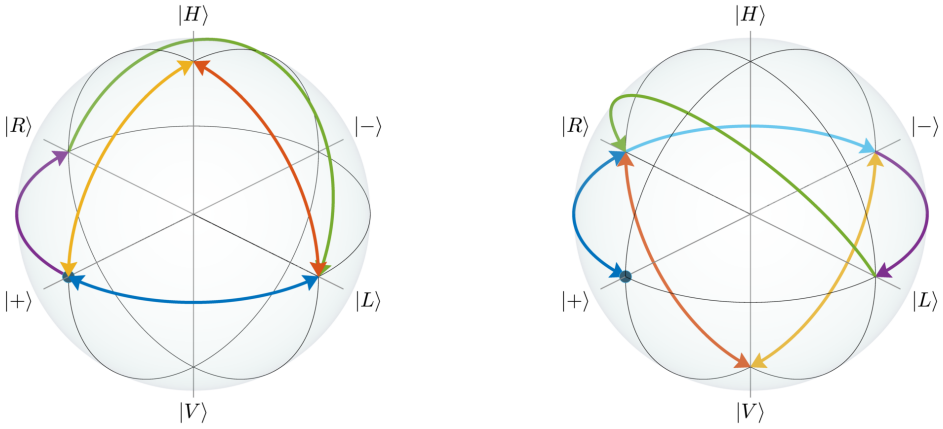
$$\begin{aligned}\mathbf{Y}^A(\phi) &= \mathbf{Q}(-\pi/4)\mathbf{Z}^A(\phi)\mathbf{Q}(\pi/4)\\ &= Y(-\pi/2)Z(\pi/2)Y(\pi/2)Z(\phi)Y(\pi/2)Z(\pi/2)Y(-\pi/2)\\ &= Y(-\pi/2)Y(\pi/2)X(-\pi/2)Z(\phi)X(\pi/2)Y(\pi/2)Y(-\pi/2)\\ &= X(-\pi/2)Z(\phi)X(\pi/2)\\ &= Y(\phi).\end{aligned} \tag{3.3-8}$$

For this gadget, the anti-symmetry follows immediately:

$$\begin{aligned}\Theta\big[\mathbf{Y}^A(\phi)\big] &= \Theta\big[\mathbf{Q}(\pi/4)\big]\Theta\big[\mathbf{Z}^A(\phi)\big]\Theta\big[\mathbf{Q}(-\pi/4)\big]\\ &= \mathbf{Q}(-\pi/4)\mathbf{Z}^A(-\phi)\mathbf{Q}(\pi/4)\\ &= Y(-\phi).\end{aligned} \tag{3.3-9}$$

The trajectories produced by the gadgets $\mathbf{Z}^A(\phi)$, $\mathbf{X}^S(\phi)$ and $\mathbf{Y}^A(\phi)$ are shown in Figs. 3.6-3.8. It can be seen that while $\mathbf{Z}^A(\phi)$ and $\mathbf{Y}^A(\phi)$ produce geodesic trajectories, $\mathbf{X}^S(\phi)$ does not. At the cost of an additional two wave plates, however, it is possible to produce an $\mathbf{X}^S(\phi)$ gadget that is entirely geodesic:

$$\begin{aligned}\mathbf{X}^S_{\text{geo}}(\phi) &= \mathbf{Q}(\pi/2)\mathbf{Y}^A(\phi)\mathbf{Q}(0)\\ &= Y(\pi)Z(\pi/2)Y(-\pi)Y(\phi)Z(\pi/2)\\ &= Y(\pi)Y(-\pi)Z(-\pi/2)Z(\pi/2)X(\phi)\\ &= X(\phi).\end{aligned} \tag{3.3-10}$$



(a) $\mathbf{X}^S_{\text{geo}}(\phi)$ (b) $\Theta\big[\mathbf{X}^S_{\text{geo}}(\phi)\big]$

Figure 3.9: **Trajectories for $|+\rangle$ under the symmetric $\mathbf{X}^S_{\text{geo}}(\phi)$ gadget for $\phi = -4\pi/3$ in both propagation directions.** The trajectory is similar to that of the $\mathbf{Y}^A(\phi)$ gadget shown in Fig. 3.8, and unlike the $\mathbf{X}^S(\phi)$ gadget shown in Fig. 3.7 the state evolves entirely along geodesics. Since the area spanned by the two half-circle arcs in the two different propagation directions is related by a rotation around the z-axis the gadget is symmetric. Note that while the spherical triangle in (b) appears to be mirrored with respect to (a), the path $|+\rangle \mapsto |H\rangle \mapsto |L\rangle$ in (a) is traversed in both directions, and the spherical triangle does therefore not contribute any signed area, and the same argument holds for (b). The state evolution in (a) is: $|+\rangle \mapsto |L\rangle \mapsto |H\rangle \mapsto |+\rangle \mapsto |R\rangle \mapsto |L\rangle \mapsto |+\rangle \mapsto |H\rangle \mapsto |L\rangle \mapsto |+\rangle$. And for (b): $|+\rangle \mapsto |R\rangle \mapsto |V\rangle \mapsto |-\rangle \mapsto |L\rangle \mapsto |R\rangle \mapsto |-\rangle \mapsto |V\rangle \mapsto |R\rangle \mapsto |+\rangle$.

The symmetry of this gadget can also be verified:

$$\begin{aligned}\Theta\big[\mathbf{X}^S_{\text{geo}}(\phi)\big] &= \Theta\big[\mathbf{Q}(0)\big]\Theta\big[\mathbf{Y}^A(\phi)\big]\Theta\big[\mathbf{Q}(\pi/2)\big] \\ &= Q(0)Y(-\phi)Q(-\pi/2) \\ &= \big[Q(\pi/2)Y(\phi)Q(0)\big]^T \\ &= X(\phi)^T \\ &= X(\phi).\end{aligned} \tag{3.3-11}$$

where the second step used $Q(\pi/2) = Q(-\pi/2)$. Examples of the geodesic trajectories under this gadget in the two propagation directions are shown in Fig. 3.9.

3.3.2 Transpose gadget

Using the Pauli gadgets presented in the previous section together with the wave plate only gadgets from Section 3.2.3 it becomes possible to construct gadgets with new symmetries, that are capable of realising any $U \in \mathrm{SU}(2)$. One of the simplest of these is the transpose gadget, that obeys

$$\Theta\big[\mathbf{G}_T\big] = (\mathbf{G}_T)^T. \tag{3.3-12}$$

To construct the gadget, one first picks an Euler or Tait-Bryan decomposition of an arbitrary unitary U, for example

$$U = Z(\alpha)Y(\beta)Z(\gamma). \tag{3.3-13}$$

One then finds the transpose of the decomposition:

$$\begin{aligned} U^T &= Z(\gamma)^T Y(\beta)^T Z(\alpha)^T \\ &= Z(\gamma)Y(-\beta)Z(\alpha). \end{aligned} \tag{3.3-14}$$

From this one sees which symmetry property each sub-gadget needs. In this case the gadget would be

$$\mathbf{G}_T = \mathbf{Z}^S(\alpha)\mathbf{Y}^A(\beta)\mathbf{Z}^S(\gamma), \tag{3.3-15}$$

which is easily confirmed:

$$\begin{aligned} \Theta\big[\mathbf{G}_T\big] &= \Theta\big[\mathbf{Z}^S(\gamma)\big]\Theta\big[\mathbf{Y}^A(\beta)\big]\Theta\big[\mathbf{Z}^S(\alpha)\big] \\ &= Z(\gamma)Y(-\beta)Z(\beta) \\ &= \big[Z(\alpha)Y(\beta)Z(\gamma)\big]^T \\ &= (\mathbf{G}_T)^T. \end{aligned} \tag{3.3-16}$$

Other sets of Euler angles, such as $U = Y(\alpha)X(\beta)Y(\gamma)$, could also be used to construct a gadget with the same symmetry, however for this decomposition the resulting gadget would be

$$\mathbf{G}_T = \mathbf{Y}^A(\alpha)\mathbf{X}^S(\beta)\mathbf{Y}^A(\gamma), \tag{3.3-17}$$

which consists entirely of the gadgets from Section 3.3.1. These gadgets contain two Faraday rotators each, while the gadget in (3.3-15) only has two such rotators in total. The construction above is therefore less efficient in terms of the number of required optical components.

It is worth noting that the transpose is a basis-dependent operation, and therefore also depends on the convention used for the Stokes parameters and Pauli matrices. As shown in Publication 3, particular choices of convention can lead to simpler realisations of the transpose that do not require any Faraday rotators.

3.3.3 Adjoint gadget

Another gadget that can be constructed in a straightforward manner is the adjoint, or inverse, gadget, which under counterpropagation transforms as

$$\Theta\big[\mathbf{G}_A\big] = \mathbf{G}_A^\dagger = \mathbf{G}_A^{-1}. \tag{3.3-18}$$

The inverse of any rotation $R_{\vec{n}}(\theta)$ is simply $R_{\vec{n}}(-\theta)$, and this condition is equivalent to

$$\vec{n} \mapsto -\vec{n}. \tag{3.3-19}$$

The gadget should therefore be constructed entirely out of anti-symmetric sub-gadgets. Since $\mathbf{X}^A(\phi)$ can be constructed just from wave plates, the most resource efficient implementation is one that utilizes two X-rotations. One possible choice is

$$\mathbf{G}_A = \mathbf{X}^A(\alpha)\mathbf{Z}^A(\beta)\mathbf{X}^A(\gamma). \tag{3.3-20}$$

The property (3.3-18) can be verified by direct calculation:

$$\begin{aligned}\Theta\big[\mathbf{G}_A\big] &= \Theta\big[\mathbf{X}^A(\gamma)\big]\Theta\big[\mathbf{Z}^A(\beta)\big]\Theta\big[\mathbf{X}^A(\alpha)\big] \\ &= X(-\gamma)Z(-\beta)X(-\alpha) \\ &= X(\gamma)^\dagger Z(\beta)^\dagger X(\alpha)^\dagger \\ &= \big[X(\alpha)Z(\beta)X(\gamma)\big]^\dagger \\ &= \mathbf{G}_A^\dagger.\end{aligned} \tag{3.3-21}$$

3.3.4 Reciprocal gadget

A particularly useful gadget is the fully reciprocal gadget, i.e. the one which is fully symmetric under counterpropagation:

$$\Theta\big[\mathbf{G}_\mathrm{R}\big] = \mathbf{G}_\mathrm{R}. \tag{3.3-22}$$

One approach to construct such a gadget would be to start with a palindromic order of symmetric sub-gadgets:

$$\mathbf{G}_\mathrm{R} \stackrel{?}{=} \mathbf{Y}^S(\gamma)\mathbf{Z}^S(\beta)\mathbf{X}^S(\alpha)\mathbf{Z}^S(\beta)\mathbf{Y}^S(\gamma). \tag{3.3-23}$$

This gadget clearly obeys (3.3-22), however since the rotations do not constitute a set of Euler or Tait-Bryan rotations it is not immediately obvious that the gadget is capable of realising any unitary. To show that it is universal it is sufficient to show that it can apply an arbitrary phase $\pm\lambda/2$ to an arbitrary eigenstate $|v\rangle$:

$$Y(\gamma)Z(\beta)X(\alpha)Z(\beta)Y(\gamma)\,|v\rangle = e^{-i\lambda/2}\,|v\rangle \tag{3.3-24}$$

as this is equivalent to being able to perform any rotation on the Bloch sphere. This can be done in the following way. Let

$$G = Y(\gamma)Z(\beta)X(\alpha)Z(\beta)Y(\gamma), \tag{3.3-25}$$

Then multiply both sides by $X(\pi)$:

$$X(\pi)G = X(\pi)Y(\gamma)Z(\beta)X(\alpha)Z(\beta)Y(\gamma). \tag{3.3-26}$$

Using (3.2-10) this can be re-written as

$$\begin{aligned}X(\pi)G &= X(\pi)Y(\gamma)Z(\beta)X(\alpha)Z(\beta)Y(\gamma) \\ &= X(\pi)Y(\gamma)X(-\pi)X(\pi)Z(\beta)X(-\pi)X(\pi)X(\alpha)Z(\beta)Y(\gamma) \\ &= Y(-\gamma)Z(-\beta)X(\alpha+\pi)Z(\beta)Y(\gamma) \\ &= Y(-\gamma)Z(-\beta)X(\alpha')Z(\beta)Y(\gamma),\end{aligned} \tag{3.3-27}$$

where $\alpha' = \alpha + \pi$. The first three rotations $X(\alpha')Z(\beta)Y(\gamma)$ are a set of Tait-Bryan rotations and are therefore universal for SU(2). Let α', β and γ be chosen such that

$$X(\alpha')Z(\beta)Y(\gamma)\,|u\rangle = e^{i\phi}\,|+\rangle\,, \tag{3.3-28}$$

for a state $|u\rangle$ that is the eigenstate of some $U \in \mathrm{SU}(2)$:

$$U|u\rangle = e^{-i\lambda/2}|u\rangle\,. \tag{3.3-29}$$

The phase ϕ can be decomposed into contributions from the first two rotations and the X rotation:

$$\begin{aligned} & X(-\alpha')X(\alpha')Z(\beta)Y(\gamma)|u\rangle = X(-\alpha')e^{i\phi}|+\rangle \\ \iff \quad & Z(\beta)Y(\gamma)|u\rangle = e^{i(\phi+\alpha'/2)}|+\rangle \\ & \qquad\qquad = e^{i\delta}|+\rangle\,. \end{aligned} \tag{3.3-30}$$

Here we identify $\phi = \delta - \alpha'/2$ and see that δ is the phase from the $Z(\beta)Y(\gamma)$ rotation. It then follows that

$$\begin{aligned} X(\pi)G|u\rangle &= Y(-\gamma)Z(-\beta)X(\alpha')Z(\beta)Y(\gamma)|v\rangle \\ &= Y(-\gamma)Z(-\beta)e^{i(\delta-\alpha'/2)}|+\rangle \\ &= e^{-i\alpha'/2}|u\rangle \end{aligned} \tag{3.3-31}$$

Setting $\alpha' = \lambda$ gives

$$X(\pi)G|u\rangle = e^{-i\lambda/2}|u\rangle = U|u\rangle\,. \tag{3.3-32}$$

Since α' and U are arbitrary this shows that $X(\pi)G$ is universal, and by choosing $U = X(\pi)V$ for any $V \in \mathrm{SU}(2)$ it follows that

$$G = X(-\pi)X(\pi)V = V, \tag{3.3-33}$$

and hence the sequence of rotations in (3.3-24) is universal as well. A direct realisation of the construction (3.3-23) only requires two Faraday rotators, since they only appear in the $\mathbf{X}^S(\alpha)$ gadget, and since any sequence of wave plates is equivalent to a three-component Simon–Mukunda gadget, at most three wave plates are required on either side of the Faraday rotators. This, however, turns out to not be optimal. To show why, it is first necessary to characterise a two-wave-plate gadget on the form $\mathbf{Q}(\phi)\mathbf{H}(\theta)$. In Ref. [161] it was shown that this gadget implements a two-parameter subset of SU(2) parametrised by $Y(\alpha)Z(\pi/2)Y(\beta)$, and this is straightforward to confirm:

$$\begin{aligned} \mathbf{Q}(\phi)\mathbf{H}(\theta) &= Y(2\phi)Z(\pi/2)Y(-2\phi)Y(2\theta)Z(\pi)Y(-2\theta) \\ &= Y(2\phi+\pi)Z(-\pi/2)Y(-2\phi-\pi)Y(2\theta)Z(\pi)Y(-2\theta) \\ &= Y(2\phi+\pi)Z(\pi/2)Y(2\phi+\pi-4\theta) \\ &= Y(\alpha)Z(\pi/2)Y(\beta). \end{aligned} \tag{3.3-34}$$

This parametrisation can equivalently be expressed in a form that is more useful for the problem at hand:

$$\begin{aligned} \mathbf{Q}(\phi)\mathbf{H}(\theta) &= Y(\alpha)Z(\pi/2)Y(\beta) \\ &= Y(\alpha-\pi)Y(\pi)Z(\pi/2)Y(\pi)Y(\beta-\pi) \\ &= Y(\alpha-\pi)Y(\pi/2)X(\pi/2)Y(3\pi/2)Y(\beta-\pi) \\ &= Y(\alpha-\pi/2)X(\pi/2)Y(\beta+\pi/2) \\ &= Y(\alpha-\pi/2)Z(\beta+\pi/2)X(\pi/2) \\ &= Y(\gamma)Z(\delta)X(\pi/2). \end{aligned} \tag{3.3-35}$$

In the backwards propagation direction this becomes

$$\begin{aligned}\mathbf{H}(-\theta)\mathbf{Q}(-\phi) &= \Theta\big[\mathbf{Q}(\phi)\mathbf{H}(\theta)\big] \\ &= Z(\pi)\big[Y(\gamma)Z(\delta)X(\pi/2)\big]^T Z(-\pi) \\ &= X(-\pi/2)Z(\delta)Y(\gamma).\end{aligned} \tag{3.3-36}$$

Consider the gadget

$$\mathbf{G}_{\mathrm{R}} = \mathbf{Q}(\phi)\mathbf{H}(\theta)\mathbf{X}^S(\psi)\mathbf{H}(-\theta)\mathbf{Q}(-\phi). \tag{3.3-37}$$

Using the parametrisation (3.3-35) the unitary it implements can be written as

$$\begin{aligned}\mathbf{G}_{\mathrm{R}} &= Y(\gamma)Z(\delta)X(\pi/2)X(\psi)X(-\pi/2)Z(\delta)Y(\gamma) \\ &= Y(\gamma)Z(\delta)X(\psi)Z(\delta)Y(\gamma),\end{aligned} \tag{3.3-38}$$

which is equivalent to (3.3-24) and therefore universal. The reciprocity of the gadget can also be verified:

$$\begin{aligned}\Theta\big[\mathbf{G}_{\mathrm{R}}\big] &= \Theta\big[\mathbf{Q}(-\phi)\big]\Theta\big[\mathbf{H}(-\theta)\big]\Theta\big[\mathbf{X}^S(\psi)\big]\Theta\big[\mathbf{H}(\theta)\big]\Theta\big[\mathbf{Q}(\phi)\big] \\ &= \mathbf{Q}(\phi)\mathbf{H}(\theta)\mathbf{X}^S(\psi)\mathbf{H}(-\theta)\mathbf{Q}(-\phi) \\ &= \mathbf{G}_{\mathrm{R}}.\end{aligned} \tag{3.3-39}$$

The explicit form of this gadget is

$$\begin{aligned}&\mathbf{G}_{\mathrm{R}} = \\ &\mathbf{Q}(\phi)\mathbf{H}(\theta)\mathbf{H}(\pi/8)\mathbf{F}_-\mathbf{Q}(\pi/2)\mathbf{H}(\psi/4-\pi/2)\mathbf{Q}(\pi/2)\mathbf{F}_+\mathbf{H}(\pi/8)\mathbf{H}(-\theta)\mathbf{Q}(-\phi),\end{aligned} \tag{3.3-40}$$

which using the wave plate reduction rules (3.2-13) and (3.2-14) can be reduced to

$$\mathbf{G}_{\mathrm{R}} = \mathbf{Q}(\phi_1)\mathbf{H}(\theta_1)\mathbf{F}_-\mathbf{Q}(\pi/2)\mathbf{H}(\psi/4-\pi/2)\mathbf{Q}(\pi/2)\mathbf{F}_+\mathbf{H}(\theta_2)\mathbf{Q}(\phi_2). \tag{3.3-41}$$

Due to the results presented in Section 3.2.4 such reductions do not change the symmetry properties of the sub-gadgets, and therefore not the overall gadget either.

3.3.5 The *XUX* gadget

As discussed in Section 3.2.4 the wave plate symmetry $U \mapsto ZU^TZ$ is equivalent to changing the sign of the X-component of the unitary in any parametrisation. The symmetry $U \mapsto XUX$ is in a sense the opposite of this, as it transforms as

$$\begin{bmatrix}\theta & n_x & n_y & n_z\end{bmatrix} \mapsto \begin{bmatrix}\theta & n_x & -n_y & -n_z\end{bmatrix}. \tag{3.3-42}$$

Such a gadget can be written down directly using a form similar to the reciprocal gadget (3.3-23) and using only the gadgets from Section 3.3.1, for example

$$\mathbf{Z}^A(\gamma)\mathbf{Y}^A(\beta)\mathbf{X}^S(\alpha)\mathbf{Y}^A(\beta)\mathbf{Z}^A(\gamma), \tag{3.3-43}$$

however such a gadget would be inefficient, requiring ten Faraday rotators in total, and with no direct way of reducing wave plates across them. A more efficient approach is given by the gadget

$$\mathbf{G}_{\perp} = \mathbf{X}^A(-\theta)\mathbf{F}_-\mathbf{L}(\alpha,\beta)\mathbf{F}_+\mathbf{X}^A(\theta). \tag{3.3-44}$$

The fact that the gadget is universal for SU(2) can be seen by writing the rotations explicitly:

$$\mathbf{G}_\perp = X(-\theta)Y(-\pi/2)Y(\beta)Z(\alpha)Y(-\beta)Y(\pi/2)X(\theta), \tag{3.3-45}$$

and doing the substitution $\beta' = \pi/2 - \beta$:

$$\mathbf{G}_\perp = X(-\theta)Y(-\alpha')Z(\beta)Y(\alpha')X(\theta). \tag{3.3-46}$$

Since $Z(\beta)Y(\alpha')X(\theta)$ is a set of Tait-Bryan angles, this form of the parametrisation is equivalent to (3.3-27), which was already shown to be universal. It remains to verify the symmetry properties of the gadget:

$$\begin{aligned}
\Theta\big[\mathbf{G}_\perp\big] &= \Theta\big[\mathbf{X}^A(\theta)\big]\Theta\big[\mathbf{F}_+\big]\Theta\big[\mathbf{L}(\alpha,\beta)\big]\Theta\big[\mathbf{F}_-\big]\Theta\big[\mathbf{X}^A(-\theta)\big] \\
&= \mathbf{X}^A(-\theta)\mathbf{F}_-\mathbf{L}(\alpha,-\beta)\mathbf{F}_+\mathbf{X}^A(\theta) \\
&= X(-\theta)Y(-\pi/2)Y(-\beta)Z(\alpha)Y(\beta)Y(\pi/2)X(\theta) \\
&= X(-\theta+\pi)X(-\pi)Y(-\pi/2)Y(-\beta)Z(\alpha)Y(\beta)Y(\pi/2)X(\pi)X(\theta-\pi) \\
&= X(-\theta+\pi)Y(-\pi/2)Z(\pi)Y(-\beta)Z(\alpha)Y(\beta)Z(-\pi)Y(\pi/2)X(\theta-\pi) \\
&= X(\pi)\big[X(-\theta)Y(-\pi/2)Y(\beta)Z(\alpha)Y(-\beta)Y(\pi/2)X(\theta)\big]X(-\pi) \\
&= X(\pi)\mathbf{G}_\perp X(-\pi).
\end{aligned} \tag{3.3-47}$$

3.3.6 Universal gadget

All the gadgets discussed so far exhibit a symmetry that, once known, can be used to relate the unitary transformation in the forwards propagation direction to the one in the backwards direction. There are, however, gadgets that do not exhibit any such symmetries. Consider, for example

$$\mathbf{G}_X(\phi,\theta) = \mathbf{X}^A(\theta)\mathbf{X}^S(\phi) \ = X(\phi+\theta) = X(\alpha). \tag{3.3-48}$$

In the counterpropagating frame the unitary is

$$\Theta\big[\mathbf{G}_X(\phi,\theta)\big] = \mathbf{X}^S(\phi)\mathbf{X}^A(-\theta) = X(\theta-\phi) = X(\beta). \tag{3.3-49}$$

Taking $\theta = \alpha + \beta$ and $\phi = \alpha - \beta$ it's clear that the rotation angles can be chosen independently in the two propagation directions, and the implemented unitaries do not obey any symmetry under counterpropagation. This can be generalized to arbitrary unitaries by a gadget on the following form:

$$\begin{aligned}
\mathbf{G}_\mathrm{U} &= \mathbf{X}^A(\phi_1)\mathbf{X}^S(\phi_3)\mathbf{Y}^A(\theta_1)\mathbf{Y}^S(\theta_2)\mathbf{Z}^A(\delta_1)\mathbf{Z}^S(\delta_2) \\
&= X(\phi_1+\phi_2)Y(\theta_1+\theta_2)Z(\gamma_1+\gamma_2) \\
&= X(\alpha)Y(\beta)Z(\gamma).
\end{aligned} \tag{3.3-50}$$

In the backwards direction the transformation realised by the gadget is

$$\begin{aligned}
\Theta\big[\mathbf{G}_\mathrm{U}\big] &= \mathbf{Z}^S(\delta_2)\mathbf{Z}^A(-\delta_1)\mathbf{Y}^S(\theta_2)\mathbf{Y}^A(-\theta_1)\mathbf{X}^S(\phi_3)\mathbf{X}^A(-\phi_1) \\
&= Z(\gamma_2-\gamma_1)Y(\theta_2-\theta_1)X(\phi_2-\phi_1) \\
&= Z(\gamma')Y(\beta')X(\alpha').
\end{aligned} \tag{3.3-51}$$

Here the six rotation angles $\alpha, \alpha', \beta, \beta', \gamma, \gamma'$ can be chosen independently, and the transformations in the two directions are therefore also independent. This gadget is once again quite costly, requiring three sets of two Faraday rotators. To construct a simpler gadget one might use the same starting idea as (3.3-50), namely that a symmetric as well as an anti-symmetric version of each Pauli rotation should be present. The simplest way to satisfy this condition is to combine the Simon–Mukunda gadget

$$\mathbf{G}_{\mathrm{SM}} = \mathbf{Q}(\alpha)\mathbf{H}(\beta)\mathbf{Q}(\gamma), \tag{3.3-52}$$

with the XUX gadget $\mathbf{G}_{\perp}$ presented in Section 3.3.5:

$$\mathbf{G}_{\mathrm{U}} = \mathbf{G}_{\mathrm{SM}}\mathbf{G}_{\perp}. \tag{3.3-53}$$

For this gadget to be equivalent to (3.3-50) it should satisfy

$$\mathbf{G}_{\mathrm{SM}}\mathbf{G}_{\perp} = U \tag{3.3-54}$$
$$\Theta\big[\mathbf{G}_{\mathrm{SM}}\mathbf{G}_{\perp}\big] = V, \tag{3.3-55}$$

for two arbitrary unitaries $U, V \in \mathrm{SU}(2)$. Let

$$\mathbf{G}_{\perp} = W, \tag{3.3-56}$$
$$\mathbf{G}_{\mathrm{SM}} = UW^{\dagger}, \tag{3.3-57}$$

By construction this satisfies (3.3-54) for any $W \in \mathrm{SU}(2)$. For the second condition (3.3-55) consider

$$\begin{aligned}
\Theta\big[\mathbf{G}_{\mathrm{SM}}\mathbf{G}_{\perp}\big] &= \Theta\big[\mathbf{G}_{\perp}\big]\Theta\big[\mathbf{G}_{\mathrm{SM}}\big] \\
&= X(\pi)WX(-\pi)Z(\pi)\big[UW^{\dagger}\big]^{T}Z(-\pi) \\
&= X(\pi)WX(-\pi)Z(\pi)W^{*}U^{T}Z(-\pi) \\
&= X(\pi)WX(-\pi)Z(\pi)Y(-\pi)WY(\pi)U^{T}Z(-\pi) \\
&= X(\pi)WX(-\pi)X(\pi)WY(\pi)U^{T}Z(-\pi) \\
&= X(\pi)W^{2}Y(\pi)U^{T}Z(-\pi).
\end{aligned} \tag{3.3-58}$$

Here the relations $Y(\pm\pi)UY(\mp\pi) = Y^{*}$ and (3.2-12) were used. For the universality we require

$$\Theta\big[\mathbf{G}_{\mathrm{SM}}\mathbf{G}_{\perp}\big] = X(\pi)W^{2}Y(\pi)U^{T}Z(-\pi) = V, \tag{3.3-59}$$

and this equation is solved by

$$\mathbf{G}_{\perp} = W = \sqrt{X(-\pi)VZ(\pi)U^{*}Y(-\pi)}. \tag{3.3-60}$$

This shows that both (3.3-54) and (3.3-55) can be satisfied simultaneously. In other words, the gadget is capable of implementing two independent unitaries in the two propagation directions.

3.3.7 Faraday mirrors and optical isolators

In this section we will discuss the two well known uses of Faraday rotators using the same framework that was used to analyse the various polarization gadget. The first device is the Faraday mirror [386], which is simply a mirror preceded by a $\mathbf{F}_\pm$ gadget. Faraday mirrors undo any linearly birefringent polarization rotation between some initial point and a mirror, and additionally rotate linearly polarized light by 90°. They are particularly useful in fiber-based Michelson interferometers, since the Faraday mirrors make the interferometer insensitive to any polarization fluctuations induced by the fibers. Viewed as a gadget, one arm of a Michaelson interferometer with a Faraday mirror is

$$\mathbf{G}_{\mathrm{FM}} = \Theta\big[\mathbf{F}_+\mathbf{G}_{\mathrm{SM}}\big]Z(\pi)\mathbf{F}_+\mathbf{G}_{\mathrm{SM}} = Z(\pi)\mathbf{G}_{\mathrm{SM}}^T Z(-\pi)\mathbf{F}_- Z(\pi)\mathbf{F}_+\mathbf{G}_{\mathrm{SM}}, \tag{3.3-61}$$

where the first $Z(\pi)$ rotation transforms the light into the counterpropagating frame. Substituting in the Tait-Bryan parametrisation $\mathbf{G}_{\mathrm{SM}} = Z(\gamma)Y(\beta)X(\alpha)$ gives

$$\begin{aligned}
\mathbf{G}_{\mathrm{FM}} &= Z(\pi)\mathbf{G}_{\mathrm{SM}}^T Z(-\pi)\mathbf{F}_- Z(\pi)\mathbf{F}_+\mathbf{G}_{\mathrm{SM}} \\
&= X(-\alpha)Y(\beta)Z(\gamma)Y(-\pi/2)Z(\pi)Y(\pi/2)Z(\gamma)Y(\beta)X(\alpha) \\
&= X(-\alpha)Y(\beta)Z(\gamma)X(-\pi)Y(-\pi/2)Y(\pi/2)Z(\gamma)Y(\beta)X(\alpha) \\
&= X(-\alpha-\pi)Y(-\beta)Z(-\gamma)Z(\gamma)Y(\beta)X(\alpha) \\
&= X(-\pi).
\end{aligned} \tag{3.3-62}$$

In the co-moving frame the diagonal polarization states are therefore preserved, while the horizontal/vertical as well as the right-/left-handed circular states are swapped. Transforming back into the original frame by taking $Z(-\pi)\mathbf{G}_{\mathrm{FM}} = Z(-\pi)X(-\pi) = Y(\pi)$ shows that the light viewed from the input is rotated by 90°, just as commonly described.

The second device is the optical isolator, which was discussed in the introduction to this chapter. Such a device consists of two polarisers:

$$\mathbf{P}_V = |V\rangle\langle V| = \mathbf{P}_+ \qquad = Y(-\pi)\,|V\rangle\langle V|\,Y(\pi) = |+\rangle\langle +|\,, \tag{3.3-63}$$

that transform the same as wave plates:

$$\Theta\big[\mathbf{P}_V\big] = Z(\pi)\,|V\rangle\langle V|\,Z(-\pi) = |V\rangle\langle V|\,, \tag{3.3-64}$$

$$\Theta\big[\mathbf{P}_+\big] = Z(\pi)\,|+\rangle\langle +|\,Z(-\pi) = |-\rangle\langle -|\,, \tag{3.3-65}$$

together with a Faraday rotator $\mathbf{F}_-$:

$$\mathbf{G}_{\mathrm{OI}} = \mathbf{P}_+\mathbf{F}_-\mathbf{P}_V. \tag{3.3-66}$$

This device fully transmits V-polarized light in the forward direction:

$$\begin{aligned}
\mathbf{G}_{\mathrm{OI}}\,|V\rangle &= |+\rangle\langle +|\,\mathbf{F}_-\,|V\rangle\langle V|\,|V\rangle \\
&= |+\rangle\langle +|\,\mathbf{F}_-\,|V\rangle \\
&= |+\rangle\langle +|\,|+\rangle \\
&= |+\rangle\,,
\end{aligned} \tag{3.3-67}$$

and blocks all light in the backwards direction:

$$\begin{aligned}\Theta\big[\mathbf{G}_{\mathrm{OI}}\big]\,|\Psi\rangle &= \mathbf{P}_V\mathbf{F}_+\mathbf{P}_-\,|\Psi\rangle \\ &= |V\rangle\langle V|\,\mathbf{F}_+\,|-\rangle\langle -|\,|\Psi\rangle \\ &= |V\rangle\langle V|\,\mathbf{F}_+\,|-\rangle \\ &= |V\rangle\langle V|\,|H\rangle \\ &= 0.\end{aligned} \tag{3.3-68}$$

It is often said that the Faraday rotator is a necessary element in the optical isolator because it introduces nonreciprocity [106], however from the above description it is clear that it is the combination of nonreciprocity in both the $\mathbf{P}_+$-polariser and the Faraday rotator that produces the isolation. Other combinations of symmetries and polarisers are possible. For example:

$$\mathbf{G}_{\mathrm{OI}} = \mathbf{P}_H\mathbf{X}^A(\pi/2)\mathbf{P}_L, \tag{3.3-69}$$

would also be an isolator if the circular polariser transforms as $\Theta\big[\mathbf{P}_L\big] = \mathbf{P}_R$. However, a circular polariser obeying time-reversal symmetry transforms as $\Theta\big[\mathbf{P}_L\big] = \mathbf{P}_L$, and a symmetric X-rotation would therefore be required in (3.3-69). In general, an isolator needs to consist of an even number of nonreciprocal components; since in time-reversal symmetric elements only the X-component is nonreciprocal, the Faraday rotator is necessary.

3.4 Gadget wave plate angles

In this section we will briefly discuss how to find the wave plate angles to implement a given unitary using some of the gadgets described in this chapter. The gadgets described in sections 3.2.3 and 3.3.1 have explicit forms that relate the wave plate angles to the rotations they realise, and they do therefore not need to be discussed further.

3.4.1 Transpose and adjoint gadgets

The transpose and adjoint gadgets discussed in sections 3.3.2 and 3.3.3 are both parametrised in terms of Euler angles:

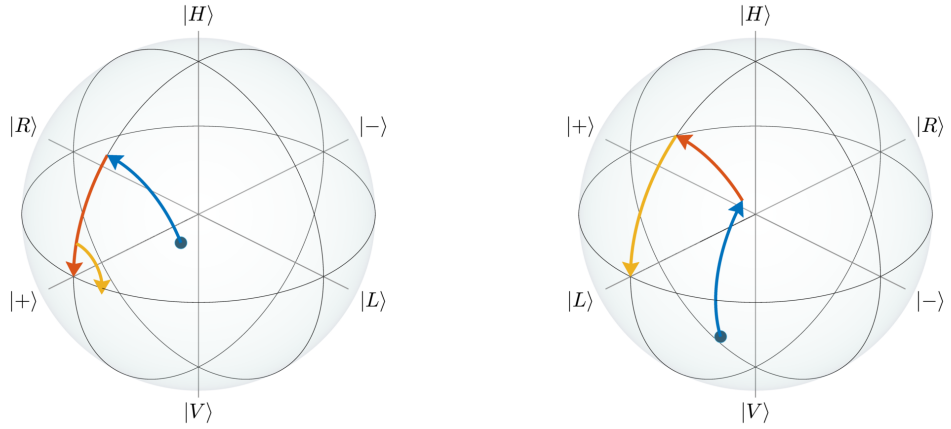
$$U = R_j(\alpha)R_k(\beta)R_j(\gamma). \tag{3.4-1}$$

For these gadgets the angles can be calculated in a way similar to the procedure for the Simon–Mukunda gadget described in 2.3.1. The basic idea, illustrated in Fig. 3.10, is to undo the inverse of the unitary one wishes to implement, by starting from the eigenstates of the chosen rotations mapped by the inverse unitary in question. More concretely, let $|j\rangle$ and $|k\rangle$ be the $+1$ eigenstates of the R_j and R_k rotations, and consider the mapped versions of these states under the inverse unitary $U^\dagger$:

$$U^\dagger\,|j\rangle = |j_1'\rangle\,, \qquad U^\dagger\,|k\rangle = |k_1'\rangle\,. \tag{3.4-2}$$

To map $|j_1'\rangle \mapsto |j_1\rangle$ first choose γ such that $|j_2'\rangle = R_j(\gamma)\,|j_1'\rangle$ lies in the plane orthogonal to k, i.e.

$$\mathrm{Tr}\big[\sigma_k\,|j_2'\rangle\langle j_2'|\big] = 0. \tag{3.4-3}$$



(a) $|j_1'\rangle = U^\dagger |+\rangle$ (b) $|k_1'\rangle = U^\dagger |L\rangle$

Figure 3.10: **$\mathbf{G}_T$ and $\mathbf{G}_A$ gadgets.** An example of a gadget with a Euler-angle decomposition implementing the unitary U for $j = x$ and $k = y$. This is equivalent to implementing the inverse of $U^\dagger$, and the figures (a) and (b) show how this is done for the initial states $U^\dagger |x_+\rangle = |j_1'\rangle$ and $U^\dagger |y_+\rangle = |k_1'\rangle$. The first rotation, $R_x(\gamma)$, takes $|j_1'\rangle$ to the plane perpendicular to y, and the $R_y(\beta)$ operation then rotates it to x_+. Finally, the second rotation about x, $R_x(\alpha)$ takes $R_y(\beta)R_x(\gamma)U^\dagger |x_+\rangle$ to y_+.

As illustrated in Fig. 3.11, this condition is satisfied by

$$\gamma = \operatorname{atan2}\big(\epsilon_{jkl}\operatorname{Tr}\big[\sigma_k |j_1'\rangle\langle j_1'|\big], \operatorname{Tr}\big[\sigma_l |j_1'\rangle\langle j_1'|\big]\big). \tag{3.4-4}$$

Next, the second angle, β, should be chosen such that $|j_2'\rangle \mapsto |j_1\rangle$, or equivalently

$$\operatorname{Tr}\big[\sigma_j |j_3'\rangle\langle j_3'|\big] = 1, \tag{3.4-5}$$

where $R_k(\beta) |j_2'\rangle = |j_3'\rangle$. This can be done by taking

$$\beta = \operatorname{atan2}\big(\epsilon_{jkl}\operatorname{Tr}\big[\sigma_l |j_2'\rangle\langle j_2'|\big], \operatorname{Tr}\big[\sigma_j |j_2'\rangle\langle j_2'|\big]\big). \tag{3.4-6}$$

At this point all that remains is to map $R_k(\beta)R_j(\gamma) |k_1'\rangle = |k_2'\rangle \mapsto |k\rangle$. Let

$$R_j(\alpha)R_k(\beta)R_j(\gamma) |k_1'\rangle = |k_3'\rangle . \tag{3.4-7}$$

Then this final condition

$$\operatorname{Tr}\big[\sigma_k |k_3'\rangle\langle k_3'|\big] = 1, \tag{3.4-8}$$

is satisfied by

$$\alpha = \operatorname{atan2}\big(\epsilon_{jkl}\operatorname{Tr}\big[\sigma_l |k_2'\rangle\langle k_2'|\big], \operatorname{Tr}\big[\sigma_k |k_2'\rangle\langle k_2'|\big]\big) \tag{3.4-9}$$

This ensures that

$$R_j(\alpha)R_k(\beta)R_j(\gamma) = \pm U, \tag{3.4-10}$$

and if needed the global phase can be corrected by adding 2π to any of the rotations.

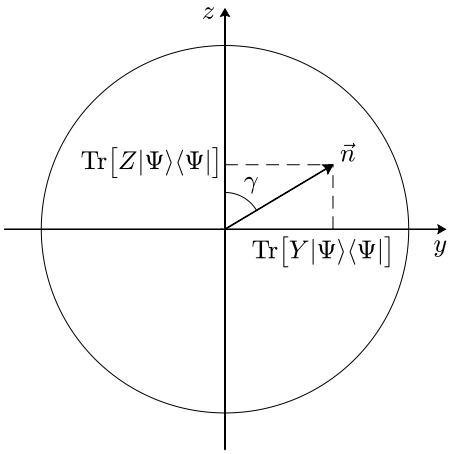


Figure 3.11: **Bloch sphere rotation angle.** The rotation angles for the $\mathbf{G}_T$ and $\mathbf{G}_A$ gadgets can be found geometrically. In this example the rotation axes are $j = x$ and $k = y$ The first rotation about the axis x should bring the state $|\Psi\rangle$ to the plane spanned by the axes x and z. The figure shows the $y - z$ plane of the Bloch sphere, and the vector $\vec{n}$ is the Bloch vector of the state $|\Psi\rangle$. The desired rotation angle γ is given by $\tan\gamma = n_z/n_y$, where $n_z = \mathrm{Tr}[Y\,|\Psi\rangle\langle\Psi|]$ and $n_y = \mathrm{Tr}[Y\,|\Psi\rangle\langle\Psi|]$. The sign-correcting factor ϵ_{jkl} that appears in the formulae in the main text accounts for the handedness of the coordinate system defined by the rotation order.

3.4.2 Reciprocal and universal gadgets

The procedure for the reciprocal gadget is slightly different. Start with the parametrisation

$$X(\pi)\mathbf{G}_{\mathrm{R}} = Y(-\gamma)Z(-\beta)X(\alpha)Z(\beta)Y(\gamma). \tag{3.4-11}$$

Given a unitary $U = \mathbf{G}_{\mathrm{R}}$ that should be implemented, define a new unitary $V = X(\pi)U$ and let $V\,|v\rangle = e^{-i\lambda/2}\,|v\rangle$. The rotation angles γ and β should then be chosen such that the eigenstate of V gets mapped to $|+\rangle$:

$$Z(\beta)Y(\gamma)\,|v\rangle = e^{i\varphi}\,|+\rangle\,, \tag{3.4-12}$$

where φ is a irrelevant phase. Once this has been done, the $R_x(\alpha)$ rotation can be used to apply the correct eigenphase, after which the last two rotations map $|+\rangle$ back to $|v\rangle$. This procedure is illustrated in Fig. 3.12. The condition above can be satisfied by first taking

$$\gamma = \mathrm{atan2}\big(\mathrm{Tr}\big[Z\,|v\rangle\langle v|\big], \mathrm{Tr}\big[X\,|v\rangle\langle v|\big]\big). \tag{3.4-13}$$

Then let $Y(\gamma)\,|v\rangle = |v'\rangle$ and take

$$\beta = -\mathrm{atan2}\big(\mathrm{Tr}\big[Y\,|v'\rangle\langle v'|\big], \mathrm{Tr}\big[X\,|v'\rangle\langle v'|\big]\big). \tag{3.4-14}$$

The angles γ and β are illustrated in Fig. 3.13. Setting $\alpha = \lambda$ ensures

$$\begin{aligned} & X(\pi)\mathbf{G}_{\mathrm{R}} = V = X(\pi)U \\ \iff \qquad & \mathbf{G}_{\mathrm{R}} = U. \end{aligned} \tag{3.4-15}$$

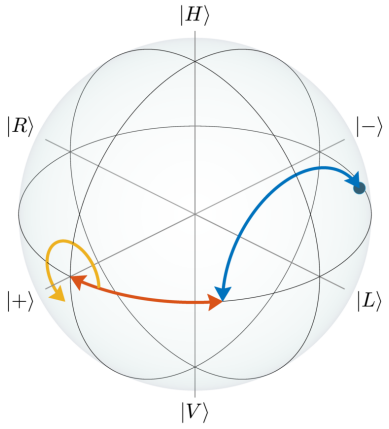


Figure 3.12: **Principle of the $\mathbf{G}_\mathrm{R}$ gadget.** The reciprocal gadget can be decomposed as $X(\pi)\mathbf{G}_\mathrm{R} = Y(-\gamma)Z(-\beta)X(\alpha)Z(\beta)Y(\gamma)$. This decomposition can be used to realise any unitary by using the first two rotations to take the eigenstate of the desired transformation to $|+\rangle$. The $X(\alpha)$ rotation then applies the eigenphase of the unitary, after which the inverse of the first two rotations take $|+\rangle$ back to the initial state, but with the added phase.

To find the wave plate angles, recall the form of the full gadget:

$$\begin{aligned}\mathbf{G}_\mathrm{R} &= \mathbf{Q}(-\phi)\mathbf{H}(-\theta)\mathbf{X}^S(\psi)\mathbf{H}(\theta)\mathbf{Q}(\phi)\\ &= Y(\gamma)Z(\beta)X(\pi/2)X(\psi)X(-\pi/2)Z(\beta)Y(\gamma).\end{aligned} \tag{3.4-16}$$

The angle ψ is given by $\psi = \alpha - \pi$, and the two wave plates on the right side should implement

$$\begin{aligned}W &= X(-\pi/2)Z(\beta)Y(\gamma)\\ &= \mathbf{H}(\theta)\mathbf{Q}(\phi)\end{aligned} \tag{3.4-17}$$

Let $|L'\rangle = W^\dagger |L\rangle = Y(-\gamma)Z(-\beta)X(\pi/2)$. Note that, as illustrated in Fig. 3.14, $|L'\rangle$ always lies in the linear plane. The quarter-wave plate can therefore be used to map it back to a circular polarization. The QWP angle ϕ should be chosen as

$$\phi = \frac{1}{2}\mathrm{atan2}\big(\mathrm{Tr}\big[X\,|L'\rangle\langle L'|\big], \mathrm{Tr}\big[Z\,|L'\rangle\langle L'|\big],\big) + \frac{\pi}{4}. \tag{3.4-18}$$

This choice of ϕ maps $|L'\rangle \mapsto |R\rangle$, and ensures that $|H'\rangle = \mathbf{Q}(\phi)W^\dagger |H\rangle$ lies in the linear plane. The HWP angle θ should then be picked so that $|H'\rangle \mapsto |H\rangle$. This is sufficient, as every choice of angle maps $|R\rangle \mapsto |L\rangle$. The correct angle θ can be found as

$$\theta = \frac{1}{4}\mathrm{atan2}\big(\mathrm{Tr}\big[X\,|H'\rangle\langle H'|\big], \mathrm{Tr}\big[Z\,|H'\rangle\langle H'|\big]\big). \tag{3.4-19}$$

The above expressions for the wave plate angles are explained in greater detail in Appendix A. The conditions above lead to an ambiguity in the global phase, however this turns out to not matter since the unitary is realised on both sides of the gadget. The reduction from three to two wave plates on each side of the gadget:

$$\begin{aligned}\mathbf{G}_\mathrm{R} &= \mathbf{Q}(-\phi)\mathbf{H}(-\theta)\mathbf{H}(\pi/8)\mathbf{Z}^A(\phi)\mathbf{H}(\pi/8)\mathbf{H}(\theta)\mathbf{Q}(\phi)\\ &= \mathbf{Q}(\phi_1)\mathbf{H}(\theta_1)\mathbf{Z}^A(\phi)\mathbf{H}(\theta_2)\mathbf{Q}(\phi_2),\end{aligned} \tag{3.4-20}$$

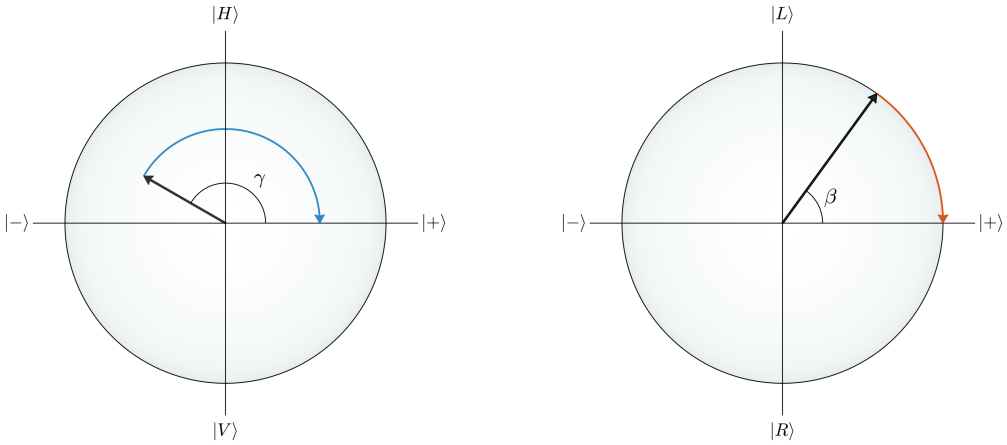


Figure 3.13: **$\mathbf{G}_{\mathrm{R}}$ gadget angles.** The first two rotations $Y(\gamma)$ and $Z(\beta)$ of the reciprocal gadget should take an initial state $|\Psi\rangle$ first to the $x-y$ plane and then to $|+\rangle$. The two figures above show these rotations in the $x-z$ and $x-y$ planes, respectively, and correspond to the transformation shown in Fig. 3.12. The left figure illustrates the need for the atan2 function, since $\mathrm{atan}(\mathrm{Tr}[Z\,|\Psi\rangle\langle\Psi|]/\mathrm{Tr}[X\,|\Psi\rangle\langle\Psi|]) = \pi - \gamma$, rather than the correct angle γ. In the right figure, the positive z-axis is pointing out of the figure, and in order to rotate the state counter-clockwise the negative angle should be used, explaining the sign in (3.4-14).

can be done using (3.2-13) and (3.2-14), giving

$$\theta_1 = \frac{\pi}{2} - \theta \tag{3.4-21}$$

$$\phi_1 = \phi - \theta + \frac{\pi}{8} - \frac{\pi}{2} \tag{3.4-22}$$

$$\theta_2 = \frac{\pi}{2} + \theta \tag{3.4-23}$$

$$\phi_2 = \theta - \phi + \frac{\pi}{8} - \frac{\pi}{2}. \tag{3.4-24}$$

For the universal gadget $\mathbf{G}_{\mathrm{U}} = \mathbf{G}_{\mathrm{SM}}\mathbf{G}_{\perp}$ the Simon–Mukunda part can be calculated using the steps in Section 2.3.1 and the XUX gadget has the same form as (3.4-11) and the same steps as the ones above can therefore be used.

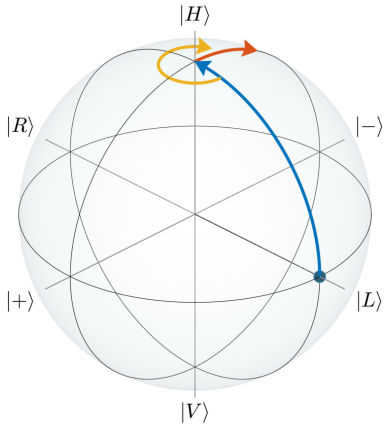


Figure 3.14: **Action of $W^\dagger$ on $|L\rangle$.** The transformation $Y(-\gamma)Z(-\beta)X(\pi/2)$ always takes $|L\rangle$ to somewhere in the linear plane, as illustrated above. This means that a QWP can always be used to map the resulting state back to a circular polarization.

3.5 Applications

In this section a few potential applications of the gadgets described throughout the chapter will discussed. The theme of these is taking something which is already possible in a non-common-path interferometer, such as a Mach-Zehnder interferometer, and translating it to a common-path interferometer that has passive phase stability. The primary example of this is a quantum switch built using a Sagnac geometry. An experimental realisation of such a device is the subject of Publication 4, and is elaborated on in Chapter 7.

3.5.1 Variable partially-polarizing beamsplitter

Partially-polarizing beamsplitters (PPBSes) show up in many places in photonic quantum information, such as in the C-Phase gate shown in Fig. 2.16, and as shown in Fig. 2.15 they can also be used to implement POVMs on polarization qubits, without resorting to inconvenient beam-displacers, or a polarization-path conversion scheme. PPBSes, however, often suffer from low tolerances in their splitting ratios, and being bulk optics components they are not tunable, which prohibits their use for implementing for example dynamically reconfigurable POVMs [387]. One way to implement a PPBS is using a Mach-Zehnder interferometer with one tunable R_z operation in each arm. This is pictured in Fig. 3.15. The first beamsplitter transforms a input state $|\Psi\rangle = (\alpha\hat{a}_H^\dagger + \beta\hat{a}_V^\dagger)\,|0\rangle$ as

$$|\Psi\rangle \mapsto \alpha\frac{\hat{a}_H^\dagger + \hat{b}_H^\dagger}{\sqrt{2}}\,|0\rangle + \beta\frac{\hat{a}_V^\dagger + \hat{b}_V^\dagger}{\sqrt{2}}\,|0\rangle \tag{3.5-1}$$

The R_z operations then introduce phases ϕ_a and ϕ_b between the polarization components,

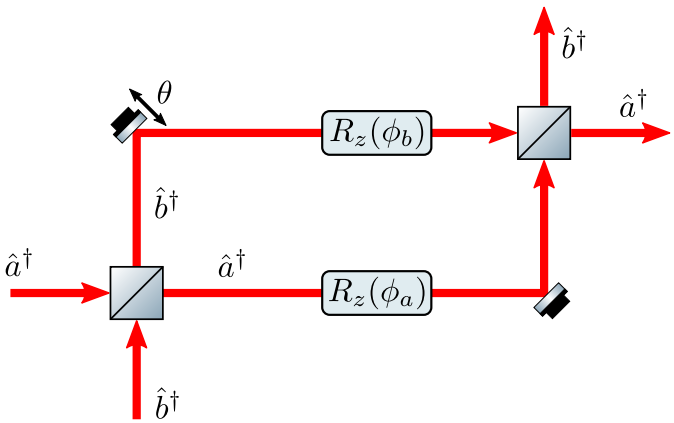


Figure 3.15: **Variable PPBS.** A regular Mach-Zehnder interferometer can be used to construct a variable partially-polarizing beamsplitter. By including polarization rotations about the z-axis in one or both of the MZI arms. This changes the interference condition for the two polarization components, however this by itself would not lead to a polarization dependent splitting ratio since the output probabilities are symmetric functions of the interferometer phase. By additionally controlling the phase θ originating in the path-length difference the interference condition can be tuned independently for each polarization component. The downside of this is that the path-length difference must be actively stabilized to keep the phase θ constant.

in addition to the interferometric phase θ between the two arms of the interferometer:

$$\begin{aligned} &\alpha\frac{\hat{a}_H^\dagger+\hat{b}_H^\dagger}{\sqrt{2}}\left|0\right\rangle+\beta\frac{\hat{a}_V^\dagger+\hat{b}_V^\dagger}{\sqrt{2}}\left|0\right\rangle \\ \mapsto\ &\frac{1}{\sqrt{2}}\left(e^{-i(\phi_a+\theta)/2}\hat{a}_H^\dagger+e^{-i(\phi_b-\theta)/2}\hat{b}_H^\dagger+e^{-i(-\phi_a+\theta)/2}\hat{a}_V^\dagger+e^{i(\phi_b+\theta)/2}\hat{b}_V^\dagger\right)\left|0\right\rangle. \end{aligned} \tag{3.5-2}$$

The total phase difference $\Delta\varphi$ for the two polarization components at the second beamsplitter is therefore

$$\Delta\varphi_H=\frac{1}{2}\left(\theta+\phi_a+\theta-\phi_b\right)=2\theta+\Delta\phi \tag{3.5-3}$$

$$\Delta\varphi_V=-\frac{1}{2}\left(\theta-\phi_a+\theta+\phi_b\right)=-2\theta+\Delta\phi, \tag{3.5-4}$$

where $\Delta\phi=\phi_a-\phi_b$. It's clear that the interference conditions for the two polarization components can be set independently by controlling the phase θ through the physical path length difference between the two arms, together with the birefringent phase difference $\Delta\phi$. At the same time that means that the path length difference between the two arms must be kept constant for a given pair of reflectivities for the PPBS.

If one wishes to make an interferometric PPBS in a Sagnac interferometer instead, where the phase θ does not fluctuate, one immediately runs into the problem this phase cannot be tuned. There is, however, a different way to make a tunable PPBS in a Sagnac configuration. Such a device makes use of a polarizing beamsplitter instead of a regular balanced beamsplitter, as well as a pair of symmetric and anti-symmetric $\mathbf{Y}(\phi)$ gadgets. The design is of this device is illustrated in Fig. 3.16.

The basic idea is that states propagating in the clockwise direction see the gadget

$$\mathbf{Y}^S(\phi_1)\mathbf{Y}^A(\phi_2) = Y(\phi_1 + \phi_2) = Y(\alpha), \tag{3.5-5}$$

and the light propagating in the counter-clockwise direction sees

$$\Theta\big[\mathbf{Y}^S(\phi_1)\mathbf{Y}^A(\phi_2)\big] = Y(\phi_1 - \phi_2) = Y(\beta), \tag{3.5-6}$$

and the rotation of the linearly polarised light can be independently tuned in the two propagation directions. Consider H-polarised light entering from the left port of the device in Fig. 3.16. This light is transmitted by the PBS, travels in a counter-clockwise fashion and is rotated by an angle α:

$$\hat{a}_H^\dagger \mapsto \cos\Big(\frac{\alpha}{2}\Big)\hat{a}_H^\dagger + \sin\Big(\frac{\alpha}{2}\Big)\hat{b}_V^\dagger. \tag{3.5-7}$$

It is then split on the PBS with a reflection coefficient of $\mathcal{R} = \sin(\alpha/2)$. Let the PBS transformation inside the Sagnac be

$$\begin{aligned} \hat{a}_H^\dagger &\mapsto \hat{c}_H^\dagger, \qquad \hat{a}_V^\dagger \mapsto \hat{d}_V^\dagger \\ \hat{b}_H^\dagger &\mapsto \hat{d}_H^\dagger, \qquad \hat{b}_V^\dagger \mapsto \hat{c}_V^\dagger, \end{aligned}, \tag{3.5-8}$$

where c and d are the counter-propagating modes in the bottom and left ports, respectively. The state after the PBS is then

$$\cos\Big(\frac{\alpha}{2}\Big)\hat{c}_H^\dagger + \sin\Big(\frac{\alpha}{2}\Big)\hat{d}_V^\dagger \tag{3.5-9}$$

Since the reflected light is V-polarized, it is necessary to swap the H and V components in the backwards propagating top mode d with a $Y(-\pi)$ rotation, giving the total transformation

$$\hat{a}_H^\dagger \mapsto \cos\Big(\frac{\alpha}{2}\Big)\hat{c}_H^\dagger + \sin\Big(\frac{\alpha}{2}\Big)\hat{d}_H^\dagger, \tag{3.5-10}$$

In contrast, V-polarized light entering the left port is reflected by the PBS, travels clockwise through the interferometer, and undergoes the rotation $Y(\beta)$:

$$\hat{a}_V^\dagger \mapsto \hat{b}_V^\dagger \mapsto -\cos\Big(\frac{\beta}{2}\Big)\hat{a}_H^\dagger + \sin\Big(\frac{\beta}{2}\Big)\hat{b}_V^\dagger. \tag{3.5-11}$$

The total transformation after the PBS and $Y(-\pi)$ rotation is

$$\hat{a}_V^\dagger \mapsto \cos\Big(\frac{\beta}{2}\Big)\hat{c}_V^\dagger + \sin\Big(\frac{\beta}{2}\Big)\hat{d}_V^\dagger. \tag{3.5-12}$$

The splitting ratios of the PPBS for H and V can therefore be set by choosing α and β. In order to use the bottom port of the PBS as an input as well, it is necessary to perform a $Y(\pi)$ operation at the input. An H-polarised input state in mode b then

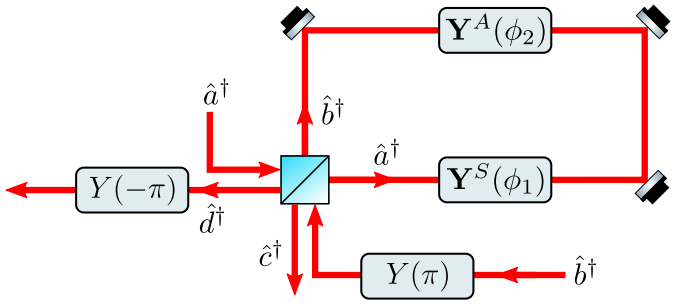


Figure 3.16: **Sagnac variable PPBS.** Using a pair of symmetric and anti-symmetric $\mathbf{Y}$ gadgets it is possible to create partially-polarizing beamsplitter in a Sagnac geometry. The PBS splits a polarization state such that the H-component input in mode a propagates in a counter-clockwise fashion, and the V-component propagates clockwise. In the counter-clockwise direction the Y-rotations add, giving a total rotation of linear polarization by an angle $\alpha = \phi_1 + \phi_2$, whereas in the clockwise propagation direction the angles subtract, giving a total rotation of $\beta = \phi_1 - \phi_2$. This leads to two different splitting ratios for the H and V polarizations when they impinge on the PBS a second time. Two additional Y-rotations in the input/output modes c and d, respectively, ensure that the device works in an all-to-all mode configuration.

becomes V-polarized, is reflected by the PBS and travels in a counter-clockwise fashion just as an H-polarized input state in mode a:

$$\begin{aligned} U_{\text{PBS}}Y(\pi): \quad & \hat{b}_H^\dagger \mapsto \hat{a}_V^\dagger \\ Y(\alpha): \quad & \mapsto \cos\left(\frac{\alpha}{2}\right)\hat{a}_V^\dagger - \sin\left(\frac{\alpha}{2}\right)\hat{a}_H^\dagger \\ U_{\text{PBS}}: \quad & \mapsto \cos\left(\frac{\alpha}{2}\right)\hat{d}_V^\dagger - \sin\left(\frac{\alpha}{2}\right)\hat{c}_H^\dagger \\ I_c \otimes Y(-\pi)_d: \quad & \mapsto \cos\left(\frac{\alpha}{2}\right)\hat{d}_H^\dagger - \sin\left(\frac{\alpha}{2}\right)\hat{c}_H^\dagger, \end{aligned} \tag{3.5-13}$$

and similarly for V-polarized light incident in the bottom mode b:

$$\begin{aligned} U_{\text{PBS}}Y(\pi): \quad & \hat{b}_V^\dagger \mapsto -\hat{b}_H^\dagger \\ Y(\beta): \quad & \mapsto -\cos\left(\frac{\beta}{2}\right)\hat{b}_H^\dagger - \sin\left(\frac{\beta}{2}\right)\hat{b}_V^\dagger \\ U_{\text{PBS}}: \quad & \mapsto -\cos\left(\frac{\beta}{2}\right)\hat{c}_H^\dagger - \sin\left(\frac{\beta}{2}\right)\hat{d}_V^\dagger \\ I_c \otimes Y(-\pi)_d: \quad & \mapsto -\cos\left(\frac{\beta}{2}\right)\hat{c}_V^\dagger + \sin\left(\frac{\beta}{2}\right)\hat{d}_V^\dagger. \end{aligned} \tag{3.5-14}$$

The total transformation of the PPBS device is therefore:

$$\begin{bmatrix} \cos\left(\frac{\alpha}{2}\right) & -\sin\left(\frac{\alpha}{2}\right) & 0 & 0 \\ \sin\left(\frac{\alpha}{2}\right) & \cos\left(\frac{\alpha}{2}\right) & 0 & 0 \\ 0 & 0 & \cos\left(\frac{\beta}{2}\right) & -\sin\left(\frac{\beta}{2}\right) \\ 0 & 0 & \sin\left(\frac{\beta}{2}\right) & \cos\left(\frac{\beta}{2}\right) \end{bmatrix}, \tag{3.5-15}$$

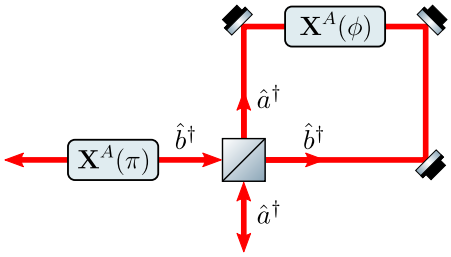


Figure 3.17: **Variable beamsplitter.** Using the polarization degree of freedom as an intermediary, a regular Sagnac interferometer can be made to act as a variable beamsplitter, while still preserving the polarization state. Viewed in the $\{H, V\}$ basis the anti-symmetric $\mathbf{X}^A(\phi)$ gadget simply changes the interference condition the same way for both polarizations. Consequently polarization states that are superpositions of the horizontal and vertical polarizations have the same interference condition as well, however in one output port the state becomes rotated. Since this rotation does not depend on the angle ϕ it can be corrected with a static $X(-\pi)$ rotation. For states input in mode b this correction would have to be applied in the opposite port, however this can instead be accounted for with an additional $X(\pi)$ rotation before the beamsplitter. Both of these corrective rotations can be realised with a single $\mathbf{X}^A(\pi)$ gadget.

written in the bases:

$$\begin{aligned} \mathcal{H}_{in}: &\quad \{\hat{a}_H^\dagger, \hat{b}_H^\dagger, \hat{a}_V^\dagger, \hat{b}_V^\dagger\} \\ \mathcal{H}_{out}: &\quad \{\hat{c}_H^\dagger, \hat{d}_H^\dagger, \hat{c}_V^\dagger, \hat{d}_V^\dagger\}. \end{aligned} \tag{3.5-16}$$

3.5.2 Variable beamsplitter

As discussed in Section 2.5.1 a basic Mach-Zehnder interferometer works as a variable beamsplitter, and in analogy with the variable PPBS discussed in the previous section, the lack of control over the interferometric phase in a Sagnac interferometer translates into a fixed splitting ratio if one tries to use such a common-path interferometer as a beamsplitter. By using the polarization degree of freedom as an intermediary, however, one can control the splitting ratio. Such a device is pictured in Fig. 3.17.

Given a polarization gadget $\mathbf{G}$ inside a Sagnac interferometer, as well an input state $|\Psi\rangle_a$, the output state is of the interferometer is

$$\frac{1}{2}\big(\mathbf{G} + \Theta[\mathbf{G}]\big)\,|\Psi\rangle_a + \frac{1}{2}\big(\mathbf{G} - \Theta[\mathbf{G}]\big)\,|\Psi\rangle_b\,, \tag{3.5-17}$$

where a and b are the two input/output modes of the Sagnac. If the gadget $\mathbf{G}$ is anti-symmetric, then the relative probability amplitude in the two output modes can be controlled. The simplest choice for achieving this is the $\mathbf{X}^A(\phi)$ gadget, since it only requires wave plates. This gives the output state

$$\begin{aligned} &\frac{1}{2}\big(\mathbf{X}^A(\phi) + \mathbf{X}^A(-\phi)\big)\,|\Psi\rangle_a + \frac{1}{2}\big(\mathbf{X}^A(\phi) - \mathbf{X}^A(+\phi)\big)\,|\Psi\rangle_b \\ &\qquad = \cos\Big(\frac{\phi}{2}\Big) I\,|\Psi\rangle_a - i\sin\Big(\frac{\phi}{2}\Big) X\,|\Psi\rangle_b \\ &\qquad = \cos\Big(\frac{\phi}{2}\Big) I\,|\Psi\rangle_a + \sin\Big(\frac{\phi}{2}\Big) X(\pi)\,|\Psi\rangle_b\,. \end{aligned} \tag{3.5-18}$$

Aside from the X operation in the second output mode, this is a tunable beamsplitter with reflection coefficient $\mathcal{R} = \sin(\phi/2)$. Since the $X(\pi)$ rotation in mode b does not depend on the rotation angle ϕ inside the Sagnac, is can be undone with a fixed $X(-\pi)$ operation, thereby making the beamsplitter polarization preserving. The total transformation for an input state in mode a with polarisation P is

$$\begin{aligned}
U_{\mathrm{BS}}: \quad \hat{a}_P^\dagger &\mapsto \frac{\hat{a}_P^\dagger + \hat{b}_P^\dagger}{\sqrt{2}} \\
\mathbf{X}^A(\phi)_a \otimes \Theta\big[\mathbf{X}^A(\phi)_b\big]: \quad &\mapsto \frac{1}{\sqrt{2}}\big(X(\phi)\hat{a}_P^\dagger + X(-\phi)\hat{b}_P^\dagger\big) \\
U_{\mathrm{BS}}: \quad &\mapsto \frac{1}{2}\Big(\big[X(\phi) + X(-\phi)\big]\hat{a}_P^\dagger + \big[X(\phi) - X(-\phi)\big]\hat{b}_P^\dagger\Big) \\
&= \cos\Big(\frac{\phi}{2}\Big)\hat{a}_P^\dagger + \sin\Big(\frac{\phi}{2}\Big)X(\pi)\hat{b}_P^\dagger \\
I_a \otimes X(-\pi)_b: \quad &\mapsto \cos\Big(\frac{\phi}{2}\Big)\hat{a}_P^\dagger + \sin\Big(\frac{\phi}{2}\Big)X(0)\hat{b}_P^\dagger \\
&= \cos\Big(\frac{\phi}{2}\Big)\hat{a}_P^\dagger + \sin\Big(\frac{\phi}{2}\Big)\hat{b}_P^\dagger,
\end{aligned} \tag{3.5-19}$$

Similar to the case of the variable PPBS, in order to make use of the second input port it is necessary to first flip the input polarization with an $X(\pi)$ rotation:

$$\begin{aligned}
X(\pi): \quad \hat{b}_P^\dagger &\mapsto X(\pi)\hat{b}_P^\dagger \\
U_{\mathrm{BS}}: \quad &\mapsto X(\pi)\frac{\hat{a}_P^\dagger - \hat{b}_P^\dagger}{\sqrt{2}} \\
\mathbf{X}^A(\phi)_a \otimes \Theta\big[\mathbf{X}^A(\phi)_b\big]: \quad &\mapsto \frac{1}{\sqrt{2}}\big(X(\phi)X(\pi)\hat{a}_P^\dagger - X(-\phi)X(\pi)\hat{b}_P^\dagger\big) \\
U_{\mathrm{BS}}: \quad &\mapsto \frac{1}{2}\big(\big[X(\phi) - X(-\phi)\big]X(\pi)\hat{a}_P^\dagger \\
&+ \quad \big[X(\phi) + X(-\phi)\big]X(\pi)\hat{b}_P^\dagger\big) \\
&= \sin\Big(\frac{\phi}{2}\Big)X(2\pi)\hat{a}_P^\dagger + \cos\Big(\frac{\phi}{2}\Big)X(\pi)\hat{b}_P^\dagger \\
I_a \otimes X(-\pi)_b: \quad &\mapsto -\sin\Big(\frac{\phi}{2}\Big)\hat{a}_P^\dagger + \cos\Big(\frac{\phi}{2}\Big)X(0)\hat{b}_P^\dagger \\
&= -\sin\Big(\frac{\phi}{2}\Big)\hat{a}_P^\dagger + \cos\Big(\frac{\phi}{2}\Big)\hat{b}_P^\dagger.
\end{aligned} \tag{3.5-20}$$

The transformation realised by device is therefore

$$\begin{bmatrix} \cos\Big(\frac{\phi}{2}\Big) & -\sin\Big(\frac{\phi}{2}\Big) \\ \sin\Big(\frac{\phi}{2}\Big) & \cos\Big(\frac{\phi}{2}\Big) \end{bmatrix}, \tag{3.5-21}$$

which shows that the splitting ratio can be controlled using the gadget phase. This is transformation is equivalent to a Sagnac interferometer with a relative phase difference

of θ between the two arms, which means that a nonreciprocal gadget can be used to set the working point of the interferometers, which is often important for uses in sensing, such as optical gyroscopes. In ring-laser gyroscopes a phase bias is typically applied using active phase modulators [388], whereas with access to the polarization degree of freedom this could be accomplished entirely passively.

4

Experimental methods for "Trace-free counterfactual communication with a nanophotonic processor"

This chapter will discuss some of the experimental methods behind Publication 1 'Trace-Free counterfactual communication with a nanophotonic processor'. As indicated by the title of the final publication, the goal of this project was to demonstrate a counterfactual communication protocol where the photon transmitted by Alice does not leave a weak trace in Bob's laboratory. The motivation for this is that the presence of a weak trace of a photon in a certain region of space can be taken to mean that the photon too was present in that region. In the context of counterfactual communication, the presence of such a weak trace can call the counterfactuality of the process into question. In Section 1.3.4 the first protocol for counterfactual communication, proposed by Salih et al. [66], was analysed using the so-called two-state vector formalism, and it was shown visually that the protocol does not eliminate the weak trace of photons sent by Alice in Bob's laboratory. To address this issue, a modified protocol was proposed by Arvidsson-Shukur and Barnes [85], and it was this protocol that was the focus of Publication 1.

4.1 The Arvidsson-Shukur and Barnes protocol

The counterfactual communication protocol proposed by Arvidsson-Shukur and Barnes differs slightly from the original Salih protocol in that photons are allowed to travel from Alice to Bob as long as any photons entering Bob's laboratory do not travel back to Alice. For this reason Salih-type protocols are sometimes referred to as type-I protocols, while the ones of the latter type are called type-II [71]. The type-II protocol discussed here is actually simpler than the Salih protocol, since it does not contain any nested Mach-Zehnder interferometers. Instead it consists of $N-1$ concatenated MZIs, with a total of N beamsplitters, as shown in Fig. 4.1.a. Bob can transmit a 0- or 1-bit to Alice by either blocking the photon path through his laboratory or not. The reflectivity of the beamsplitters is not $50:50$, but is instead chosen to be

$$R = \cos^2\Big(\frac{\pi}{2N}\Big). \tag{4.1-1}$$

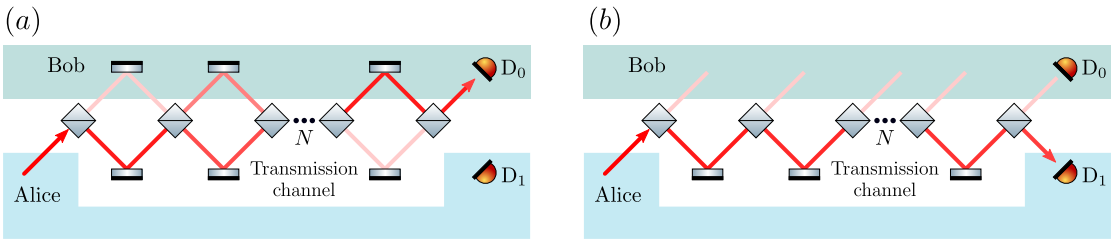


Figure 4.1: **Arvidsson-Shukur & Barnes CFC protocol.** The trace-free counterfactual communication protocol proposed in Ref. [85] allows for two parties, Alice and Bob, to exchange a classical bit. The two shaded areas indicate which parts of the setup belong to the laboratories of the two parties, while the unshaded area is considered a transmission channel belonging to neither party. (a) If Bob wants to send a 0-bit to Alice he leaves the mirrors in his lab in place. This causes the photon amplitude to be gradually transferred from the transmission channel to Bob's laboratory, and at the end of the interferometer chain it is detected by Bob with unity probability. Alice records the non-detection of a photon as a 0-bit. (b) When Bob wants to send a 1-bit he blocks or removes the mirrors in his lab. This prevents the photon amplitude leaking into his lab from returning to the transmission channel, and the majority of the wavefunction will simply reflect off the high reflectivity beamsplitters and return to Alice, who records a 1-bit when she successfully detects the photon she injected into the transmission channel. In the limit of infinitely many beamsplitters this occurs with unity probability. Note that in neither (a) nor (b) does the photon ever travel from Bob's laboratory back to Alice, and this is why the protocol is counterfactual.

The corresponding beamsplitter transformation is

$$U = \begin{bmatrix} \cos\left(\frac{\pi}{2N}\right) & i\sin\left(\frac{\pi}{2N}\right) \\ i\sin\left(\frac{\pi}{2N}\right) & \cos\left(\frac{\pi}{2N}\right) \end{bmatrix} = \exp\left[i\frac{\pi}{2N}X\right], \tag{4.1-2}$$

implicitly written such that reflection is described by the identity operator, since this corresponds to the photon staying in the transmission channel (see Fig. 4.1.a). If Bob wishes to transmit a 0-bit to Alice he leaves the mirrors in his laboratory unblocked. The photon is propagated by the transformation U a total of N times:

$$\begin{aligned} U^N &= \exp\left[iN\frac{\pi}{2N}X\right] \\ &= \exp\left[i\frac{\pi}{2}X\right] \\ &= iX. \end{aligned} \tag{4.1-3}$$

This is a swap operation, meaning that the photon transits from the transmission channel into Bob's laboratory with unity probability. In the communication protocol this corresponds to the transmission of a 0-bit, which on Alice's side is registered as the non-detection of a photon injected into the transmission channel. When Bob wants to transmit a 1-bit to Alice he blocks or removes the mirrors in his laboratory, as shown in Fig. 4.1.b. Alice records a 1-bit when she successfully detects the photon she injected into the transmission channel, and the probability P_1 for the photon to return to Alice is equal to the probability that the photon successfully reflects off all the N beamsplitters:

$$P_1 = \cos^N\left(\frac{\pi}{2N}\right). \tag{4.1-4}$$

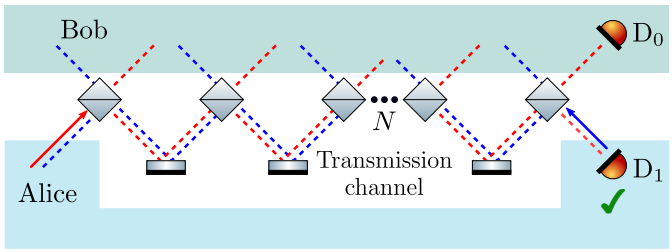


Figure 4.2: **Weak trace in the Arvidsson-Shukur & Barnes protocol.** As discussed in Section 1.3.4, the weak trace of a photon is non-zero only in regions where the forwards and backwards propagating trajectories overlap. The figure shows the case when Bob successfully transmits a 1-bit by removing the mirrors in his lab. In this scenario, the forwards and backwards propagating states, indicated by the red and blue dashed lines respectively, have zero overlap in Bob's lab, and the photon does therefore not leave a weak trace there. The diagram for the 0-bit is not shown here, since in that configuration the photon never travels back to Alice.

In the limit of infinitely many beamsplitters the 1-bit success probability P_1 tends to unity. This can be seen by evaluating the limit

$$\lim_{N\to\infty} \cos^{2N}\Big(\frac{\pi}{2N}\Big), \tag{4.1-5}$$

which can be done in several ways. One straightforward way is to take the natural log of the expression above:

$$\begin{aligned}\lim_{N\to\infty} \ln P_1 &= \lim_{N\to\infty} 2N \ln\Big[\cos\Big(\frac{\pi}{2N}\Big)\Big] \\ &= \lim_{N\to\infty} \frac{\ln\Big[1-\Big(1-\cos\Big(\frac{\pi}{2N}\Big)\Big)\Big]}{1-\cos\Big(\frac{\pi}{2N}\Big)} \frac{1-\cos\Big(\frac{\pi}{2N}\Big)}{1/2N}.\end{aligned} \tag{4.1-6}$$

Using the substitution $x = 1-\cos\Big(\frac{\pi}{2N}\Big)$ the first fraction can be evaluated as

$$\lim_{x\to 0} \frac{\ln[1-x]}{x} = \lim_{x\to 0} \ln\Big[(1-x)^{\frac{1}{x}}\Big] = \ln\frac{1}{e} = -1, \tag{4.1-7}$$

while the second term can be evaluated using l'Hopital's rule:

$$\lim_{N\to\infty} \frac{1-\cos\Big(\frac{\pi}{2N}\Big)}{1/2N} = \lim_{N\to\infty} \frac{\frac{\pi}{2}\sin\Big(\frac{\pi}{2N}\Big)\frac{\pi}{2N^2}}{1/2N^2} = 0. \tag{4.1-8}$$

Hence

$$\lim_{N\to\infty} \ln P_1 = -1\cdot 0 = 0, \tag{4.1-9}$$

and

$$\lim_{N\to\infty} P_1 = 1. \tag{4.1-10}$$

The success probability of transmitting a 1-bit can therefore be made arbitrarily high by concatenating a sufficient number of beamsplitters.

4.2 Evaluation of counterfactuality

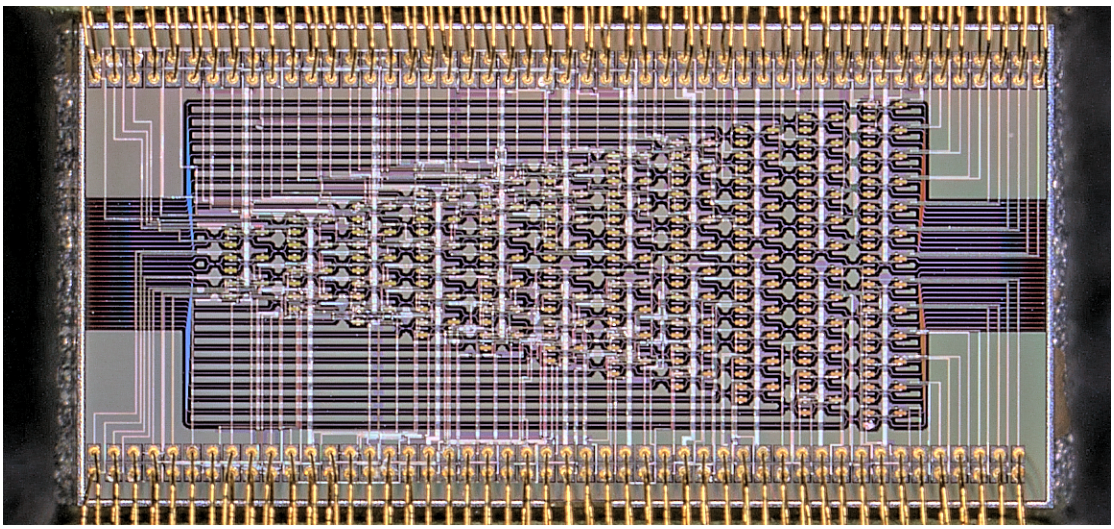


Figure 4.3: **Nanophotonic processor.** The figure shows a micrograph of the photonic waveguide used to implement the counterfactual communication protocol. The gold wires at top and bottom connect the thermal phase-shifters of the waveguide to a printed circuit board. The input of the waveguides is on the left side of the picture, and the MZIs are arranged along several diagonals. The thick vertical traces are the ground wires of the thermal phase-shifters, while the thin wires are used to set the individual voltages across the heaters. Image courtesy of Nicholas Harris.

Since the purpose of the above protocol is to allow for counterfactual communication without the photon leaving a weak trace in Bob's laboratory, we will here motivate the absence of a trace using the same methods as described in Section 1.3.4. Note that it is only necessary to analyse the outcome in which the photon returns to Alice, since in this context a violation of counterfactuality means that some part of the photon wavefunction propagated from Bob to Alice. An illustration of the forwards and backwards propagating states for the 1-bit is shown in Fig. 4.2. Since the trajectories of these states never overlap in Bob's laboratory the photon does not leave a weak trace there, and the protocol is counterfactual. Note, however, that if there are imperfections present in the interferometers, such as phase noise or incorrect splitting ratios in the beamsplitters, then there is a finite probability that a photon enters Bob's laboratory and travels back to Alice, leading to a violation of the counterfactuality in the 0-bit case.

4.3 Nanophotonic processor

There are two main challenges associated with an experimental implementation of the Arvidsson-Shukur and Barnes CFC protocol. The first is constructing a stable and precisely tunable beamsplitter that has exactly the reflectivity given by (4.1-1). As discussed in Section 3.5.2 this could be done using a Sagnac interferometer, however this still leaves the second problem, which is the stabilization of the concatenated MZIs. This

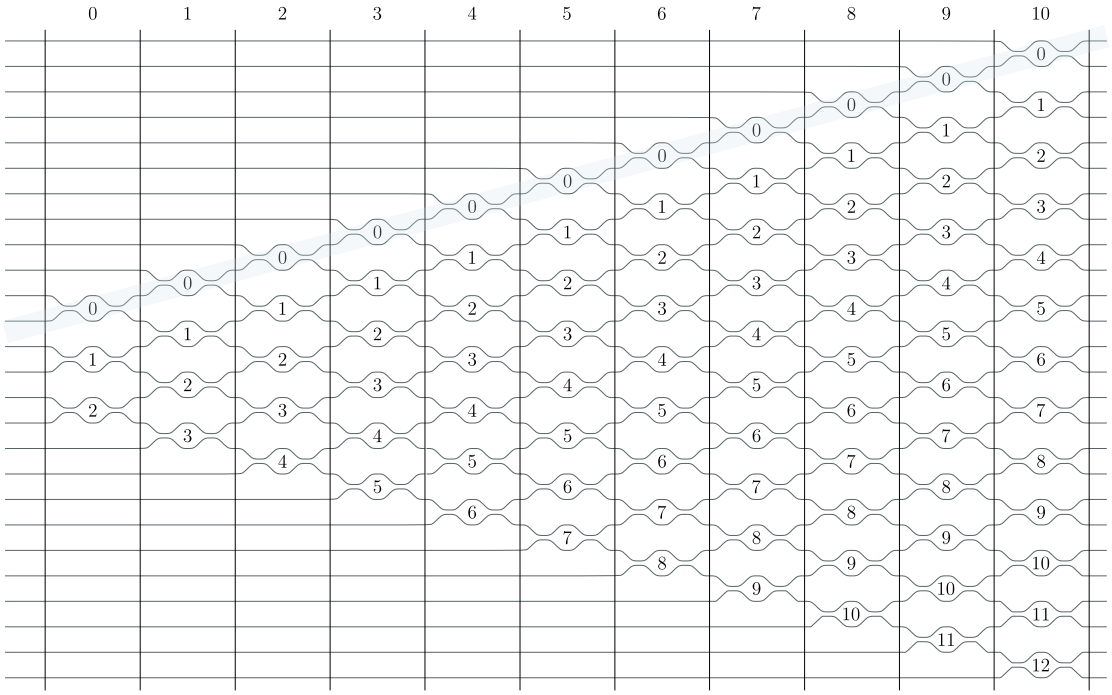


Figure 4.4: **Programmable waveguide layout.** The waveguide used in the experiment contains 88 programmable Mach-Zehnder interferometers spread out over 11 columns, and arranged in a wedge-like fashion. Each Mach-Zehnder interferometer contains two thermal phase shifters; one between the two directional couplers, and one at the output of the MZI, corresponding to the phase shifters θ and γ in Fig. 2.37. The lightly shaded area indicates a diagonal along which the internal phase shifters are calibrated.

is not an insurmountable challenge, however in order to reduce the experimental complexity the decision was made to implement the protocol in a programmable waveguide, also referred to as a programmable nanophotonic processor (PNP). Since the waveguide exhibits passive phase stability it can be used to realise complex interferometric structures, and this is indeed the point of such waveguides.

The waveguide used in this experiment was designed by Dr. Nicholas C. Harris, under the supervision of Professor Dirk Englund at MIT [389]. This photonic chip was also used in the publication [390], which is not elaborated on in this thesis. A micrograph of the device is shown in Fig. 4.3. The chip contains 26 silicon waveguides with a silica cladding, and the high refractive index contrast between the two materials results in a very strong mode confinement. This allows a total of 88 Mach-Zehender interferometers to fit into the approximately 5 mm × 2.5 mm footprint. These MZIs have one external and internal thermo-optic phase-shifter each, which are capable of modulation speeds approaching 100 kHz [335]. The 88 MZIs are arranged along several (anti-)diagonals, forming an isosceles trapezoid, or a wedge shaped structure. The full structure of the device is shown in Fig. 4.4. The bulk-optics picture of the CFC protocol shown in Fig. 4.1 can be translated into the waveguide by mapping every beamsplitter and mirror to an MZI. The act of Bob blocking a mirror to transmit a 1-bit is then modelled by making the MZI fully transmissive, which corresponds to Bob simply removing the

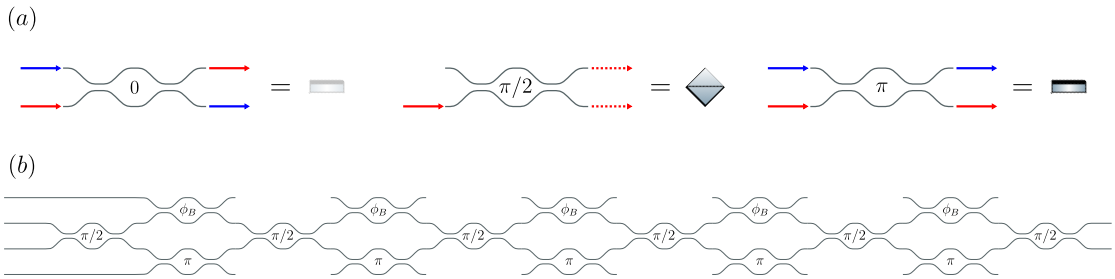


Figure 4.5: **Waveguide implementation of CFC protocol.** (a) The different basic element of the protocol are beamsplitters and mirrors, as well as the absence of a mirror. These can be implemented by MZIs with the appropriate internal phase, as indicated in the figure. (b) The 11 columns of the chip allow for an implementation of the CFC protocol with $N = 6$ beamsplitters. The phase ϕ_B is chosen by Bob according to which bit value he wants to transmit to Alice. $\phi_B = \pi$ corresponds to leaving the mirrors in place and sending a 0-bit, while $\phi_B = 0$ removes the 'mirrors' and transmits a 1-bit by returning the photon to Alice.

mirror instead of blocking it. The number of beamsplitters is limited by the number of columns of MZIs in the chip, and as shown in Fig. 4.5 the 11 columns of the chip lead to a maximum number of beamsplitters $N = 6$.

4.3.1 Device calibration

The thermal phase shifters in the chip were controlled via a voltage source, and in order to perform well-defined operations inside the chip it is necessary to calibrate each phase shifter. In practice this calibration consists of determining a voltage-to-phase map by observing the optical power at the output of the chip as a function of a bias voltage. The calibration method used for the chip was adapted from [391]. The internal phase shifters are first calibrated by attempting to send light along the diagonals with constant indices, one of which is indicated in Fig. 4.4. First, the voltage across the internal phase-shifter in the first column is scanned, while the power in the two photodetectors at the end of the diagonal is monitored. Since the only way for the light to get to these two detectors is to travel along the diagonal, this measurement is equivalent to directly observing the output power in the top port of the first MZI. One then repeats this procedure for each MZI along the diagonal, before moving on to the next diagonal. The voltage-to-phase map is created by fitting a sine function to the observed output power. In practice phase is a nonlinear function of the voltage, and a good fit requires a model incorporating electron drift velocity saturation [391].

The calibration for the external phase shifters follows the same idea. However, in order to be sensitive to the phase they induce it is necessary to create a compound MZI, consisting of four physical MZIs. This is illustrated in Fig. 4.6. One difference with respect to the internal phase shifters is that each compound MZI contains two phase shifters, and there is hence an overall phase ambiguity. This can be solved by simply treating one of the phases as a constant reference, and calibrating the remaining ones with respect to this one [391]. A second difference is that imperfections in the device can lead to imperfect information about the interferometric phase, since it is possible for the light to take multiple paths to the same detector. In practice this manifests itself as

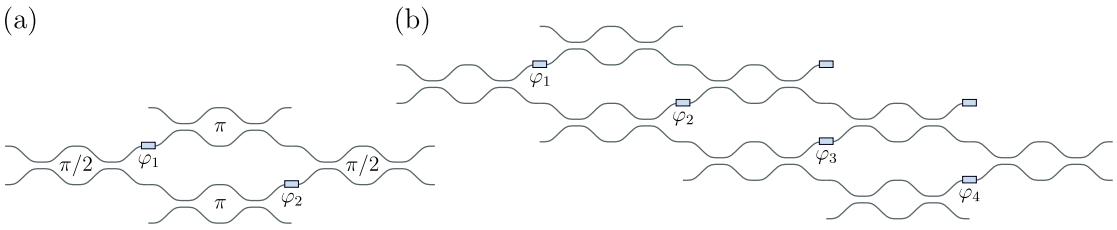


Figure 4.6: **External phase shifter calibration. (a)** To calibrate the external phase shifters of the individual MZIs a compound MZI is constructed using four physical MZIs. Two of these, with a $\pi/2$ internal phase, act as beamsplitters, and two with an internal phase of π act as mirrors. **(b)** The calibration of the external phase shifters proceeds along a diagonal and the light is routed to/from the desired compound MZI using SWAP operations. Compound MZIs that lie on a diagonal share phase shifters, and this leads to an ambiguity in the phase when performing the calibration, which can be remedied by treating one phase as a constant reference and applying an appropriate offset to the remaining phases [391]. Also note that imperfect visibility in the constituent MZI can negatively influence the calibration of the external phase shifters, as light that leaks through the 'mirror' MZIs can make its way to the detector and influence the phase measurement.

a lower visibility of the compound MZIs as compared to the physical ones, and degrades the calibration.

4.3.2 MZI visibility

In order to perform high fidelity operations the MZIs in the chip must exhibit high visibility, or equivalently, a high extinction ratio. One of the main imperfections affecting the operation of the device is the variance in the splitting ratio of the directional couplers in the MZI unit cells. The coupling constant $\mathscr{C}$ from (2.5-3) is inversely proportional to the wavelength [106, 391] and the directional couplers therefore only have a finite linewidth over which they produce a balanced splitting ratio. This balanced splitting ratio is a necessary condition for achieving high extinction ratios, which is easy to understand through the following argument: in one exit port of an MZI light will either have been reflected twice by a beamsplitter, or transmitted twice. To observe full destructive interference in this output port of the MZI, one therefore needs that:

$$\mathcal{R}^2 + e^{i\varphi}\mathcal{T}^2 = 0, \tag{4.3-1}$$

which is only possible for $\mathcal{R} = \mathcal{T}$. Note that in the other port of the MZI light is always reflected by one beamsplitter and transmitted through the other, and destructive interference is therefore possible as long as

$$\mathrm{R}_1\mathcal{T}_2 + e^{i\varphi}\mathrm{T}_1\mathcal{R}_2 = 0, \tag{4.3-2}$$

which only requires that the two directional couplers be identical, not necessarily unbiased. Due to the highly miniaturized nature of the waveguide used in this work, the interaction regions in the directional couplers is very short, and this increases their sensitivity, in absolute terms, to errors in the length of this interaction region.

In practice the various directional couplers present in the device have a center wavelength at which they produce a balanced splitting ratio, and this center wavelength

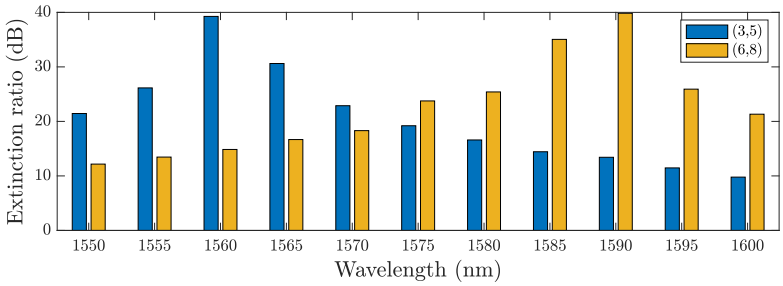


Figure 4.7: **Wavelength dependence of MZI SWAP operation extinction ratio.** When performing a SAWP operation the extinction ratio depends on the optical wavelength. This dependence is illustrated above for two different MZIs performing SWAP operations. The MZIs are labelled as (c, d), where c is the column and d is the diagonal, as indicated by Fig. 4.4. It can be seen that the extinction ratio drops by as much as a factor of 10 within 5 nm, and the linewidth of the MZIs is significantly smaller than the difference between their center wavelengths.

follows some distribution. As long as the width of this distribution is smaller than the linewidth of the individual directional couplers there exists a wavelength region within which the device can operate well. In a waveguide the operation corresponding to the condition (4.3-1) is the SWAP operation, which sends a photon entering in the top input mode of an MZI to the bottom output mode, and vice versa. Fig. 4.7 shows the extinction ratio $P_{\min}/P_{\max}$ for two MZIs on the bottom diagonal of the device used for the experiment. It can be seen that the best operating points for the two MZIs differ by much more than the effective wavelength bandwidths of the MZIs. This means that there is fundamentally no wavelength at which all the different MZIs on the chip can operate with a high degree of fidelity.

While two MZIs shown in Fig. 4.7 are outliers in the overall distribution of center wavelengths, choosing a wavelength at which the average performance is maximized still results in significantly hampered performance compared to the what each MZI is capable of individually. The mean extinction ratio for all the MZIs on the lower diagonal is shown in Fig. 4.8. This diagonal was chosen because it allows for direct measurement of the SWAP-operation splitting ratio, that is not affected by stray light travelling through alternate paths through the chip. From the figure it can be seen that the best average performance is achieved at a wavelength around 1575 nm.

For the actual experiment, a single-photon wavelength of $\lambda = 1565.8$ nm was chosen. This was due to the fact that the CW laser available at the time was limited to the telecom C-band, and was not able to perform a wavelength sweep sufficiently wide to identity the optimal working wavelength. Since the act of Bob removing a mirror is mapped onto a SWAP operation in the experiment, the relatively low fidelity of this operation is one of the major sources of error in the experiment. A secondary impact of the low fidelity SWAP operations is that the quality of the calibration of the external phase shifters is degraded, since as described in the previous section this calibration involves routing the light using SWAP operations.

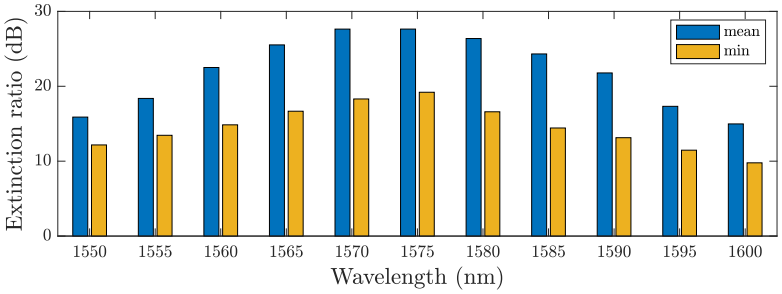


Figure 4.8: **Mean and minimum MZI extinction ratio.** The optimal wavelengths of the MZIs in the nanophotonic processor follows a wide distribution. Optimal operation of the device requires a trade-off between the average and worst-case performance. Above the mean and minimum extinction ratios for the 11 MZIs lying on the lowest diagonal of the photonic chip are shown. The best performance for this particular set of MZIs is reached around $\lambda = 1575\,\text{nm}$, with mean and minimum extinction ratios of $\text{ER}_{\text{mean}} = 27.6\,\text{dB}$ and $\text{ER}_{\text{min}} = 19.2\,\text{dB}$.

4.3.3 Crosstalk

In an ideal device setting a certain phase in a given phase shifter would not have any effect on the other phase shifters. In practice, however, this is not the case and the unintended influence of one phase shifter on another is known as crosstalk. There are two main sources of crosstalk in the waveguide. The first one is thermal crosstalk, which occurs due to the fact that the different elements on the chip are in close proximity. The heat produced by one thermal phase shifter will inevitably diffuse throughout the entire device, to at least some degree. This can be mitigated by engineering the device in such a way that most of the heat is dissipated perpendicularly to the waveguide, and by attempting to minimize the power consumption of the resistive elements. The second, and dominant source of crosstalk in the waveguide is voltage crosstalk. This crosstalk originates from the fact that all of the resistors that constitute the thermal phase shifters are connected to the same ground. Since there is a finite resistance in the ground wires this can lead to a floating ground for some or most of the phase shifters on the chip. A circuit modelling this behaviour is shown in Fig. 4.9.

In practice this means that the true voltage across a phase shifter is a function of the voltage applied on every other phase shifter in the device:

$$\begin{aligned} V_1' &= f_1(V_1, V_2, \ldots, V_N) \\ V_2' &= f_2(V_1, V_2, \ldots, V_N) \\ &\vdots \\ V_N' &= f_N(V_1, V_2, \ldots, V_N). \end{aligned} \tag{4.3-3}$$

This set of coupled equations can in principle be solved by measuring all the voltage correlation functions, truncating the functions f_k to a given order, and then inverting (4.3-3). If the functions are truncated to the first order, then the problem reduces to inverting a matrix; for higher orders methods for nonlinear systems of equations have to be employed, for example Newtons method, increasing the computational cost. While the problem is numerically tractable, the experimental determination of the functions f_k

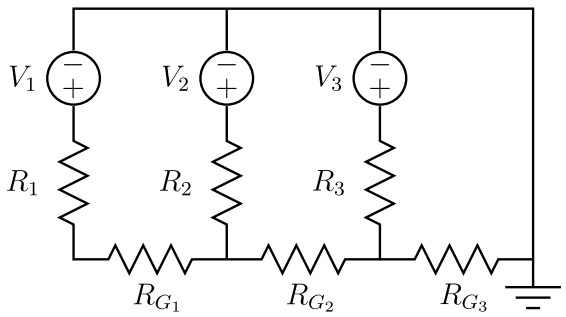


Figure 4.9: **Crosstalk model.** The voltage crosstalk in the programmable waveguide can be modelled by a single electrical circuit. In an ideal circuit the voltage sources V_k would directly apply a voltage across the resistors R_k inside the photonic chip, as well as the corresponding ground resistance R_{G_k}. In practice, some or most of the ground wire is shared by the different voltage sources. The problem that this creates is that when say the third voltage source is active and current is flowing through R_3 and R_{G_3} this causes the ground for V_1 and V_2 to float, changing the current through R_1 and R_2. This in turn changes the dissipated power, and therefore the thermo-optic phase shift.

requires N^2 measurements, where N is the number of phase shifters in the device. Since there are two phase shifters per MZI, this works out to $N = 176$ and $N^2 = 30976$. In the setup used for the experiment this measurement could not be automated, and due to the extremely high number of measurements required it was deemed infeasible. In Ref. [392] the crosstalk for an equivalent photonic waveguide was measured, and found to be as high as $100\,\mu\mathrm{V\,V^{-1}}$. Individual phase shifters in the chip can be biased as high as 10 V, giving a crosstalk of 1 mV. Considering the high number of phase shifters on the chip this can, in a worst-case scenario, be enough to reduce the extinction ratio of the affected MZI by a factor of two. As a cumulative effect the voltage crosstalk therefore significantly degrades the performance of the device.

Since this effect was not pre-corrected for in the experiment, it was not possible to directly implement the circuit for the CFC protocol simply from the voltage-to-phase maps obtained from the calibration. Instead, the correct phases were found by iteratively optimizing the circuit, in particular the external phase shifters, to obtain the correct output distribution for the 0-bit case. The downside of this approach is that one cannot independently verify that every single component in the circuit implements the correct operation, since there are many possible parameters that lead to the same output distribution. However, under the assumption that the initial parameters from the device calibration are sufficiently close to the correct minimum of the optimization, this process will yield voltage/phase values that do indeed implement the correct circuit. This is to some extent validated by the fact that the 1-bit circuit was not optimized over, and still produced the expected output distribution.

4.3.4 Input/output coupling

The dimensions of the ridge waveguides inside the device are $500\,\mathrm{nm} \times 220\,\mathrm{nm}$, and the transverse mode of the light is largely contained within the waveguide. This mode is more than an order of magnitude smaller than the mode of a standard telecom SMF-28 fiber, and can therefore not be efficiently out-coupled directly. To address this problem,

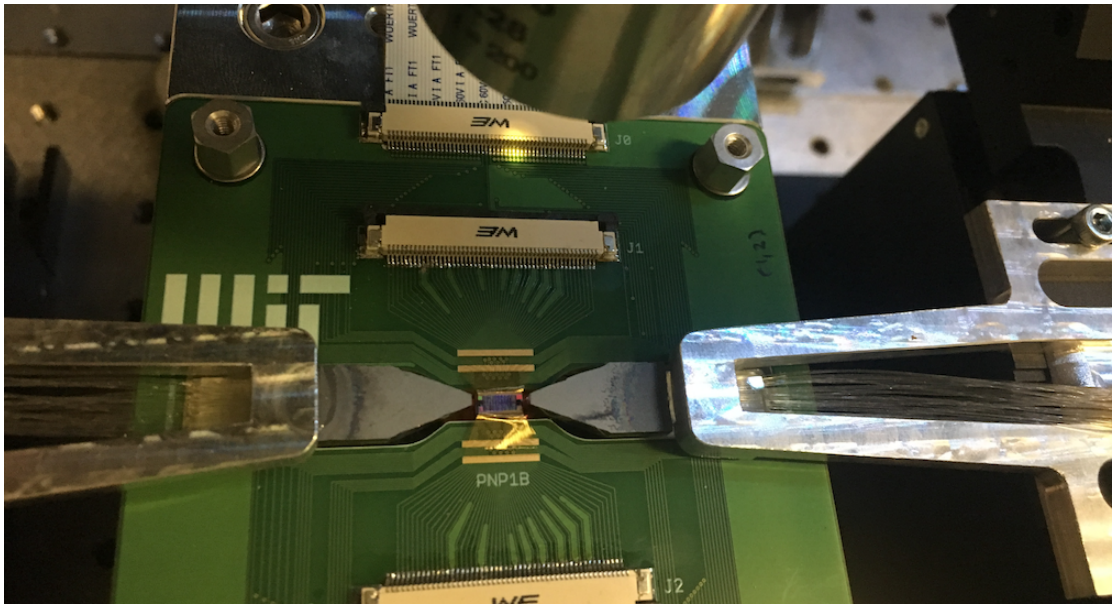


Figure 4.10: **Waveguide setup.** The picture shows the setup used to couple and control the nanophotonic processor. The wedge-shaped grey waveguides are the interposer chips that convert the mode from the 10 µm SMF-28 fiber mode, to the 2 µm PNP input mode. Between the two interposers the PNP itself is visible, along with the wirebonds to the PCB, which in turn is connected to a digital voltage controller using four 50-pin FFC cables (one pictured). At the top of the image a 10x Mitutoyo long working distance objective can be seen, which was used together with Optem Fusion lens system to image the PNP, and give visual feedback during the coupling process.

the waveguides on the input/output facets of the chip are inverse tapered to a mode field diameter of 2 µm. Since this is still too small to efficiently couple into single-mode fiber, an additional mode converter is necessary. To this end a Si_3N_4 waveguide with an SiO_2 cladding was designed by Nicholas Harris and manufactured by LioniX. This mode converter, or interposer, has a 10 µm mode field diameter with an 127 µm pitch on the input side, and is tapered to 2 µm with a 25.4 µm on the output, matching the programmable waveguide. Light is then coupled in and out of the waveguide by butt-coupling the interposer with the PNP. An image of the experimental setup, including the interposer, PNP and wirebonded PCB is shown in Fig. 4.10.

While these interposers serve the purpose of decreasing the coupling loss, the mode-converter waveguides nevertheless introduce some excess loss. A histogram of excess loss for the two interposers is shown in Fig. 4.11. The mean loss across all modes is approximately 2 dB, for a total 4 dB fiber-to-fiber loss. The highest transmission measured through the device was 18 %, equivalent to approximately 7.5 dB of loss. Out of this at least 1 dB comes from the intrinsic propagation loss of the waveguides [389], giving a coupling efficiency of at least 75 %, or −1.25 dB.

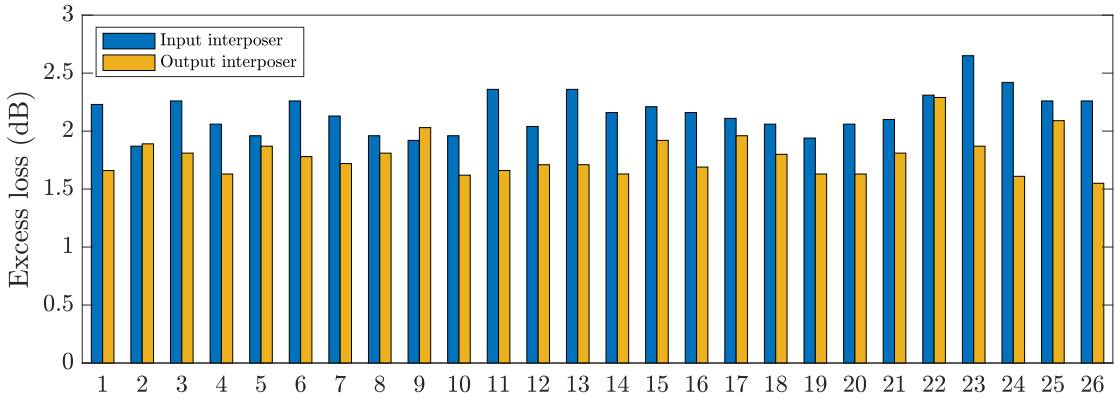


Figure 4.11: **Interposer loss.** The interposers used to couple in and out of the waveguide were characterised in terms of the per-mode loss. The measured loss in dB is shown above on a per mode basis, for both the input and output interposers. The mean loss for the input (output) interposer is 2.16 dB (1.78 dB).

4.3.5 Black-body radiation

Since the thermal phase shifters in the PNP heat up the waveguide 80 °C or more when applying a π phase shift [335] the waveguide emits a significant amount of black-body radiation that can cause accidental photon detection events in NIR single-photon detectors. This is exacerbated by the fact that many of these NIR and IR photons are emitted directly into the waveguide mode(s). During normal operation of the device the background count rate from the black-body photons was in the MHz range. When using heralded single-photon source, such as an SPDC source, this problem can be alleviated by exploiting the strong temporal correlation between the signal and idler photons, and using this to filter the detection events in time. Despite this, the black-body radiation contributes excess noise to the experiment in the form of accidental coincidences and needs to be filtered. One way of doing this is by exploiting the strong wavelength dependence of bending losses in optical fibers [393,394], to create a makeshift short-pass filter. This can be further enhanced by placing the bent fiber inside a box made from, or coated with, a low emissivity material, to decrease the effective temperature of the surroundings.

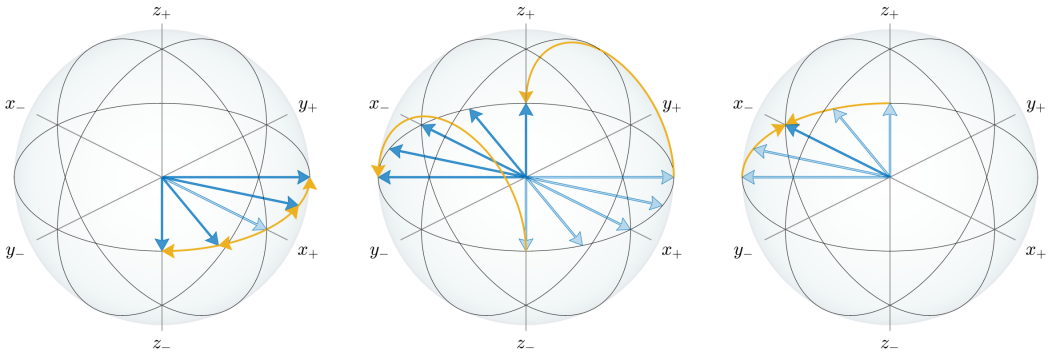
While this approach was initially considered, for the experiment it was found that a simpler solution using coarse wavelength-division multiplexers (CWDM) sufficed to filter out the background radiation. While CWDMs are not designed to provide broad-band spectral filtering, it was assumed that most of the black-body photons generating spurious detection events were in or close to the telecom C-band, and that the CDWMs would provide sufficient filtering in this wavelength range. In order to achieve sufficiently high suppression of the background two CWDMs were used in series for each of the two detected output modes.

5 Experimental methods for "Demonstration of universal time-reversal for qubit processes"

In this chapter the experimental methods and design decisions behind Publication 2, as well as some of the theoretical background, will be elaborated on. The goal of the experiment was to perform a demonstration of the time translation protocol presented in Refs. [395,396] for a qubit system. This protocol allows one to transform the unitary evolution of an arbitrary qubit state into its inverse, without requiring any knowledge of the initial quantum state of the qubit, the Hamiltonian governing its time evolution, or even the effect that one's interventions have on the system in question; the only requirement is that these interventions are repeatable. The fact that the protocol works in this very general setting is what bestows on it the quality of being universal.

5.1 The rewinding protocol

While the question of how to transform a unitary evolution into its inverse may seem abstract at first, it is in fact a problem that has been studied in various forms for a long time. The most well known example of a rewinding protocol is the spin-echo technique in nuclear magnetic resonance [397]. This protocol is applied to an ensemble of spins precessing on the equator with some distribution of precession frequencies, causing them to decohere. By performing a 180° rotation about an axis in the equatorial plane at a time $t = \tau$ the spins begin to refocus, and become fully aligned at $t = 2\tau$. This process, illustrated in Fig. 5.1, works independently of the distribution of precession frequencies, and is a simple example of a quantum state being made to reverse its time evolution. Generalizations of spin echoes led to the development of so-called dynamical decoupling techniques, which remove the influence of decohering environmental noise through the application of suitably tailored control pulses [398,399]. Such methods can also be used to refocus a quantum system independently of the free time evolution it is undergoing, albeit probabilistically [400]. These protocols, however, require the ability to perform well characterised operations on the system in question. For example, in the case of the spin echo one needs to perform exactly a 180° rotation for the ensemble to become refocused. In practice this means that control imperfections limit the accuracy of these protocols [401]. Independently of this though, one can ask the fundamental question of whether such rewinding protocols could be universal. In this context, a protocol is said to be universal if it, in addition to working for any input state and free time



(a) $t = \tau$ (b) $t = \tau + dt$ (c) $t = 2\tau + dt$

Figure 5.1: **Spin echo**. A spin echo is a technique for refocusing a quantum system precessing on the equator of the Bloch sphere. The technique is independent of the precession frequency, and therefore also works on ensembles of quantum states with different precession frequencies. (a) an ensemble of states, initially all aligned with the $|x_+\rangle$ state at time $t = 0$, spreads out as it evolves for $t = \tau$ units of time, due to the different states rotating around the z-axis at different frequencies. (b) a π rotation about the y-axis is applied on the ensemble. This rotation is assumed to take $dt << \tau$ units of time. (c) since the states have effectively been mirrored in the $x - z$ plane, they now behave as though their precession were reversed, and the ensemble is refocused such that at time $t = 2\tau$ all the states are once again aligned. The dark (light) blue arrows show the final (initial) states, and the yellow arrows indicate the trajectory traced out by the time evolution.

evolution, does not require the interventions on the system to be well characterised, or even known. This question whether such a protocol exists was first affirmatively answered in Ref. [402], and the authors later later built upon this work in Refs. [395,396]. The experimental realisation of the protocol proposed in the latter two papers is the subject of Publication 2.

The rewinding protocol is based on the following relation:

$$[U, V]U^N[U, V] \propto U^{-N}, \tag{5.1-1}$$

which holds for any invertible 2×2 matrix U, and any 2×2 matrix V. This relation can be proved by first noting that due to the cyclic property of the trace, the trace of the commutator between two matrices is always zero:

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA) \implies \mathrm{Tr}(AB - BA) = 0. \tag{5.1-2}$$

Then, let $C = [A, B]$. The characteristic polynomial of this matrix can be written

$$\begin{aligned} p_C(\lambda) &= \lambda^2 - \mathrm{Tr}(C)\lambda + \det(C)I \\ &= \lambda^2 + \det(C)I. \end{aligned} \tag{5.1-3}$$

Through the Cayley-Hamilton theorem this implies that

$$C^2 = -\det(C)I. \tag{5.1-4}$$

The desired relation can now be found by considering the commutator

$$[U, U^N V] = U^N[U, V], \tag{5.1-5}$$

and squaring it:

$$[U, U^N V]^2 = U^N [U, V] U^N [U, V] \propto I. \tag{5.1-6}$$

Multiplying this expression with U^{-N} from the left recovers (5.1-1). If one interprets U as the free time evolution of a quantum system, and V as a perturbed time evolution, then it's clear that if one is able to realise the commutator between the free and perturbed time evolutions, then one can propagate the system backwards in time. Since the relation holds for any 2×2 matrix V, the perturbed time evolution can be both arbitrary and unknown.[1] The perturbed time evolution V can therefore be generated by an uncharacterised interaction, and this, together with the fact that the relation holds for any U, is what makes the protocol universal.

The fact that the left-hand side in (5.1-1) is only proportional to the inverse unitary is a reflection of the fact that the commutator cannot be applied deterministically, since in general the commutator is not unitary. This means that any attempt to rewind the system using this relation will be probabilistic. As will be discussed in the next section, one way to probabilistically apply the commutator is to prepare a controlled superposition of the commutator and the anti-commutator $\{U, V\} = UV + VU$ by using an auxillary control system C:

$$\frac{1}{2}[U, V] \left|\Psi\right\rangle_T \otimes \left|0\right\rangle_C + \frac{1}{2}\{U, V\} \left|\Psi\right\rangle_T \otimes \left|1\right\rangle_C . \tag{5.1-7}$$

Projecting the control qubit onto the state $\left|0\right\rangle_C$ then applies the commutator. Here the subscript T labels the target system on which the gates U and V act. If the projection of the control qubit yields the outcome '1' the anti-commutator gets applied instead of the commutator, and the relation (5.1-1) can seemingly no longer be used. However, in Ref. [396] it was shown, that using the two following relations, an inadvertently applied anti-commutator can always be turned back into a commutator:

$$[U, V]^2 \propto I \tag{5.1-8}$$

$$\{U, V\}^m [U, V] \{U, V\}^m \propto [U, V], \tag{5.1-9}$$

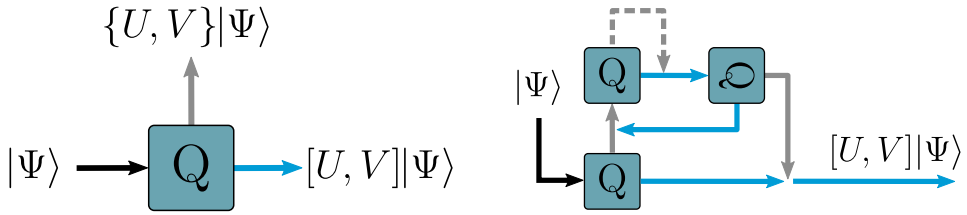
and these relations can be used to boost the success probability to an arbitrary level. The first relation is simply (5.1-4), while the second one is less obvious. Here we will restate the proof from [396]. Consider

$$\mathrm{Tr}\big([U, V]\{U, V\}^n\big). \tag{5.1-10}$$

For $n = 0$ we already saw that this expression evaluates to zero, and for $n = 1$:

$$\begin{aligned} \mathrm{Tr}\big([U, V]\{U, V\}\big) &= \mathrm{Tr}\big(UVUV + UVVU - VUUV - VUVU\big) \\ &= \mathrm{Tr}\big(UVUV - VUVU\big) + \mathrm{Tr}\big(UVVU - VUUV\big) \\ &= 0, \end{aligned} \tag{5.1-11}$$

[1] If U and V commute the protocol fails, however the measure of the set of commuting 2×2 matrices is zero, and for two time evolutions picked at random the probability for them to commute is therefore infinitesimally small.



(a) Commutator device (b) Virtual roadmap

Figure 5.2: **Adaptive error correction.** (a) The basic building block of the rewinding protocol is a device, labelled here as Q, that probabilistically applies the commutator of two unitaries U and V on an input state $|\Psi\rangle$. In the figure the port corresponding to the commutator is indicated by a light blue color, and the anti-commutator port is shown with a grey color. (b) The protocol requires the commutator to be applied, but this only happens with some probability. In the event that a Q-device applied the anti-commutator instead additional uses of Q can be used to convert the anti-commutator into a commutator. The process for doing this is illustrated using a virtual roadmap, and as long as the state exits in the lower right of this maze the commutator will have been applied. The most direct path consists of applying the commutator and then the anti-commutator, however this can also only be done probabilistically. If the anti-commutator is applied in the second step, this too needs to be corrected, and the dashed line represents recursive applications of the adaptive correction protocol.

due to the cyclic property of the trace. Since $\{U, V\}$ is a 2×2 matrix the Cayley-Hamilton theorem implies that $\{U, V\}^n$ can be written as a linear combination of I and $\{U, V\}$, and (5.1-10) can therefore be reduced to the $n = 0, 1$ cases. Hence

$$\mathrm{Tr}\big([U, V]\{U, V\}^n\big) = 0. \tag{5.1-12}$$

Since, as noted earlier, the square of any traceless 2×2 matrix is proportional to the identity matrix, we have:

$$[U, V]\{U, V\}^n[U, V]\{U, V\}^n \propto I, \tag{5.1-13}$$

and multiplying this expression with $[U, V]$ from the left recovers (5.1-9). The way in which (5.1-8) and (5.1-9) can be used to adaptively correct unwanted applications of an anti-commutator can be illustrated using a virtual roadmap shown in Fig. 5.2. Whenever the state exits in the lower right port of this 'maze' the commutator will have been applied. The most direct path corresponds to applying (5.1-9) directly with $m = 1$, however the application of the commutators and anti-commutators in the desired order is naturally also probabilistic. An additional result presented in Ref. [396] is that a photon entering the virtual maze will always exit in a finite number of steps, and this means that the protocol can achieve an arbitrarily high probability of success at the cost of a constant overhead in the running time. From (5.1-1) it can be seen that the time it takes to rewind a system N timesteps grows linearly with N, and the running time of the protocol is therefore $T + O(1)$, where the second term accounts for the error correction overhead. This linear scaling in the running time means that the system can essentially be rewound in 'real time'.

5.1.1 Comparison with similar protocols

Before moving on to the experimental implementation of the protocol, it is worth putting the underlying theory into the proper context by briefly comparing it to other similar protocols. The most closely related protocol is the resetting protocol presented in Ref. [402], and demonstrated in Refs. [403, 404]. This protocol works by preparing a set of four probes in a product of two singlet states:

$$|\Psi^-\rangle_{1,2}\,|\Psi^-\rangle_{3,4}\,. \tag{5.1-14}$$

These probes are then made to interact with the target system one at a time, after which they are projected onto a set of quasi-symmetric states, constructed from the symmetric two-qubit states. More concretely, these states are on the form

$$\frac{1}{A}(|k\rangle\,|l\rangle + |l\rangle\,|k\rangle), \tag{5.1-15}$$

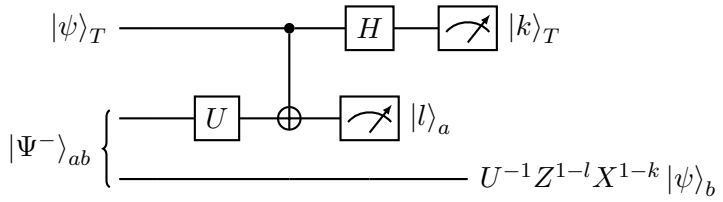
where A is a normalization constant and

$$|k\rangle\,, |l\rangle \in \{|0,0\rangle\,, |1,1\rangle\,, (|0,0\rangle + |1,1\rangle)\sqrt{2}\}, \tag{5.1-16}$$

are the symmetric two-qubit states. If the projection succeeds the target system is propagated by the identity operator. Similarly to the rewinding protocol that is the focus of this chapter, the probability of success depends on the particular interaction between the target and probe system.[2] If the joint target-probe evolution is separable then the success probability is exactly zero, in analogy to how the rewinding protocol fails for perfectly commuting free and perturbed time evolutions. When the projection fails the protocol can still be continued in order to increase the success probability, however in contrast to [396] there is no proof that the success probability approaches unity.

The second major difference between the resetting protocol of [402] and the rewinding protocol is the running time. As mentioned in the previous section, the running time of the rewinding protocol is $T + O(1)$, whereas the resetting protocol has a running time of $3T + O(1)$. Finally, the resetting protocol only propagates a system back to its initial state, hence the name. This is in contrast to the rewinding protocol which can realise higher powers of the inverse unitary.

A second protocol worth discussing is the method for reversing unknown unitaries presented in Ref. [405]. The protocol is based on gate teleportation, and for qubits it can be represented by the following circuit:

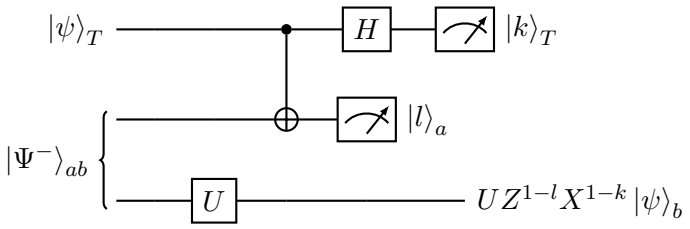


[2]The formulation of the rewinding protocol in the previous section did not explicitly include a probe system to induce the perturbed time evolution V, however such a description is given in the supplementary material of Publication 2.

Whenever the Bell-state measurement on the top two qubits results in the outcome $|\Psi^-\rangle$ the protocol succeeds:

$$(\langle\Psi^-|_{T,a} \otimes \mathbb{1}_b)\,|\psi\rangle_T \otimes (U \otimes \mathbb{1})\,|\Psi^-\rangle_{ab} \propto U^{-1}\,|\psi\rangle_b\,. \tag{5.1-17}$$

The fact that the circuit above works can be understood by starting from a standard gate teleportation circuit:



and noting that:

$$\mathbb{1} \otimes U\,|\Psi^-\rangle = U^{-1} \otimes \mathbb{1}\,|\Psi^-\rangle\,. \tag{5.1-18}$$

This relation is a direct consequence of the fact that the singlet state is anti-correlated in every basis:

$$\begin{aligned}\langle\Psi^-|U \otimes U|\Psi^-\rangle &= \langle\Psi^-|[\cos\theta\,\mathbb{1} - i\sin\theta(n_xX + n_yY + n_zZ)]^{\otimes 2}|\Psi^-\rangle \\ &= \langle\Psi^-|\cos^2\theta\,\mathbb{1}^{\otimes 2} - \sin^2\theta(n_x^2X^{\otimes 2} + n_y^2Y^{\otimes 2} + n_z^2Z^{\otimes 2})|\Psi^-\rangle \\ &= 1.\end{aligned} \tag{5.1-19}$$

since:

$$\begin{aligned}1 &= \langle\Psi^-|U \otimes U|\Psi^-\rangle \\ &= \langle\Psi^-|(U \otimes \mathbb{1})(\mathbb{1} \otimes U)|\Psi^-\rangle \\ &= (U^\dagger \otimes \mathbb{1}\,|\Psi^-\rangle)^\dagger(\mathbb{1} \otimes U)\,|\Psi^-\rangle\,,\end{aligned} \tag{5.1-20}$$

which is equivalent to (5.1-18). Similar relations exist for the other three Bell states. Writing these states as $|\psi\rangle = \mathbb{1} \otimes Z^lX^k\,|\Psi^-\rangle = X^kZ^l \otimes \mathbb{1}\,|\Psi^-\rangle$ one finds:

$$\begin{aligned}\mathbb{1} \otimes U\,|\psi\rangle &= \mathbb{1} \otimes UZ^lX^k\,|\Psi^-\rangle \\ &= X^kZ^lU^\dagger \otimes \mathbb{1}\,|\Psi^-\rangle \\ &= X^kZ^lU^\dagger Z^lX^kX^kZ^l \otimes \mathbb{1}\,|\Psi^-\rangle \\ &= X^kZ^lU^\dagger Z^lX^k \otimes \mathbb{1}\,|\psi\rangle\,.\end{aligned} \tag{5.1-21}$$

The transformation above can be understood as performing a specific set of sign flips in the Bloch vector $\vec{n}$ defining the rotation axis of U. More specifically, these sign flips correspond exactly to the minus signs of the $\sigma_i \otimes \sigma_i$ terms in the density matrices $|\psi\rangle\langle\psi|$. This offers a very intuitive way of understanding the protocol: since the density matrix of $|\Psi^-\rangle$ has a minus sign in all of its $\sigma_i \otimes \sigma_i$ terms:

$$|\Psi^-\rangle\langle\Psi^-| = \frac{1}{4}\left[I - X \otimes X - Y \otimes Y - Z \otimes Z\right], \tag{5.1-22}$$

the corresponding transformation of a unitary when exchanging the subspace that U acts on is:

$$\begin{aligned} U &= \cos\theta\,\mathbb{1} - i\sin\theta(n_x X + n_y Y + n_z Z) \\ &\mapsto \cos\theta\,\mathbb{1} - i\sin\theta(-n_x X - n_y Y - n_z Z) = U^{-1} \end{aligned} \tag{5.1-23}$$

Similarly, the density matrix of the state $|\Phi^+\rangle$ only has a minus sign in the $Y \otimes Y$ term, and hence the transformation of the unitary is a sign flip in the Y term, which corresponds to the transpose:

$$\mathbb{1} \otimes U\,|\Phi^+\rangle = U^T \otimes \mathbb{1}\,|\Phi^+\rangle\,. \tag{5.1-24}$$

The teleportation circuit of the protocol is probabilistic, since 3/4 of the time the Bell-state measurement will not project the state onto $|\Psi^-\rangle$, but one of the other three Bell states, thereby giving the output state:

$$U^{-1} Z^{1-l} X^{1-k}\,|\psi\rangle\,, \tag{5.1-25}$$

where at least one of the Pauli terms is present. These Pauli terms cannot be undone, since they act before U^{-1}. These failed outcomes can still be corrected though, by applying U, then the Pauli terms defined by the Bell-state measurement in the reverse order and finally attempting the teleportation again. Unlike the resetting scheme described at the beginning of the section, it's obvious that the success probability of this method converges to 1. However, the error correction requires the ability to implement specific Pauli operations in accordance with the error syndrome, and the protocol does therefore not meet the definition of universality used here.

5.2 Photonic realisation of the rewinding protocol

The rewinding protocol described in the previous section requires the ability to implement the commutator between two time evolutions. This can be done using what is known as a quantum switch [54, 55], and the fact that this device has a relatively straightforward optical realisation makes the protocol particularly suited to photonic implementations [11, 406]. A quantum switch acts on two unitaries U and V and transforms them into a controlled superposition of the operations being applied in different orders:

$$(U, V) \mapsto UV \otimes |0\rangle\langle 0|_C + VU \otimes |1\rangle\langle 1|_C\,. \tag{5.2-1}$$

To see how this transformation can be used to generate the commutator, consider the superposition of operations above acting on the state $|\Psi\rangle_T \otimes |+\rangle_C$:

$$\begin{aligned} &\big(UV \otimes |0\rangle\langle 0|_C + VU \otimes |1\rangle\langle 1|_C\big)\,|\Psi\rangle_T \otimes |+\rangle_C \\ &= \frac{1}{\sqrt{2}} UV\,|\Psi\rangle_T \otimes |0\rangle_C + \frac{1}{\sqrt{2}} VU\,|\Psi\rangle_T \otimes |1\rangle_C \\ &= \frac{1}{2} UV\,|\Psi\rangle_T \otimes (|+\rangle_C + |-\rangle_C) + \frac{1}{2} VU\,|\Psi\rangle_T \otimes (|+\rangle_C - |-\rangle_C) \\ &= \frac{1}{2}\{U, V\}\,|\Psi\rangle_T \otimes |+\rangle_C + \frac{1}{2}[U, V]\,|\Psi\rangle_T \otimes |-\rangle_C\,. \end{aligned} \tag{5.2-2}$$

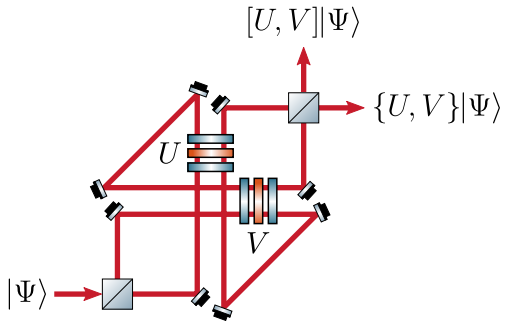


Figure 5.3: **MZI quantum switch.** A quantum switch is a device which realises a coherent superposition of two or more gate orders. The most common way of building a quantum switch is using a Mach-Zehnder interferometer with a folded geometry, aligned in such a way that the two arms of the interferometer pass through the polarization gadgets implementing the unitaries U and V in different orders. Depending on which port of the output beamsplitter the photon exits in either the commutator or anti-commutator is applied.

Projecting the control degree of freedom on the state $|-\rangle$ successfully implements the commutator. The most common way to generate (5.2-1) in a photonic setting is by spatially superposing a single photon in two different modes using a balanced beamsplitter. The discrete time evolutions U and V can be realised using polarization gadgets. Aligning the two photon paths in such a way that they propagate through the gadgets in different orders then generates the desired transformation. In such an implementation, the photon path acts as the control degree of freedom, while the photon polarization acts as the target. A second beamsplitter can then be used to perform a Hadamard transformation on the control degree of freedom, thereby correlating the (anti-)commutator with the spatial mode of the photon. This is simply a Mach-Zehnder interferometer containing two polarization gadgets, and an example drawing of such setup is shown in Fig. 5.3.

The rewinding protocol could be realised by constructing two such interferometric quantum switches, and performing the free time-evolution transformation U^N on the polarization as the photon transits from one quantum switch to the other. However, such an implementation is conceptually unappealing, since it would necessitate multiple physical realisations of the transformations U and V. This would in turn require the transformations to be known, which stands in contrast to the universality of the protocol. It was therefore decided to only use a single polarization gadget for each of the two time evolutions U and V, and to actively route the photons through the setup such that they could pass through the same polarization gadgets, as well as the quantum switch, multiple times. This routing will be discussed more in Section 5.2.3.

Active routing could also be used to implement the adaptive error correction protocol shown in Fig. 5.2, however this was not done in the experiment. Instead, a single-step version of the protocol, which, as shown in Fig. 5.4, fails if the anti-commutator is applied, was implemented. The reason for this is that it was deemed sufficiently challenging and interesting to implement the protocol even in the absence of the adaptive corrections. The two main challenges associated with the adaptive feed-forward is the additional photon loss, and the inability to herald a failure. The latter would require access to

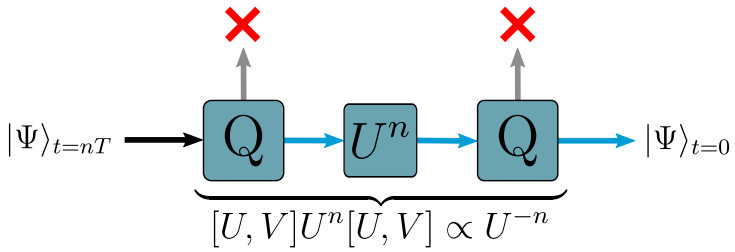


Figure 5.4: **Single-step protocol.** For the rewinding protocol to succeed the commutator between the free and perturbed time evolutions has to be applied twice, and this is done probabilistically by the quantum switches, labelled Q in the figure. In a single-step version of the protocol the execution is terminated if the anti-commutator is applied instead of the commutator.

non-demolition measurements on single photons, that would reveal their spatial mode without collapsing the polarization state or absorbing the photon in question. Without access to such measurements, every set of implemented feed-forward corrections would have to be done in series, so as to not interfere with one another. This would in turn add a significant experimental overhead, particularly in the form of photon loss.

5.2.1 Set of implemented parameters

The goal of the experiment was to certify that the rewinding protocol works independently of the choice of free time evolution U, perturbed time evolution V, input target state $|\Psi\rangle$, as well as the number of time steps to be rewound n. This could be accomplished by sampling these parameters randomly and observing that the inverse time evolution is realised with high fidelity. In our experiment, though, we instead opted for realising a specific, predefined set of parameters. This was done in order to facilitate a comparison with a 'classical' experimenter, given access to the same resources as a quantum experimenter but unable to superpose time evolutions.[3] The implemented pairs of unitaries were picked from the following set:

$$\mathcal{S}_U = \{U_p \in \mathrm{SU}(2) | U_p = e^{-i \arcsin(\alpha)\sigma_z},\ \alpha = \frac{p}{10},\ p = 1, \ldots, 10\} \tag{5.2-3}$$

$$\mathcal{S}_V = \{V_q \in \mathrm{SU}(2) | V_q = i\cos(\theta)\sigma_z + i\sin(\theta)\sigma_y,\ \theta = \frac{q\pi}{11},\ q = 1, \ldots, 10\}. \tag{5.2-4}$$

Here the free time evolution U is defined in terms of the Hamiltonian generating the time evolution, V is a perturbed time evolution, and p and q are parameters that change the degree to which U and V commute. Since p and q can be chosen independently the set contains 100 pairs of unitaries. Not all of these pairs were implemented in the experiment, since the success probability for pairs of unitaries that almost commute would be too small. Instead the only pairs of unitaries with a success probability greater or equal to 1 % were chosen. Let

$$N_c(U,V) = 1 - \left(\frac{||[U,V]||_2}{2}\right)^2, \tag{5.2-5}$$

[3] This classical strategy is detailed in the Supplementary Material of Publication 2 and will not be elaborated on in this chapter.

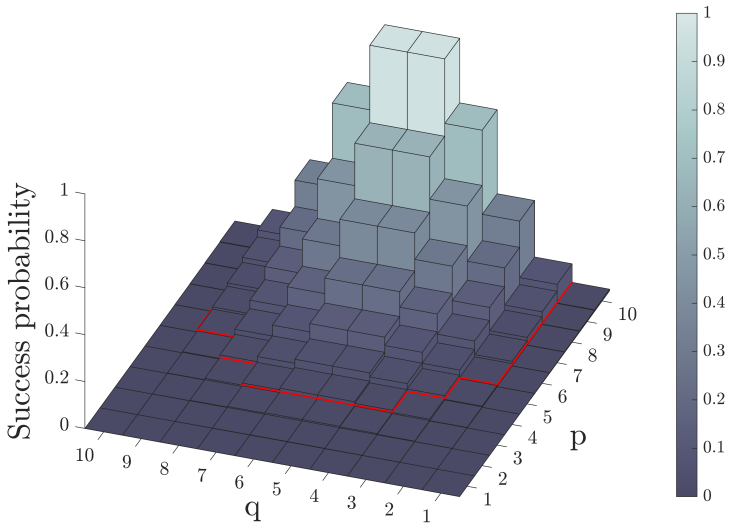


Figure 5.5: **Success probabilities.** The success probability of the protocol depends on the magnitude of the commutator between the free and perturbed time evolutions. In the experiment, these evolutions are parametrised by the integers p and q, and the figure shows the dependence of the success probability on these parameters. The red trace encloses the parameter set for which the success probability is strictly greater than 1 %. The corresponding set of 50 pairs of unitaries were the ones implemented in the experiment.

be the commutativity of two unitaries. Then the set $\mathcal{S}$ of implemented pairs of unitaries can be expressed as

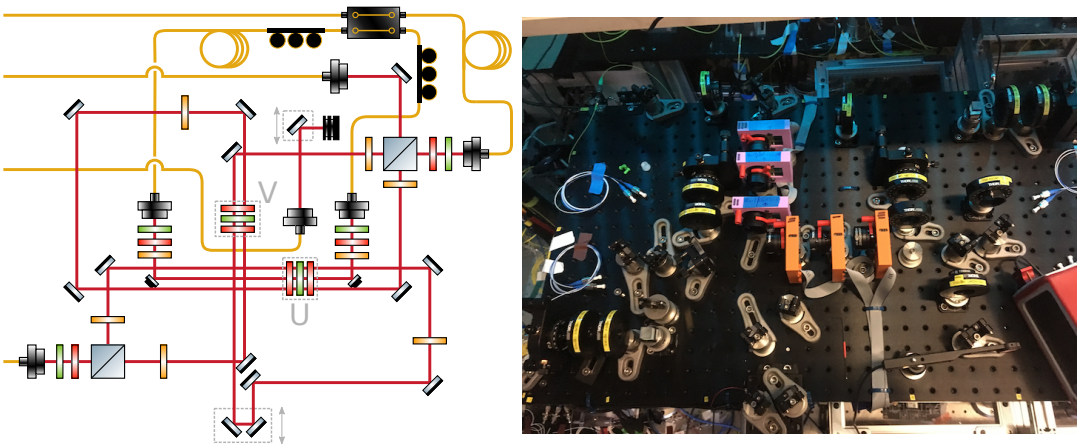
$$\mathcal{S} = \{U \in \mathcal{S}_U, V \in \mathcal{S}_V | N_c(U, V) \leq 0.9\}. \tag{5.2-6}$$

In total there are 50 pairs of unitaries in this set, and their relation to the parameters p and q is illustrated in Fig. 5.5. The input states were chosen such that they formed a tomographically complete set, more specifically:

$$|\Psi\rangle \in \{|H\rangle, |+\rangle, |-\rangle, |R\rangle\}. \tag{5.2-7}$$

A tomographically complete set of input state allows for the process matrix to be reconstructed, which in turn could be used to verify that the process indeed corresponds to the inverse unitary to some power. This will be discussed in Section 5.3.3, however for the publication a different metric for evaluating the performance of the protocol was chosen. Finally, the number of time steps n was chosen to be $n = 1, 2, 3$. These values were chosen in order to limit the total measurement time of the experiment. Higher values of n make the protocol take slightly longer, but more significantly lead to higher photon loss as the total optical path length through the experiment is increased.

Every combination of these parameters was implemented three independent times, giving a total of $50 \cdot 4 \cdot 3 \cdot 3 = 1800$ experimental runs. As mentioned in the beginning of this section, the quantum switch needed to apply the commutator of the unitaries U and V can be realised using a Mach-Zehnder interferometer. This type of implementation was initially pursued for the experiment, because the set $\mathcal{S}_V$ originally contained unitaries on



(a) Setup sketch

(b) Pre-box setup

Figure 5.6: **MZI design.** The initial experimental setup used a Mach-Zehnder interferometer for the quantum switch, due to the assumed need to implement σ_x rotations on the polarization degree of freedom. As shown in (a), both arms of the interferometer pass through both polarization gadgets. An additional path through the U-gadget, implementing the free time evolution, allows this unitary to be applied an additional n times. Two mirrors mounted on a translation stage connected to a piezo-actuator in one arm of the interferometer allow for control over the relative path length and interferometer phase. (b) is a photograph of the first iteration of the setup, which did not include thermal or acoustic insulation. Note that a few optical elements in (a) were not present in (b) at the time of the photo.

the form $V_q = \cos(\theta)\sigma_z + \sin(\theta)\sigma_x$, which, as discussed in Chapter 3, are not reciprocal when implemented using wave plates. A sketch as well as a photo of the initial MZI setup is shown in Fig. 5.6.

The main challenge associated with a Mach-Zehnder based quantum switch is the phase stabilization, since path length fluctuations on the order of nanometers will appreciably change the interference condition. Since the core idea of a quantum switch is that a non-vanishing commutator between two unitaries induces a relative phase shift between the two arms of the interferometer, one cannot easily use active phase stabilization to servo out phase drifts. This is because the locking light would travel through the same polarization optics as the single photons, and therefore acquire the same phase shift. The locking condition would therefore change depending on which polarization unitaries are implemented, and one would like the device to function without requiring any knowledge of the unitaries being applied.

Instead of active stabilization, one therefore has to rely on passive stabilization of the interferometer phase, and periodically reset the phase to prevent it from drifting too far from the desired set point. Due to space constraints, the interferometer for the quantum switch had to be built on an elevated aluminium breadboard. This further exacerbated the problem, because of the poor rigidity of the breadboard which decreased the mechanical stability of the setup.

To improve the passive stability of the bulk interferometer, it was housed inside an acrylic box, coated with Sorbothane and Thinsulate for acoustic and thermal insulation. This was successful in reducing the passive phase drift of the interferometer. A measure-

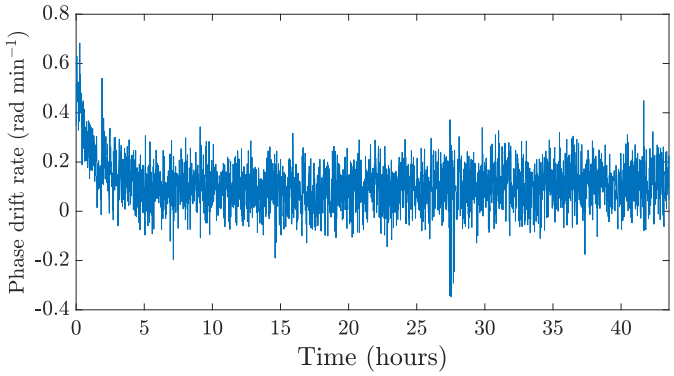
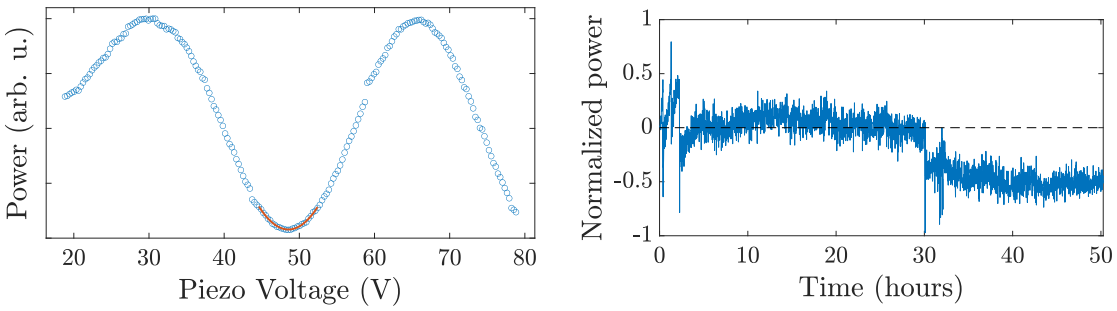


Figure 5.7: **MZI phase drift.** The drift rate of the Mach-Zehnder interferometer was measured as a function of time. The observations were done by scanning the MZI phase, and then resetting it to the quadrature point every 60 seconds while observing the output power in one port of the interferometer. At the quadrature point the relation between phase and power is linear. The drift in each minute long observation window was obtained by first rescaling the power measurement to the min / max of the phase scan, and then doing a linear fit to the time trace of the power. In the plot, a transient approximately of 5 h is present, and this is due to the thermalization of the box within which the interferometer was housed. After this initial transient the drift rate of the interferometer settles at around 97 mrad min^{-1}.

ment of the drift is shown in Fig. 5.7 where it can be see that the drift rate settles at around 97 mrad min^{-1}. This measurement was performed by scanning the phase of the interferometer every minute, and then resetting it to the quadrature point to observe the drift. The observed drift rate, while by itself not insignificant, essentially just quantifies the frequency noise at DC. The mean standard deviation of the phase in the 1-minute measurement intervals was 73 mrad, which indicates that the higher frequency phase noise is non-negligible.

Another factor that affects the phase of the interferometer is the actual phase resetting procedure. For the Mach-Zehnder, this was done by first rotating all the wave plates inside the quantum switch to a neutral position where they collectively implement an identity operation on both interferometer arms, and then moving translation stages to inject laser light into the setup, as well as a powermeter and beam-blocks to protect the single-photon detectors. The voltage of a piezo-electric actuator connected to a mechanical translation stage was then scanned while observing the output power. An example of such a phase scan is shown in Fig. 5.8a. After performing the scan, a quadradic polynomial was fitted to the minimum of the sine curve, and the piezo voltage was set to the minimum of this fit.

In practice, this turned out to work less well than expected. This can be seen in Fig. 5.8b, which shows the mean of the normalized power, taken over the first five seconds after resetting the phase. Ideally, the normalized power should be at exactly 0, since in the figure this corresponds to the quadrature point to which the phase was attempted to be set. However, there are quite substantial deviations from this point, indicating a large phase offset. This could partially be attributed to the hysteresis in the piezo actuator, which would cause the voltage-to-phase map generated in a scan to be different from the voltage-phase relation when resetting the piezo voltage. In an



(a) MZI phase scan

(b) MZI average power

Figure 5.8: **MZI phase resetting**. (a) The phase of the MZI was scanned by applying a ramping voltage to a piezo actuator connected to an optical trombone, consisting of a linear translation stage on which two mirrors were mounted. The piezo was driven using a high-voltage amplifier (Thorlabs MDT693B) and the voltage was ramped digitally in 0.3 V increments, stopping to acquire data after each step. The blue circles represent the measured power, while the red curve shows a quadratic fit to the minimum of the scan, used to determine the location of the minimum in terms of the piezo voltage. The jumps in the data are likely caused either by random phase noise unrelated to the scan, or friction in the translation stage. (b) As described in Fig. 5.7, the interferometer drift was measured by periodically resetting the phase of the MZI to the quadrature point. If the power is normalized to the range $[-1, 1]$, as is the case in this figure, the quadrature point corresponds to the zero crossing of the power. The graph above shows the output power in one interferometer port averaged over the first five seconds of measurement. If the phase is set correctly then the power in this time interval should be close to zero. As can be seen in the figure, the phase reset procedure does not consistently reset the phase to the correct value. This may be caused by a combination of hysteresis effects in the piezo actuator, as well as in the translation stage used to change the path length difference between the interferometer arms.

attempt to counteract this, the voltage was always reset to zero before ramping it to the intended value. It was also observed that rapid changes to the piezo voltage would excite mechanical modes in the spring of the translation stage, which could also give rise to phase offsets.

An alternative phase resetting procedure would have been a lock-and-hold approach, in which feedback control is used to set the interferometer phase to the desired value, after which the control loop is switched off while maintaining the control signal. This could likely have yielded improved performance, however locking an interferometer on a dark fringe is unfavourable due to the even symmetry of the sine curve around the minimum. This approach would furthermore not have alleviated a second problem problem with the periodic phase resetting procedure, which is simply the time overhead. While scanning the phase of the interferometer only took a few seconds, the time spent moving the various translation stages, and in particular rotating the wave plates, contributed a significant amount to the overall measurement time since not all of these operations could be performed in parallel, due to the serial nature of the control electronics.

It was found that the data acquisition time only constituted between one half and one third of the total measurement time, and this, in combination with the aforementioned phase control challenges, led to replacing the Mach-Zehnder with a Sagnac-based quantum switch. While this would not work with σ_x unitaries without the gadgets discussed

in Chapter 3, simply replacing the σ_x terms with σ_y does not change the commutativity of the unitaries in the two sets (5.2-3) and (5.2-4), and enables a common-path implementation.

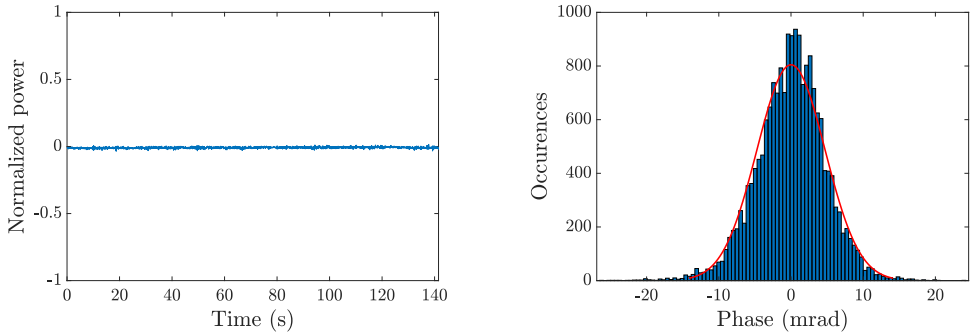
In order to retain the option of going back to the Mach-Zehnder based quantum switch, the Sagnac interferometer was constructed using the existing MZI by inserting additional mirrors that routed the light back to the same beamsplitter and combined the two interferometer arms into one. This led to a more complicated structure of the interferometer than otherwise necessary. The full experimental setup is shown in Fig. 5.10, which can be compared with the initial MZI-based quantum switch in Fig. 5.6. The Sagnac-based quantum switch exhibited no discernible phase drift, but did still have a small amount of phase noise, as shown in Fig. 5.9, however laser intensity noise was not accounted for in the stability measurements. The visibility of the interferometer was checked by changing the interference condition using the wave plates, and observed visibilities were typically in excess of 0.995.

Though not part of the quantum switch itself, the interferometer was also designed to accommodate a second beam path through the polarization gadget implementing the free time evolution U, which was applied to the photon between the two commutators as part of the rewinding protocol. This beam path was vertically offset from the interferometer beam path. In order to ensure that the polarization transformations in the different beam paths were identical, six randomly chosen unitary transformations were implemented using the wave plates and measured along four different spots on the wave plates. Each unitary was implemented only once, to ensure that the variability in the wave-plate angles did not enter into the measurement results. The fidelity of the reconstructed unitaries with respect to the ideal ones were then calculated. The results of these measurements are shown in Fig. 5.11. It can be seen that the difference between individual unitaries is greater than the standard deviation of the different measurements of the same unitaries, indicating that the wave plates are spatially uniform.

In order to reduce the amount of polarizing noise in the interferometer the mirrors were chosen to have a low polarization-dependent reflection coefficient. At 1546 nm, which was the wavelength of the single photons in the experiment, regular gold mirrors exhibit remarkably low polarization dependent loss, with reflection coefficients of 0.966 and 0.965 for the S- and P-polarizations respectively. The downside of this is the relatively high loss of 3.5 % per reflection. For $n = 3$, meaning the photons propagate through the loop containing U three times, the total amount of reflections is 22, which contributes a total loss of $0.965^{22} = 0.46$, which while not dominant, is still non-negligible.

5.2.2 Photon routing

In order to only have a single physical realisation of the time evolution operators U and V it was necessary to actively route the single photons through the experiment. The routing was performed using two high-speed fiber electro-optic switches from Agiltron (Premium NanoSpeed 2x2). The switches have a nominal 3 dB rise time of about 100 ns and repetition rate of 1 MHz. However, in practice the rise time was measured to be significantly higher, and the time needed to reach a sufficiently high contrast was closer to 150 ns (see Fig. 5.12a). Two switches, S1 and S2, were used in the experiment, and



(a) Power time trace (b) Phase histogram

Figure 5.9: **Sagnac phase stability**. A Sagnac interferometer was used for the quantum switch in the experiment in order to increase the passive phase stability, and optical power measurements were performed in order to quantify the noise of the interferometer. (a) shows a time trace of the output power in one port of the interferometer, normalized to the range $[-1, 1]$. Due to the common path geometry, no phase drift can be observed, although a small amount of phase noise is present. (b) quantifies this phase noise and shows a histogram of the phases in (a). The standard deviation of this distribution is 4.7 mrad, and the mean has been offset to zero. The red curve is a Gaussian fit to the data.

programmed to implement the following steps:

1. S1 routes the photon from the state-preparation stage into the quantum switch.
2. S2 sends the photon to U a second time.
3. S2 traps the photon in a loop for n round trips.
4. S2 releases the photon from the loop
5. S1 sends the photon back to the quantum switch.
6. S2 and S1 route the photon to a tomography stage.

The corresponding switch states during these steps shown in Table 5.1.

Switch \ Step	1	2	3	4	5	6
S1	low	low	low	low	high	low
S2	low	low	high	low	low	high

Table 5.1: Switching voltages for the two EO-switches during the different steps of the protocol.

Steps 1 and 2, as well as 4 and 5, can both be combined into single steps, and this logic sequence therefore corresponds to four distinct states. An oscilloscope trace of the FPGA pulse sequence for $n = 2$ is shown in Fig. 5.12b, while a graphical illustration of the four switch states and the photon routing is shown in Fig. 5.13. For simplicity, the switches were connected in such a way that the voltage applied to the switches was

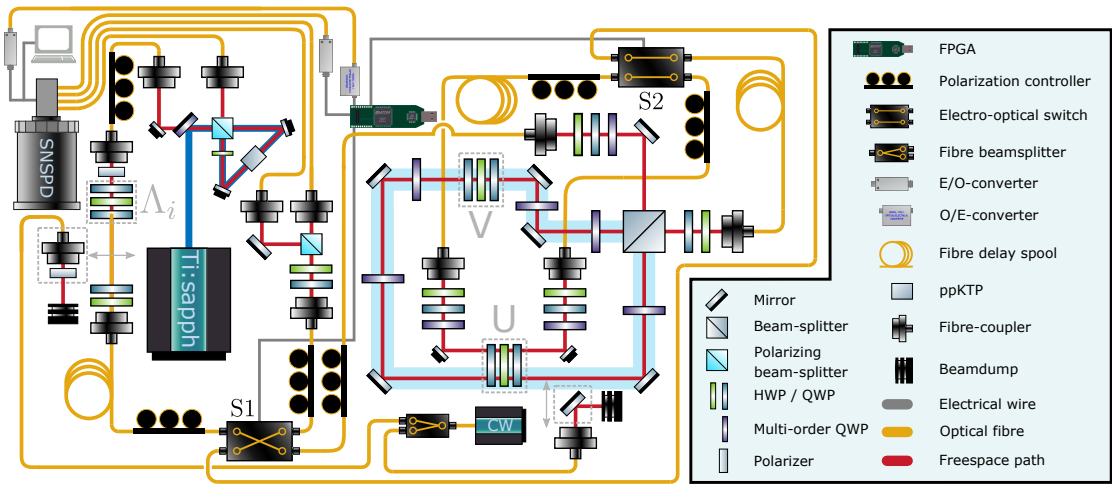


Figure 5.10: **Experimental seutp.** The experimental seutp consisted, broadly speaking, of four different parts. The first of these is the single-photon generation part shown in the top left, in which pairs of single photons are generated in a collinear type-II SPDC process using a ppKTP crystal inside a Sagnac interferometer. Since the experiment did not make use of two-photon entanglement the source was pumped using a single polarization and direction. The second part is the quantum switch in the centre of the figure, for which the photon path is highlighted in blue. This quantum switch is simply a Sagnac interferometer with two polarization gadgets inside of it. The somewhat circuitous photon path was a result of non-intrusively modifying the original MZI design in Fig. 5.6, and few mirrors have been omitted for simplicity. The third principal part of the setup is the feed-forward elements, consisting of digital logic in the form of an FPGA, as well as two ultra-fast electro-optical switches S1 and S2. These routed the photons and allowed them to propagate through the quantum switch as well as the U-gadget multiple times. The final part of the setup is the quantum tomography stage, drawn next to the single-photon source. Here the polarization state of the single photons was measured and later reconstructed. Additionally, the output of a continuous-wave (CW) laser was split using a fiber beamsplitter connected to a coupler and a mirror on motorized translation stages. These were used to inject laser light into the state-preparation stage as well as the Sagnac during the pre-measurement polarization compensation procedure.

low (zero) in the first step. The switching logic was implemented on a Lattice iCEstick FPGA clocked at 200 MHz, and programmed in Verilog.

In order to deterministically route the single photons, it was necessary to synchronise the FPGA with the photon generation. The first step of this was to trigger the FPGA pulse sequence on the detection of a heralding photon. This was done by splitting off the output signal of the corresponding superconducting nanowire single-photon detector (SNSPD) after amplification, but before the time-tagger. This split-off signal was then sent to a comparator (Pulse Research Lab PRL-350TTL) to ensure a consistent and sufficiently strong output signal. Since the cryostat housing the SNSPDs was situated in a different laboratory, separated by approximately 100 m of optical fiber, this electrical signal was converted to an optical signal at 1310 nm using a Highland Technology J720 electrical-to-fiberoptic converter. After propagating back through the fiber separating the laboratories the optical signal was then converted back into an electrical signal using a corresponding fiberoptic-to-electrical converter (Highland Technology J730), and this signal was finally fed to the FPGA. If the FPGA was already executing a pulse sequence

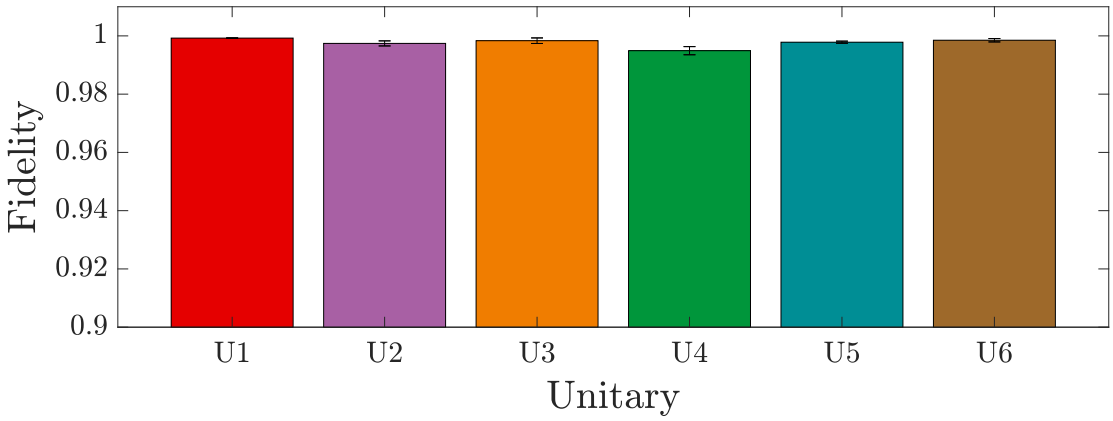
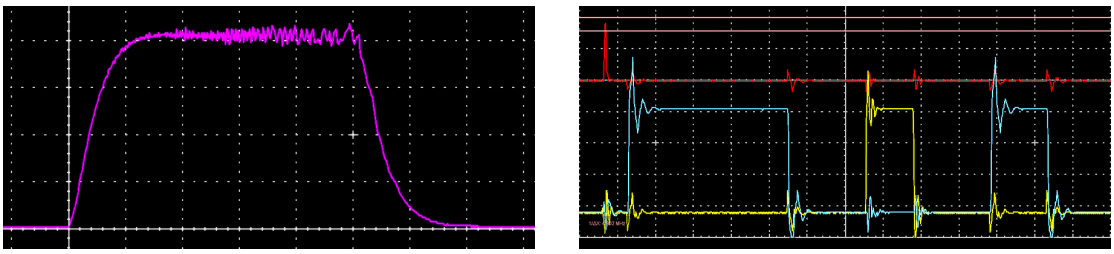


Figure 5.11: **Wave plate characterisation.** The spatial uniformity of the wave plates was verified by performing tomography on six different unitaries at four different beam positions along an axis transverse to the beam. Since the wave-plate angles are different for each unitary the wave plates are in effect sampled on a set of points lying in a plane, not a line. The fidelities are defined as the mean fidelity of the two inputs states, $|H\rangle$ and $|+\rangle$, under the reconstructed unitaries with respect to the theoretical unitary: $\mathcal{F} = 1/2[\langle H|U_{\text{exp}}^{\dagger}U|H\rangle + \langle +|U_{\text{exp}}^{\dagger}U|+\rangle]$. The mean standard variation of this fidelity across the different measurement points was $7 \cdot 10^{-4}$. The error bars show the standard deviation of the fidelities for each unitary.

the trigger signal would be ignored, and otherwise a new sequence would be initiated. Upon triggering, the FPGA would output a trigger signal of its own, which was converted to an optical signal and sent back to the lab housing the SNSPDs and time tagger. There it was detected using an InGaAs photodiode (ThorLabs DET08CFC/M) and recorded using the time tagger.

In order to give time for the heralding photon to travel to the detectors, the resulting trigger signal to propagate back to the laboratory, as well as for the EO-switches to change state, it was necessary to include several fiber delay lines in the experiment. An initial fiber delay of approximately $\tau_1 = 519\,\text{ns}$, corresponding to 106 m of fiber, was used between the state-preparation stage and the quantum switch. This delay fiber is only traversed once by every photon, and served to compensate for part of the trigger delay. A second delay of $\tau_2 = 533\,\text{ns}$ (109 m) was placed between the output of the quantum switch and the second EO-switch. This delay was traversed twice by every photon. The first pass compensated for the remaining trigger delay, while the second pass gave the switch S1 time to change state in order to send the photon to the tomography stage. A final fiber delay of $\tau_3 = 760\,\text{ns}$ (155 m) was placed inside the loop which trapped the photon, the purpose of which was to give S2 time to trap / release the photons.

For the routing to work properly, it was also necessary to synchronise the switch states with the photon as it propagated through the experiment. Since the photon cannot be directly observed as it is propagating through the setup, the right switch timings were found by varying the electrical delays and searching for peaks in two-photon histograms at expected delays.



(a) EO-Switch response (b) FPGA output, $n = 2$

Figure 5.12: **Photon routing oscilloscope time traces.** An oscilloscope was used to characterise the response of the EO-switches, as well as the FPGA pulse sequence. In both figures the horizontal grid division is 100 ns. (a) shows the EO-switch response to a 1 MHz square wave, measured using a CW laser and a fast photodiode. The rise time of the switch is approximately 150 ns, and the fall time is 200 ns. (b) The FPGA pulse sequence is triggered directly by a split-off and amplified detector signal, shown as the red trace in the figure. The yellow and blue traces correspond to the FPGA outputs connected to S1 and S2, respectively, and are shown for the $n = 2$ pulse sequence.

5.2.3 Polarization compensation

Since the target qubit on which the rewinding protocol was applied was encoded in the polarization state of a single photon, it was necessary to remove any undesired polarization rotations induced by the various optical components. Although single-mode optical fibers exhibit fairly low birefringence, they are in most cases sufficiently long that they in effect act as polarization scramblers, and these random polarization transformations need to be undone. The full-width half-max wavelength width of the photons used in the experiment was approximately 1 nm, which is narrow compared to the typical dispersion of SMF-28 fiber. The polarization transformation could therefore be considered unitary.

To undo the transformations of the various fibers in the experiment, a combination of fiber polarization controllers and wave plates were used. A fiber polarization controller typically consists of three small loops of fiber, spooled with a radius chose such that each loop acts as a quarter- or half-wave plate. In this way, a Simon–Mukunda gadget can be realised directly in fiber, and if the retardances of these fiber loops are correct they can be used to implement any unitary on the fiber. In practice, however, the spool radius is fixed, and the retardances are only correct for a single wavelength, which means that the set of implementable unitaries is restricted. Fiber polarization controllers also have an additional problem. To undo an arbitrary unitary U, it is sufficient to implement some unitary W such that:

$$WU\,|H\rangle = |H\rangle \tag{5.2-8}$$

$$WU\,|+\rangle = |+\rangle \tag{5.2-9}$$

The choice of states here is arbitrary, and the only condition is that the two states are not orthogonal or parallel. The conditions (5.2-8) and (5.2-9) have a simple operational interpretation: they correspond to preparing the state $|H\rangle$ ($|+\rangle$) before the device implementing U, and minimizing the transmission through a $|V\rangle\langle V|$ ($|-\rangle\langle -|$) polarizer after

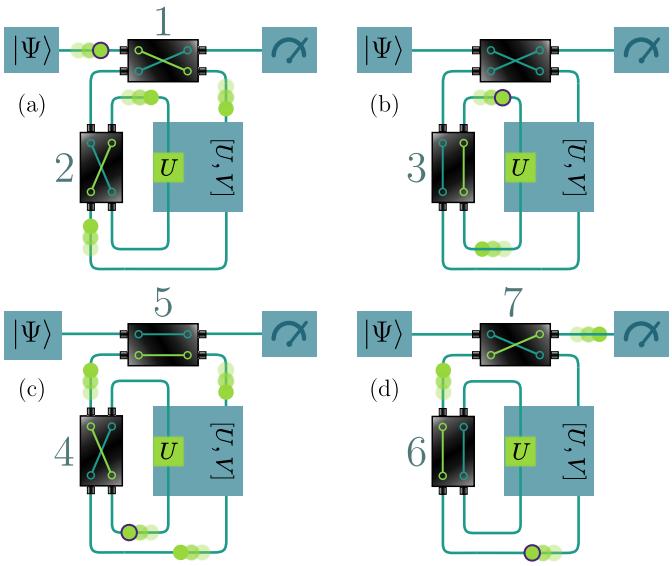


Figure 5.13: **Photon routing.** The switching logic consists of four distinct states. (a) The 'off' state of the EO-switches was chosen such that a photon from the source is sent first through the quantum switch and then to U a second time. (b) The state of the second switch changes to trap the photon in a loop, where it propagates through U an additional $n-1$ times. (c) The two switches change their state to release the photon back from the loop and send it to the quantum SWTICH a second time. (d) The photon is routed to a measurement device. In each diagram the photon is represented by the green circles, with the initial state being marked by a dark contour, while the other circles represent the photon at later times. The bright green lines in the EO-switches indicate the photon path through the switch, and the numbers show the order in which the switches are traversed.

the device. The problem that arises when attempting to simultaneously satisfy these two conditions using a fiber polarization controller is that they are not decoupled. This is because the transformation implemented by the fiber polarization controller, as well as the state it acts on, are uncharacterised.

By using a quarter- and half-wave plate in combination with a fiber polarization controller, however, it becomes possible to decouple the two conditions. This method is illustrated in Fig. 5.14 and Fig. 5.15, and works the following way: an $|H\rangle\langle H|$ polarizer is used to initialize the state $|H\rangle$, after which it propagates through a QWP fixed at 45°, turning $|H\rangle$ into $|L\rangle$. The light then propagates through a HWP, and gets transformed into $|R\rangle$ independently of the angle of the HWP. A fiber polarization controller can then be used to minimize the transmission of the fiber output through a $|V\rangle\langle V|$ polarizer.

Next, a $|+\rangle\langle +|$ polarizer is used to prepare the state $|+\rangle$, which is unaffected by the QWP since it is aligned with its optical axis. The HWP can now be used to map $|+\rangle$ to any other linearly polarized state. Since the fiber polarization controller has been to used to ensure that $|H\rangle$ gets mapped to itself after the fiber, the great circle that is the image of $|+\rangle$ under the HWP therefore gets mapped to the equatorial plane after the fiber. Simply rotating the HWP thus sets the phase between the horizontal and vertical polarization components after the fiber, without affecting their magnitudes and (5.2-9) can be set independently of (5.2-8).

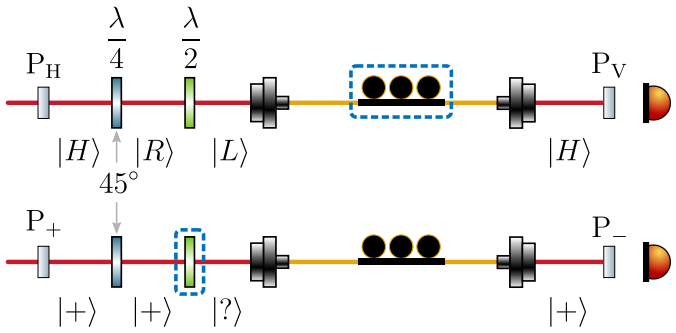
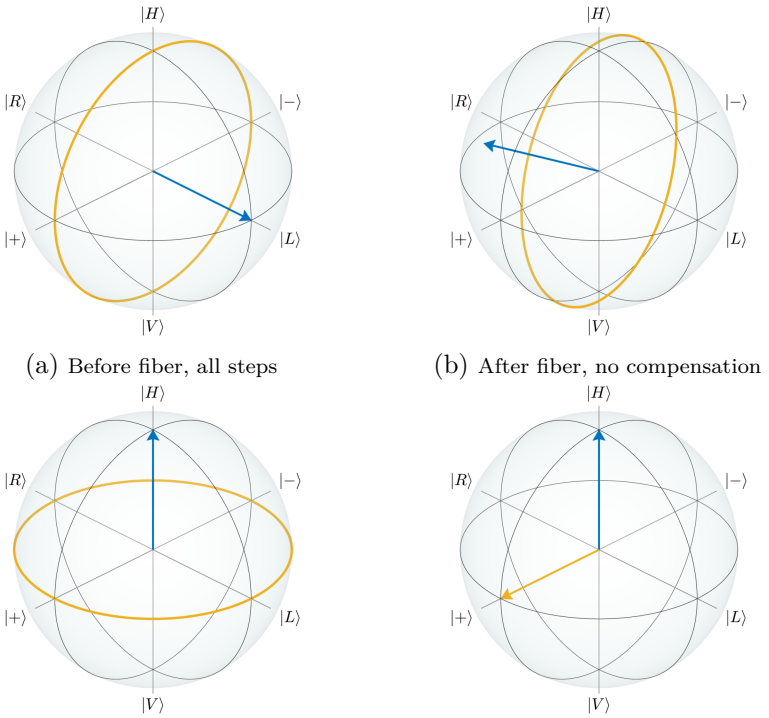


Figure 5.14: **Polarization compensation method.** A sufficient condition to ensure that a fiber implements the identity operation on the polarization degree of freedom is that it maps $|H\rangle \mapsto |H\rangle$ and $|+\rangle \mapsto |+\rangle$. Satisfying this condition using a fiber polarization controller alone can be challenging, because the two mappings cannot be adjusted independently. By using two wave plates in addition to the fiber paddles, however, the two measurement bases can be decoupled [407]. One begins by using a polarizer to prepare the state $|H\rangle$, which then propagates through a QWP fixed at 45°, transforming it to $|R\rangle$. The HWP then, independently of its angle, transforms this state to $|L\rangle$, and the fiber paddles can be used to extinguish the transmission through the P_V-polarizer after the fiber. Next, one prepares the state $|+\rangle$, which is left unchanged by the QWP. The HWP can then be used to transform $|+\rangle$ into any linear state. Since the fiber maps $|L\rangle \mapsto |H\rangle$, the set of linear polarizations is mapped to the equator of the Bloch sphere. This guarantees that there is a linear polarization that gets mapped to $|+\rangle$, and the P_--polarizer can therefore be extinguished by simply rotating the HWP. The evolution of the polarization states during the compensation procedure is visualized in Fig. 5.15.

This decoupling of the two polarization bases requires the quarter- and half-wave plates to be ideal, and also necessitates that the angles of the polarizers are set correctly. Any imperfections will induce a small coupling between the polarizations, meaning that when the HWP is rotated to set the phase, the state $|H\rangle$ experiences a slight rotation at the output of the fiber. In practice it was found that this effect can be mitigated by employing an iterative compensation procedure, in which one alternates attempting to satisfy the two conditions. In the experiment, two rounds of compensation was usually sufficient to ensure a polarization contrast of at least 40 dB for both components.

Due to the photon propagating through some parts of the setup several times, the polarization compensation procedure was more involved than it would be for a straight linear path. As shown in Fig. 5.10, translation stages were used to inject CW laser light at two different points in the setup, in order to be able to perform the necessary polarization measurements. The full fiber compensation procedure is illustrated in Fig. 5.16.

While standard single-mode fiber such as SMF-28 is not strongly dispersive, the same was not true for the fiber based active EO-switches. A consequence of this dispersion is that the random polarization transformations applied to the light as it propagates through a switch is also wavelength dependent. A measurement dispersive nature of the polarization rotations is shown in Fig. 5.17. This means that to effectively compensate these polarization transformations using only unitary transformations, the laser light used in the compensation procedure has to closely match the wavelength of the single photons. Since this effect was unexpectedly large it was not initially tested or accounted for. As a result, the polarization compensation in the measurement runs for $n = 1, 2$



(a) Before fiber, all steps

(b) After fiber, no compensation

(c) After fiber, paddle compensation

(d) After fiber, full compensation

Figure 5.15: **State evolution during polarization compensation.** The polarization compensation method uses two input states, $|H\rangle$ and $|+\rangle$, which are shown in blue and yellow, respectively. In (a) the two states are drawn immediately before the fiber to be compensated, after they have passed through the two wave plates (cf. Fig. 5.14). The input state $|+\rangle$ is drawn as a great circle representing all the possible states it could have been mapped to by the HWP. (b) shows the states in (a) after the fiber, which performs a random rotation. In (c) the fiber polarization controller is used to map $|H\rangle \mapsto |H\rangle$, and this leaves $|+\rangle$ somewhere on the equator. (d) Setting the appropriate HWP angle ensures $|+\rangle \mapsto |+\rangle$.

was done with a laser wavelength of approximately 1549 nm, which differs from the single-photon wavelength by a few nm.

Before the measurements for $n = 3$ this effect was discovered, and the polarization compensation was instead done with a tunable laser. Care was then taken to ensure that the polarization contrast was maximized for the laser and single photons simultaneously. We believe that this accounts for the consistently higher fidelities observed for $n = 3$, even though the higher loss and longer optical path make the measurements more challenging than for $n = 1, 2$.

In addition to the compensation of the rotations induced by the fibers, it was also necessary to undo the polarization-dependent phase shifts caused by the mirrors inside the Sagnac interferometer. Since these phase shifts have a preferred basis, defined by the orientation of the mirror, it is sufficient to realise polarization phase shifts in this particular basis to perform the compensation. In the experiment this was achieved using multi-order quarter-wave plates mounted on z-axis rotation mounts. Rotating the wave

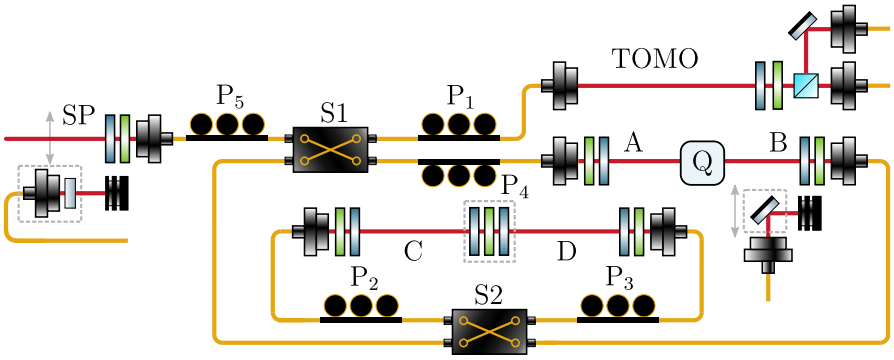


Figure 5.16: **Polarization compensation procedure.** The polarization compensation is done in five steps between the various points indicated above. In the figure, Q represents the quantum switch. The path S2 → C → D → S2 is the loop containing U, and has been drawn separately from Q for clarity. The procedure begins by using a translation stage to inject laser light at the point B, and into the output coupler of the quantum switch. Using the fiber polarization controller, or paddle, P_1 as well as the wave plates at B the path B → TOMO can be compensated. This guarantees that the polarization transformation between these two points is the identity operation. Next, the pair of wave plates at C are used together with P_2 to compensate the path B → C. Then the same procedure is applied using P_3 on the path D → TOMO. Since B → C, B → TOMO and D → TOMO are now compensated this implies that D → C is also compensated. At this point the light source is switched to the state-preparation stage (SP), and the path SP → TOMO is compensated using P_5 and the wave plates in SP. Finally, SP → A is compensated using P_4 and the wave plates at A. These conditions now imply that D → A is also compensated and the experiment can proceed.

plate changes the effective birefringence and can be used to tune the phase between two linear polarization components.

While this approach does not guarantee full 2π tunability of the phase unless the birefrigent material is sufficiently thick, it turned out to work quite well in practice, especially since the sign of phase shift can be change by rotating the wave plate by 90°. A total of five multi-order wave plates were used for polarization compensation inside the Sagnac, with some compensating for a single mirror and some for multiple mirrors. An additional two were used inside the U-loop. It was observed that the multi-order QWPs only compensating a single mirror were able to reach higher polarization extinction ratios, regularly in excess of 50 dB. This was likely caused by different mirrors giving phase shifts along slightly different axes. The use of additional multi-order QWPs to compensate each mirror separately was however prevented by space constraints.

5.2.4 Depolarizing noise

The polarization compensation procedure described in the previous section requires the polarization transformations of the various components in the setup to be unitary. This is a valid assumption as long as the total delay between the polarization components is small compared to the coherence length of the light, which for our approximately 3 ps long photons is the case even for several km of fiber, due to the low polarization mode dispersion (PMD) exhibited by single-mode telecom fibers. The same does not necessarily hold true for all the components in the setup though. In particular, the electro-optic

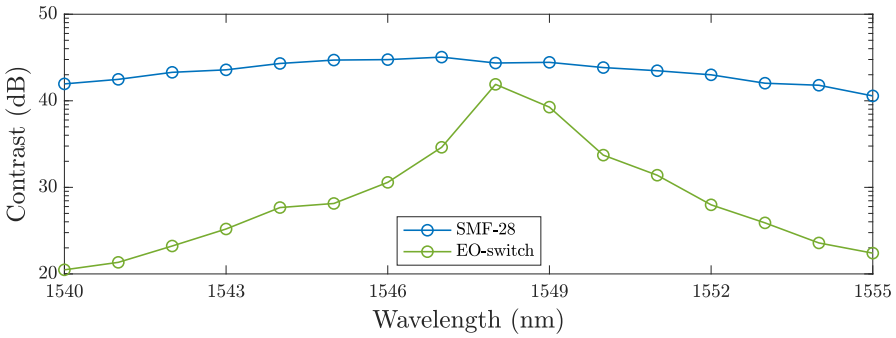


Figure 5.17: **Wavelength-dependent polarization contrast.** Measurements of polarization contrast as a function of wavelength were performed for two different optical components: a standard SMF-28 fiber (blue), approximately 10 m in length, and a high-speed electro-optic fiber switch (green). The data was taken by performing polarization compensation in one basis at a single wavelength, in this case 1548 nm, and then scanning the laser wavelength to see how the extinction for the compensated polarization changes. As can be seen in the figure the SMF-28 fiber exhibits only negligible wavelength dependence, and maintains a polarization contrast of 10,000:1 over 15 nm. The EO-switch, on the other hand, exhibits a strongly wavelength-dependent behaviour, with a drop in contrast by as much as a factor of five in a single nm step. For this reason it is critical that the wavelength of the laser used in the polarization compensation step of the experiment closely matches that of the single photons.

switches used in the experiment exhibit a non-negligible PMD. This is not due to the inherent birefringence of the $LiNbO_3$ crystal used since, according to the manufacturer, the light propagates along the optical axis of the crystal, meaning that both polarization components see the ordinary refractive index. Instead, the PMD is likely caused by the waveguide structure used in the device. A rectangular ridge waveguide, for example, will typically have polarization-dependent propagation constants.

The manufacturer-specified PMD for the EO-switches was 0.1 to 0.3 ps, and this number should be understood as the maximum delay between two polarization components propagating through the device. For $n = 3$ the photons travel through the EO-switches a total of seven times, and the total delay between the polarization components can therefore be on the order of the coherence length of the photons. In practice the polarization state of a photon as it enters an EO-switch is essentially random, since the polarization compensation procedure does not enforce any condition inside the switches. The PMD of the EO-switchess therefore acts as a random walk, instead of coherently building up in a certain basis.

In addition to the EO-switches the multi-order quarter-wave plate also introduce some amount of depolarizing noise. Each quarter-wave plate delays one polarization component by 5.25 wavelengths, corresponding to a PMD of approximately 0.027 ps, which is roughly ten times less than the EO-switches. However, for $n = 7$ the photons propagate through a total of 16 multi-order QWPs, several of which will give phase shifts that add, and the wave plates may therefore contribute a small amount to the overall depolarizing noise. Finally, the wavelength-dependent polarization rotations in the EO-switches discussed in the previous section can also lead to depolarizing noise. However due to the relatively narrow spectral bandwidth of the photons this is likely a

tertiary effect.

5.3 Data acquisition and analysis

For each value of n the experiment was aligned and the polarization compensation was performed. Data was then acquired for the different values of p and q in descending order, starting with $(p, q) = (10, 9)$, then $(10, 8)$ and so on. For each value of p, q the four different input states $|H\rangle$, $|+\rangle$, $|-\rangle$, $|R\rangle$ were prepared in sequence, and quantum state tomography was performed on the resulting output state. After performing the tomography for every combination of input state and p, q the whole process was repeated two more times, for a total of three repetitions. These were carried out without interruption or re-alignment of the setup. The only interruptions came from the need to re-condense the helium inside the closed-cycle cryostat once per day, limiting to the daily measurement time to between 12 and 16 hours. One additional 10 day interruption occurred for $n = 1$ due to an air-conditioning failure, however the measurements were successfully resumed after the lab was once again able to be cooled.

The quantum state tomography was implemented using two motorized wave plates and a polarizing beamsplitter, which meant that both eigenvalues of a qubit observable could be measured simultaneously, and three measurement settings would have sufficed to perform complete tomography. In practice, however, the fiber-coupled reflection port of the PBS tended to have lower detection efficiency, and additionally had a worse extinction ratio. These differences need to be accounted for in order to faithfully reconstruct the input state.

One can alternatively implement an additional three measurement settings, corresponding to the negated observables for which the eigenvalues are swapped. This cancels any imbalance between the two ports, since each eigenvalue is measured in reflection half the time, and in transmission the other half of the time. In the experiment this latter approach was used due to its simplicity, despite the downside of a slightly larger experimental overhead. A total of six measurement settings were therefore used in each tomography.

As can be inferred from the preceding discussion the total number of measurement settings was quite high. More concretely, considering three runs, three different values of n, 50 different pairs of unitaries, four different input states and six tomography settings, the total number of measurement settings was $3 \cdot 3 \cdot 4 \cdot 6 \cdot 50 = 10800$. Since the success probability, and therefore final detected photon rate, varies greatly with p and q, the choice was made to tailor the integration time for each setting to the expected photon rate. This helped to distribute the total measurement time more efficiently, and ensure that the signal-to-noise ratio for the low count rate settings was sufficient, while avoiding excess integration past the point of diminishing returns. In order to mitigate the effect of long term drifts, as well as ensuring that the measurement time was not dominated by the overhead of rotating wave plates, the integration time for each setting was constrained to be between 45 and 240 seconds.

Due to the relatively long photon flight time through the experiment, over 4 µs for $n = 3$, a lot of trigger events were ignored by the FPGA. This is because the FPGA only

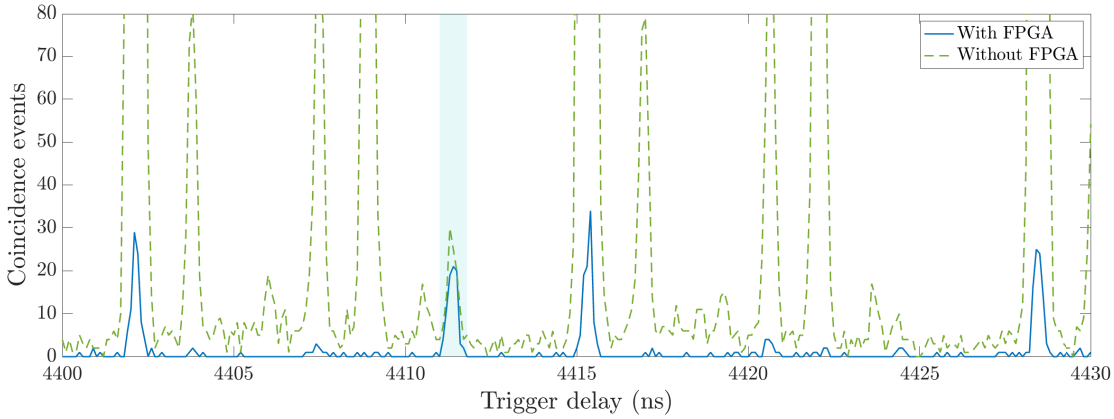
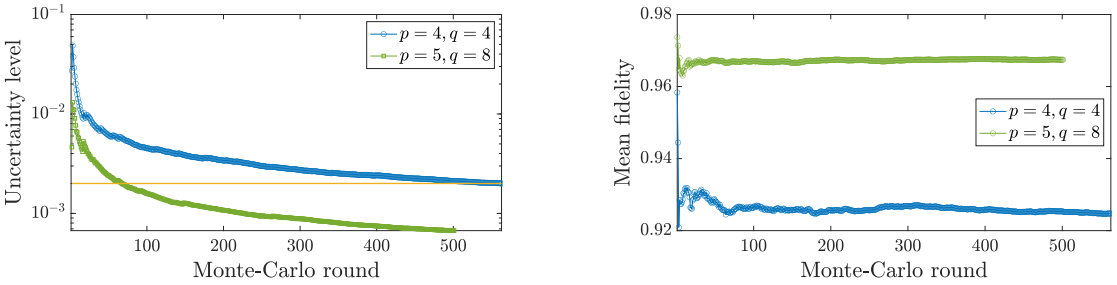


Figure 5.18: **Effect of FPGA triggering on background.** The use of active switches in the experiment meant that there were multiple paths from the photon source to the detector. This in turn led to a background of single photon detection events in the tomography stage. While most of these events would be filtered out by their arrival time when conditioning the signal on the heralding photon, an idler photon from one pair generation event can arrive within the coincidence window defined by a heralding photon from a second pair. Due to the use of a pulsed pump laser the various background peaks in the coincidence histogram were periodic with a period of 13.2 ns defined by the laser repetition rate of 76 MHz. As a first step, these noise peaks were offset from the signal peak by adding a short fiber delay to the U-loop, that differentially delayed the signal and the background. This was done separately for each value of n. In order to further filter the noise the two-fold coincidence events were also conditioned on an FPGA trigger signal, thereby filtering out any heralding photons that were not used to start a new pulse sequence. As can be seen in the figure, this results in a strong reduction of the background events. The solid blue curve shows the coincidence histogram after conditioning on the FPGA signal, and the dashed green line shows the signal before. The light blue highlight indicates the region of the histogram containing the experimental signal. While the filtering does remove a small amount of real signal, since some ignored heralding events will have produced photons close enough in time to the real heralding signal to pass through the EO-switches in the correct order, this step was nevertheless found to improve the signal-to-noise ratio. The bin size in the histogram is 0.1 ns.

starts a new pulse sequence after the previous one has been completed. The corresponding idler (input) photons for these events, however, may still contribute background noise. This is possible due to the multiple photon paths through the experimental setup, which allow photons to arrive with various delays. This was exacerbated by the relatively low extinction ratio of the EO-switches of around 13 dB, which contributed to a fairly high rate of photons going straight through S1 to the detectors, even when this switch was set to route photons from the source to the quantum switch. To mitigate this effect the pump power of the SPDC source was decreased to lower the rate of trigger photons while still keeping the rate of actual trigger events by the FPGA nearly constant. This additionally allowed the trigger detector to be biased higher without latching, thereby increasing its detection efficiency.

An additional consequence of the long photon path through the experiment was that the time tagger could not filter coincidence events in real time, due to the trigger delay being too long. All the coincidence analysis was therefore done on the raw time tags in software. With integration times of upwards of four minutes, and a heralding



(a) $n = 1$ (b) $n = 2$

Figure 5.19: **Convergence of the Monte-Carlo simulation.** The convergence rate of the fidelity estimation varies significantly between measurements. This is illustrated in (a), where the uncertainty in the estimate of the mean fidelity is shown as a function of the number of iterations in the Monte-Carlo simulation. The green curve reaches the halting condition for the confidence, indicated by the yellow line, in less than 70 iterations, but the simulation continues until the minimum number of iterations. On the other hand, the blue curve converges much slower and takes 565 iterations to reach an uncertainty below $2 \cdot 10^{-3}$. (b) shows the corresponding evolution of the mean fidelity during the simulation. In this case, the faster converging measurement also has a higher fidelity, as well as higher photon counts, but these two factors were only found to be weakly correlated with the convergence rate.

photon rate of approximately 1 MHz, the amount of data collected was fairly high. As an example, a 1 MHz rate of time tags consisting of two 32-bit integers, one for the detection channel and one for the actual time tag, results in a raw data rate of about 1 GB per minute, due to the raw text output of the time tagger. For this reason the data for each measurement setting was converted to a binary format and compressed as soon as it was fully acquired.

The large amounts of trigger events also made finding coincidence events in the time tags a somewhat computationally intensive process, and a lot of effort, mostly by Peter Schiansky, was spent on optimizing this procedure. In the end it was found that the fastest way to perform the coincidence analysis was using standard set-intersection algorithms, by first binning the time tags into 0.1 ns bins. Performing the set intersection with a range of relative offsets between two channels then generated a coincidence histogram, which could be integrated to yield the desired coincidence window. By distributing this analysis across several different computers in the lab, it was possible to perform it in real time, so that the data was being processed at the same rate at which it was being acquired.

To increase the signal-to-noise ratio the coincidence analysis also included the trigger signal from the FPGA, thereby filtering out any heralding photons that were unused. This, effectively three-fold coincidence analysis, was done as a second set intersection between the initial two-fold coincidence events and the FPGA signal. The analysis was complicated slightly by the fact that the FPGA signal had a significantly longer temporal width than the detector output signal. This was caused by the 200 MHz clock of the board, limiting the pulse width to 10 ns. The trigger signal furthermore had a non-negligible jitter due to the digital nature of the device timing. Conditioning the two-photon coincidences on the FPGA trigger signal nevertheless significantly reduced

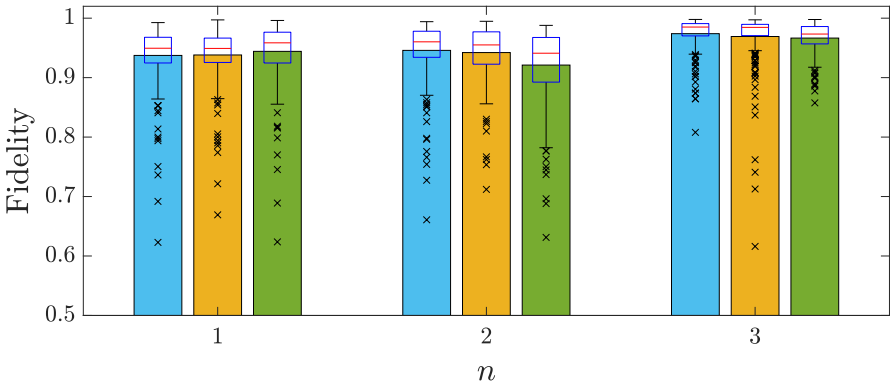


Figure 5.20: **Average fidelities.** The bars show the average fidelity for each of the three runs for each value of n. Each sub-group of bars is ordered according to the measurement order. While a slight downwards trend in the fidelity for later rounds can be seen for $n = 2, 3$, these runs nevertheless exhibit high repeatability despite the setup being operated continuously for as long as two weeks. The highest average fidelity is seen for $n = 3$, which can be attributed to the higher polarization contrast throughout the setup during the measurement run. The red lines and blue boxes of the superimposed box plot indicate the median as well as the first and third quartiles, respectively. The black crosses show the outliers, and the whiskers indicate the highest and lowest values within 1.5 times the interquartile range.

the background noise, as illustrated in Fig. 5.18.

An additional noise reduction method that was explored was averaging the background signal at coincidence windows separated by 13.2 ns, and then drawing Poisson distributed samples from this mean background to estimate the background noise in the signal. This step had a significant effect on the signal to noise prior to conditioning the coincidences on the FPGA trigger signal, since most of the noise was either periodic or uniform. As discussed above, after employing the FPGA condition most of the periodic noise was removed and the influence of the background sampling was thereby reduced as the mean background at the sampled points was very small.

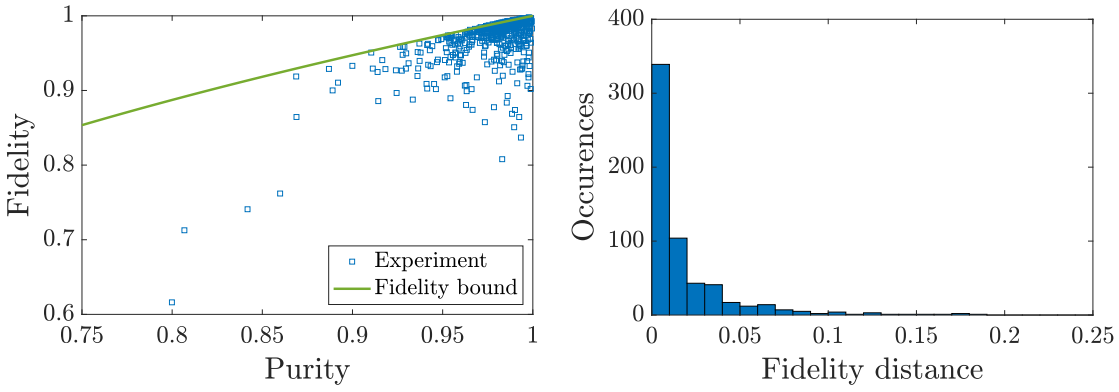
Finally, the coincidence windows were also chosen to maximize the signal-to-noise ratio. For $n = 1$ the coincidence window was 0.6 ns, while for $n = 2, 3$ it was a mere 0.3 ns due to the need to offset the decreased signal-to-noise ratio resulting from the lower photon count rate.

5.3.1 State fidelity

The performance metric chosen for the evaluation of the protocol's performance was the fidelity of the output states with respect to the ideal ones:

$$\mathcal{F} = \mathrm{Tr}\left[\sqrt{\sqrt{\rho}\rho_{\mathrm{exp}}\sqrt{\rho}}\right], \tag{5.3-1}$$

where $\rho = U^{-n}\left|\Psi_i\right\rangle\!\left\langle\Psi_i\right|\left(U^{-n}\right)^{\dagger}$, and $\left|\Psi_i\right\rangle \in \{\left|H\right\rangle, \left|+\right\rangle, \left|-\right\rangle, \left|R\right\rangle\}$ is the input state. The experimental density matrices were reconstructed using a maximum likelihood method, and the state fidelities and their uncertainties were estimated using a Monte-Carlo simulation. More concretely, this was done the following way:



(a) Fidelity bound (b) Distance from bound

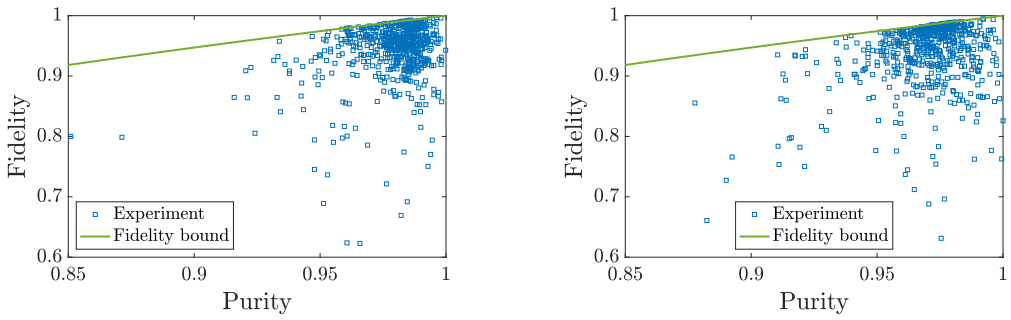
Figure 5.21: **State fidelity and purity for $n = 3$.** Depolarizing noise in the experiment necessarily decreases the fidelity of the output state, as the purity of a state imposes a bound on the maximum state fidelity. This bound is indicated by the green curve in (a), while the experimentally obtained fidelities are shown in blue. The data points that lie on the green curve correspond to states that have the highest possible fidelity for their purity, meaning that the observed fidelity is caused solely by depolarizing noise and not imperfect unitary transformations. Conversely, the data points lying on the right-hand vertical axis correspond, in principle, to pure states that were not subjected to the correct unitary transformations. In practice many of the purities very close to 1 were likely a result of noise. It can be seen that the data points are, by and large, clumped close to the green curve, indicating that the depolarizing noise was a significant contributor to the observed fidelities. This is quantified in (b) by a histogram that shows the distribution of the difference in fidelity between the measured value and the bound defined by the purity. This distribution is peaked around 0, indicating that many of the observed fidelities were limited by the state purity.

1. The observed photon counts for each measurement outcome are used to define a set of Poisson distributions, from which new photon counts are randomly drawn.
2. A maximum likelihood estimation is used to fit a density matrix to the photon counts.
3. The fidelity (5.3-1) is calculated, and appended to a list $\mathcal{F}_{\mathrm{L}}$.
4. The procedure is repeated for at least $r = 500$ rounds, and until the condition $1.96\dfrac{\sigma(\mathcal{F}_{\mathrm{L}})}{r} < 2 \cdot 10^{-3}$ is satisfied.

Here $\sigma(\mathcal{F}_{\mathrm{L}})$ is the standard deviation of the generated fidelities, and this is used as the uncertainty estimate of the mean fidelity. As discussed in Section 1.2.5, the halting condition above is equivalent to lower bounding the probability P of the estimated mean fidelity $\overline{\mathcal{F}_{\mathrm{L}}}$ being wrong by more than $1/500$ by $P < 0.05$:

$$P(|\overline{\mathcal{F}_{\mathrm{L}}} - \overline{\mathcal{F}}| > 1/500) < 0.05. \tag{5.3-2}$$

Stated more simply, it implies a 95 % confidence that the estimated mean is within $1/500$ of the true mean $\overline{\mathcal{F}}$. Here the term 'true mean' refers to the mean of the distribution of fidelities defined by the Poissonian sampling of the photon counts, which is not



(a) $n = 1$ (b) $n = 2$

Figure 5.22: **State fidelity and purity for $\mathbf{n = 1, 2}$.** In contrast to $n = 3$, the data for $n = 1$ and $n = 2$ exhibit deviations from the ideal state fidelity that are frequently much larger than the fidelity bound imposed by the state purity. This was likely a result of the lower polarization contrast in the setup during these measurements, leading to unwanted polarization rotations.

necessarily the true fidelity. The purpose of a minimum number of iterations as part of the halting condition is to ensure that the central limit theorem can be invoked. Two examples of the convergence of Monte-Carlo simulation runs are shown in Fig. 5.19.

The mean fidelities for each value of n and each of the three runs is and visualized in Fig. 5.20. It can be seen that $n = 2$ produces a slightly lower fidelity than $n = 1$, which is not unexpected due to the longer photon path introducing more noise and photon loss. The runs for $n = 3$, despite using an even longer photon path, nevertheless exhibit higher average fidelities. This can be attributed to the improved polarization compensation enabled by tuning the CW laser used for the compensation to the single-photon wavelength, as discussed in Section 5.2.3.

As will be explored in the following section, the reconstructed states were, in most cases, not pure states, and this in turn affected their fidelities. In the special case of two qubits, ρ_1 and ρ_2, one of which is in a pure state, the fidelity between the two states can be expressed in simpler way than (5.3-1):

$$\mathcal{F} = \mathrm{Tr}[\rho_1 \rho_2]. \tag{5.3-3}$$

Suppose that ρ_2 is a depolarized version of ρ_1:

$$\rho_2 = \lambda \rho_1 + \frac{I}{2}(1 - \lambda), \qquad \lambda \in [0, 1]. \tag{5.3-4}$$

The purity $\mathcal{P}$ of this state is:

$$\mathcal{P}(\rho_2) = \frac{1 + \lambda^2}{2}, \tag{5.3-5}$$

and, using (5.3-3), it is straightforward to show that:

$$\mathcal{F}(\rho_1, \rho_2) = \frac{\lambda + 1}{2}. \tag{5.3-6}$$

Since ρ_1 and ρ_2 correspond to parallel vectors on the Bloch sphere, the fidelity $\mathcal{F}(\rho_1, \rho_2)$ is an upper bound for the fidelity between ρ_1 and any state with a purity of $(1 + \lambda^2)/2$. By inverting (5.3-5):

$$\lambda = \sqrt{2\mathcal{P} - 1}, \tag{5.3-7}$$

this bound can be expressed as

$$\mathcal{F} \leq \frac{\sqrt{2\mathcal{P}-1}+1}{2}. \tag{5.3-8}$$

In Fig. 5.21a and Fig. 5.22 the experimental fidelities are visualized with respect to this bound. For $n = 3$ many of the data points lie close to the bound. This is further quantified in Fig. 5.21b, which shows the distribution of the distances from the bound. The fact that the data is close to the fidelity bound indicates that depolarizing noise had a dominant role to play in decreasing the output state fidelities. For $n = 1, 2$, on the other hand, the data is not consistently close to the bound, meaning that other experimental imperfections, such as polarization rotations in the optical fibers, contributed to the degradation of the fidelities.

Since the input states in the experiment were initialized with a high-contrast linear polariser their purity was essentially unity. Despite this the purity of the reconstructed states at the output of the experiment was less than 1, falling as low as 0.8 for some measurement settings. This drop in purity could have been caused by several different effects. One possibly contributing effect is the depolarizing noise described in Section 5.2.4, which will be discussed momentarily. Such noise causes the actual purity of the polarization state to drop, but there are also effects which only cause the apparent purity to drop by introducing noise in the measurement. This can be in the form of background detection events, or noise arising from the finite counting statistics. Additionally, time-dependent effects can cause the state to change during the measurement itself, and averaging over these different states will also lead to an effective impurity.

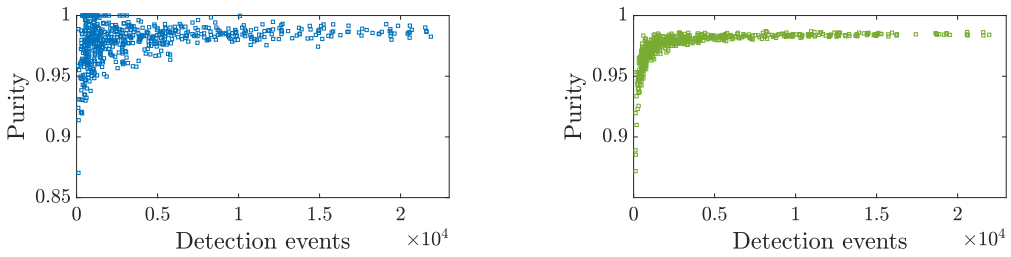
A scatter plot of the observed state purities for $n = 1$ is shown in Fig. 5.23a. It can be seen that the purities decrease with lower photon count rates. Additionally the variance of the purities increases. As shown in Fig. 5.23b, Poissonian noise arising from a combination of the finite counting statistics and a background signal explains the decrease in purity with decreasing photon counts, but does not explain the great increase in the variance. The distribution of state purities is therefore caused by a different noise source.

5.3.2 State purity

The results of an alternative noise model is shown in Fig. 5.24 and Fig. 5.25. This model accurately reproduces both the distribution of the purities, shown in (a) and (b) of the two figures, and the behaviour of the purity as a function of the mean photon counts, shown in (c) and (d). The model consist of simply adding normal-distributed noise to the photon counts:

$$P(N) = \frac{1}{\sigma\sqrt{\pi}}\exp\left[-\frac{1}{2}\left(\frac{N-\mu}{\sigma}\right)^2\right]. \tag{5.3-9}$$

The simulations shown in Fig. 5.24 and Fig. 5.25 were done by iterating through the measurements and using the mean photon counts to define a Poissonian distribution, from which a new mean value was drawn. This was then used to simulate the measurement process by drawing Poisson-distributed samples from a fixed density matrix, using



(a) Experiment (b) Simulation

Figure 5.23: **State purities and Poissonian noise, n = 1.** (a) shows the purities of the reconstructed quantum states for $n = 1$ as a function of the average number of photon counts in each basis. It can be seen that the state purity drops as fewer photons are collected, but the spread of the purities also increases. (b) shows a model of the tomography that accounts for the Poissonian noise in the photon counting, as well as an additional Poisson-distributed background that acts as depolarizing noise. While this model predicts a drop in the state purity as a function of decreasing photon counts, it does not accurately describe the widening of the distribution of purities. The tomography was simulated for a slightly impure state, which is why the purity does not asymptotically approach one.

the new mean value. Additional noise was then sampled from the distribution above and added to the six measurement outcomes. The simulations of both $n = 2$ and $n = 3$ used the same noise distribution, more specifically one with a mean of zero ($\mu = 0$), and a variance with a term proportional to the observed photon number plus a constant offset:

$$\sigma = 8 + \frac{N}{300}. \tag{5.3-10}$$

This noise model is qualitatively different from the background-noise model shown in Fig. 5.23 since the noise can be negative. Physically, this can be caused by a several different effects. For example, errors in the wave-plate angles in the tomography can cause the observed counts for a measurement outcome to drop, although this would not explain the observed results in the experiment since the precision and repeatability of the wave plate motors was sufficiently high for this to be a negligible effect. More likely causes are various forms of time-dependent noise. Since the integration time for certain measurement settings was upwards of four minutes, there was time for slow coupling drifts to change the efficiency of the setup, which introduces a variance in the photon counts with respect to the measurement settings. Slow time dependent polarization noise that rotates the output state could also affect the measurement in a similar way.

This type of time dependent noise justifies the term in (5.3-10) that is proportional to the photon rate. The constant term is needed to account for the bigger variance at lower photon counts. This could partially be explained by the measurement time being longer for these settings, making them more susceptible to longer term drifts. In this scenario the dependence of the count rate on the noise would actually be nonlinear, however the simple constant offset is sufficient to reproduce the behaviour. Ultimately, however, a more quantitative noise model would be needed to elucidate exactly which process caused the observed purity distribution.

While the simulations in Fig. 5.24 and Fig. 5.25 show good agreement with the experimental data, the correct asymptotic behaviour is achieved by using an impure

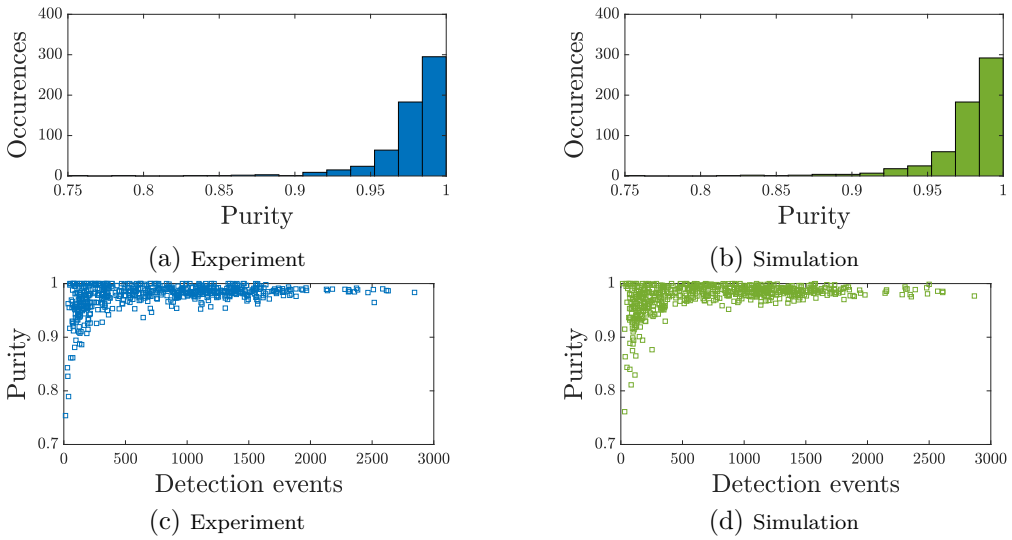


Figure 5.24: **Observed and simulated state purities, n = 3.** The measured final state purities for $n = 3$ (left) were compared with a simulation of the quantum state tomography using the noise model described in the main text (right). Good agreement between the simulation and experiment can be seen. (a) and (b) show histograms of the distribution of purities, while (c) and (d) are scatter plots of the state purity as a function of the average photon counts per observable. In the experimental data the purity converges to a value around 0.99 as the number of detection events increases, suggesting the presence of depolarizing noise sources. In the simulation this behaviour is reproduced by assuming a slightly impure input state.

input state, and the input purity is a free parameter in the simulation. In the actual experiment this depolarization may have been caused by the polarization mode dispersion in the EO-switches discussed in Section 5.2.4. A Monte-Carlo simulation was used to quantify the effect of this noise. The simulation models the photon wavepacket as a superposition of discrete time bins, and then applies unitary operations as well as polarization-dependent time displacements on the bins. The purity of the output state is found by simulating an idealized quantum state tomography that traces out the arrival time of the photon. In each run of the simulation different fiber polarization unitaries are generated to account for the random orientation of the polarization state as it propagates through the EO-switches. An example of a simulated wavepacket is shown in Fig. 5.26.

Naturally, the final purity depends on the random polarization unitaries inside the fibers connecting the EO-switches to the free-space parts of the experiment. These random rotations define the walk-off axis, and can cause the walk-off to add or cancel. The unitaries U_p and V_q also rotate the polarization state, which too could indirectly influence the purity. However, the choice of q turns out to not matter, since this parameter only affects the magnitude of the commutator $[U_p, V_q]$, and does not change the fact that it is proportional to σ_x. This means that whenever a photon successfully exits in the commutator port of the quantum switch the same transformation will always have been applied. The choice of p, though, does affect the orientation of the polarization state inside the EO-switches since unitary U_p is applied outside the quantum switch as

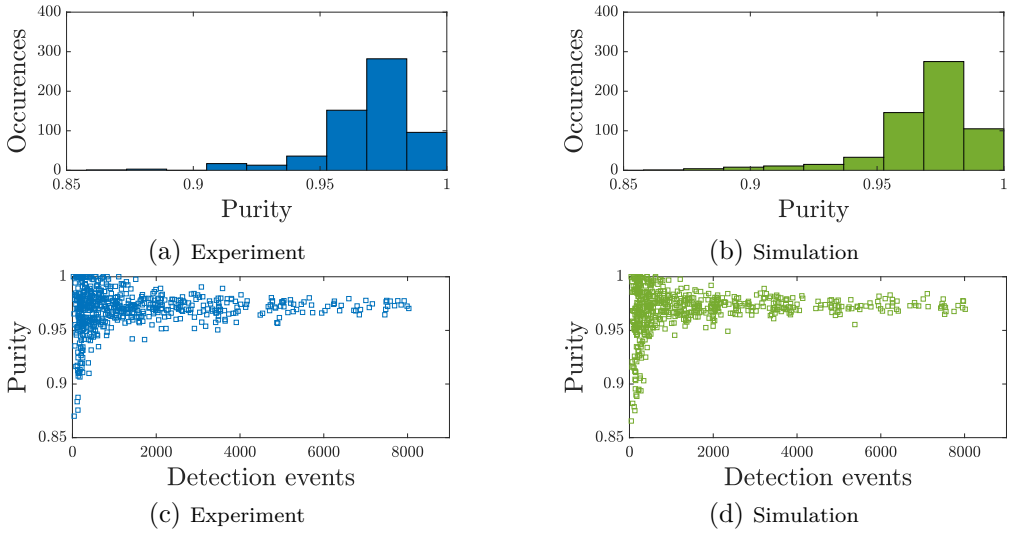


Figure 5.25: **Observed and simulated state purities, n = 2.** The left-hand figures show the distribution of purities for $n = 2$, while the right-hand figures show the results of a simulation using the noise model described in the main text, with the same parameters as in Fig. 5.24. Good agreement can be seen for both the relative frequency of purities, as well as the relationship between the state purities and the number of detected photons. The correct asymptotic behaviour in the simulation is achieved by simulating the tomography on a state with a purity of approximately 0.97.

well. For this reason the simulation was performed for the seven different choices of p.

An example of the simulated dependence on p in the state purity for $n = 1$ is shown in Fig. 5.27a. The simulation shows that the variance in the purity is large. This makes it hard to judge whether the depolarizing noise observed in the experiment was due to polarization mode dispersion in the EO-switches, or was caused by something else. Fig. 5.27b shows two examples of simulation outputs, differentiated by their random fiber unitaries. While the blue curve shows a purity that is largely independent of p, the yellow curve displays larger variations. These two simulated datasets were used as input to simulations of the measurement process, incorporating the noise model described earlier in the section. The result of these simulations is shown in Fig. 5.27c and Fig. 5.27d, and it can be seen that while the simulation corresponding to the blue curve in Fig. 5.27b shows a purity distribution that agrees well with the experimental results, using the yellow curve as input results in a distribution that does not fit the experiment.

The challenge in evaluating these results is that, since the polarization compensation was only done once for each run of $n = 1, 2, 3$, the experiment only sampled three different noise configurations, and does therefore not give any insight into the distribution of the depolarizing noise. While the simulation in Fig. 5.27c does fit the experiment reasonably well, it is unclear how likely this outcome is using the noise model. Fig. 5.28 attempts to quantify this by visualizing the distribution of the standard deviation of the purities over the different values of p. In the figure it can be seen that for relatively high purities, close to 0.99, the variance across p is low. However, for purities closer to 0.98, as observed in the $n = 2$ measurement, the typical variance is already larger. Since the observed

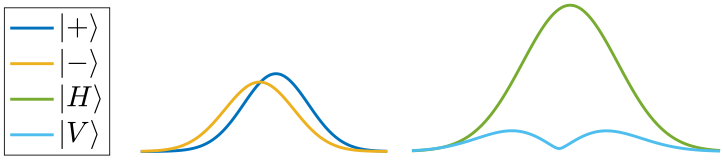


Figure 5.26: **Wavepacket polarization mode dispersion.** The figure shows an example of a simulated wavepacket at the output of the experiment. Due to the polarization mode dispersion in the EO-switches two polarization components of the wavepacket separate. This example was simulated for $p = 4$ and the magnitude of the wavepacket is shown in the $+/-$ basis on the left, and in the H/V basis on the right. Due to the imperfect overlap of the wavepackets in the diagonal basis they do not show full destructive interference in the H/V basis, leaving a residual $|V\rangle$-component. The purity of the state is approximately 0.95.

purities for this measurement is well modelled using depolarizing noise that does not depend on p, it makes the walk-off model slightly less plausible.

The influence of phase noise inside the Sagnac was also simulated, using the measured phase distribution shown in Fig. 5.9b, however the observed phase noise was found to be at least one order of magnitude to small to act as a significant source of depolarizing noise. Furthermore, the qualitative behaviour of this type of depolarizing noise does not agree with observations when taken on its own. In particular, the amount depolarization depends strongly on the particular choice of p and q, and for pairs of nearly anti-commuting unitaries the phase noise only produces negligible depolarization. Ultimately, there is not enough data to conclusively ascertain the cause of the depolarizing noise.

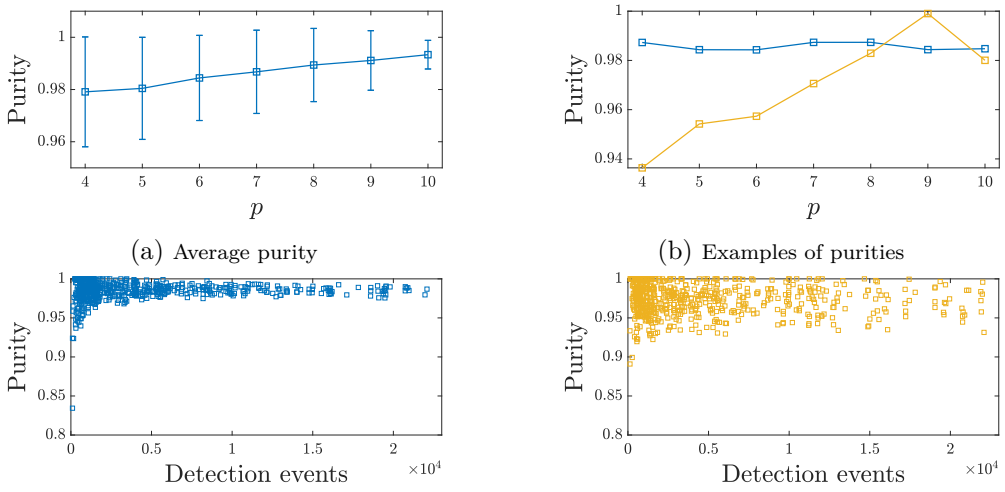
5.3.3 Process tomography

Since the experimentally implemented rewinding process was probed by a tomographically complete set of input states, the idea of performing quantum process tomography on the protocol was naturally investigated. The basic steps for doing this follow the theory presented in Section 1.2.9. Since the protocol has a finite success probability the experimentally realised process was not trace preserving. However, accurately estimating the degree to which the process was trace reducing was beyond the scope of the experiment.[4] Instead, the process tomography only used post-selected and normalized data. More concretely, instead of fitting a process matrix to the raw observed photon counts, it was fitted to the expectation values of the operators measured in the quantum state tomography.

Similar to the estimation of the quantum state fidelities described in Section 5.3.1, the process matrices were reconstructed as part of a Monte-Carlo simulation that in each step sampled Poissonian noise using the photon counts, and performed a maximum likelihood fit to the data. The fitting amounts to minimizing

$$\sum_{kl} \frac{\left(\mathrm{Tr}\left[|P_k\rangle\langle P_k|\,\mathcal{E}(|\Psi_l\rangle\langle\Psi_l|)\right] - M_{kl}\right)^2}{\mathrm{Tr}\left[|P_k\rangle\langle P_k|\,\mathcal{E}(|\Psi_l\rangle\langle\Psi_l|)\right]}, \tag{5.3-11}$$

[4]The relative success probabilities are presented in the publication, however this does not quantify the absolute trace reduction.



(a) Average purity (b) Examples of purities

(c) Good agreement (d) Poor agreement

Figure 5.27: **Simulation of depolarizing noise.** The amount of depolarization induced by the EO-switches depends on the value of p. This dependence is shown in (a) for $n = 1$ and a per-switch PMD of 0.2 ps. The data is averaged over 250 different runs of the simulation, and while the average purity changes slowly as a function of p the standard deviation is large. (b) shows two examples of purities from single runs of the simulation. The blue curve has a nearly constant purity, while the yellow curve shows much larger variations. (c) and (d) show simulations of the apparent state purities using the noise model discussed in the main text, as well as the two sets of purities from (b) as input. In (c) the simulation using the blue curve in (b) is shown, and good agreement with the experiment can be seen, cf. Fig. 5.23 and Fig. 5.24. On the other hand, the corresponding simulation for the yellow curve in (b), which is displayed in (d), shows a distribution of purities that is much too broad.

where $|P_k\rangle\langle P_k|$ are the measurement projectors, $|\Psi_l\rangle\langle\Psi_l|$ are the input states, M_{kl} is the experimentally measured expectation value of the k-th projector on the l-th input state and $\mathcal{E}$ is the quantum channel defined by the process matrix being fitted. During the fitting procedure the map $\mathcal{E}(\rho)$ needs to be evaluated a great number of times. We recall that the definition of this map is

$$\mathcal{E}(\rho) = \sum_{mn} \chi_{mn} \widetilde{E}_m \rho \widetilde{E}_n^\dagger. \tag{5.3-12}$$

Since the operators $\widetilde{E}_j$ and input states ρ_i are fixed, all the terms $\widetilde{E}_m \rho_j \widetilde{E}_n^\dagger$ can be precomputed and represented as a 16×4 matrix A_i, whose rows are $|\widetilde{E}_m \rho_i \widetilde{E}_n^\dagger\rangle\rangle^T$:

$$A_i = \begin{bmatrix} |\widetilde{E}_1 \rho_i \widetilde{E}_1^\dagger\rangle\rangle^T \\ |\widetilde{E}_2 \rho_i \widetilde{E}_1^\dagger\rangle\rangle^T \\ \vdots \\ |\widetilde{E}_4 \rho_i \widetilde{E}_1^\dagger\rangle\rangle^T \\ |\widetilde{E}_1 \rho_i \widetilde{E}_2^\dagger\rangle\rangle^T \\ \vdots \\ |\widetilde{E}_4 \rho_i \widetilde{E}_4^\dagger\rangle\rangle^T \end{bmatrix}. \tag{5.3-13}$$

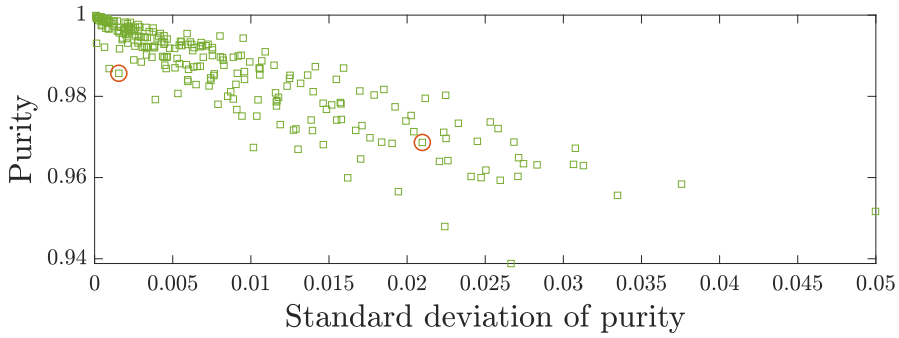


Figure 5.28: **Polarization walk-off and state purity.** Monte-Carlo simulations of the polarization mode dispersion in the EO-switches were performed to estimate the effect this has on the final quantum state purity, and the results of one such simulation is shown above. The data points marked with red circles correspond to the two curves in Fig. 5.27b. The decrease in state purity as a result of the walk-off depends on the choice of p, but not on the choice of q. The points in scatter plot correspond to a single run of the simulation, in which random fiber unitaries were generated, and each point has additionally been averaged over the seven values of $p = 4 \dots 10$. The x-axis of the graph shows the standard deviation of the purity over this set of values of p. For relatively high average purities the standard deviation is low, meaning that the amount of depolarization is roughly constant as p is changed.

The vectorized density matrices can then be efficiently computed as

$$|\mathscr{E}(\rho_i)\rangle\!\rangle = |\chi\rangle\!\rangle^T A_i. \tag{5.3-14}$$

Since the matrix χ is Hermitian it can be parametrised using the lower-triangular matrix decomposition for density matrices defined in (1.2-15) and (1.2-17). A problem with this decomposition, though, is that it does not ensure that the quantum channel described χ matrix is trace preserving. More specifically, the channel is only trace preserving if

$$\sum_{mn} \chi_{mn} \widetilde{E}_n^\dagger \widetilde{E}_m = \mathbb{1}. \tag{5.3-15}$$

While the actual experimental process is not trace preserving, since there is photon loss and the application of the commutators is probabilistic, the process defined by the post-selected states is. One would therefore like to have an optimization process that respects this condition. To do this, it turns out to be more convenient to work with the Choi-matrix representation of the channel instead. Let

$$\mathcal{C}_{\mathscr{E}} = \big(\mathbb{1}_A \otimes \mathscr{E}\big)(|\Omega\rangle\langle\Omega|_{AB}\big), \tag{5.3-16}$$

Be the Choi matrix of the map $\mathscr{E}$. Suppose $\mathscr{E}$ is trace preserving, then [408]:

$$\begin{aligned} \mathrm{Tr}[\rho] &= \mathrm{Tr}\big[\mathscr{E}(\rho)\big] \\ &= \mathrm{Tr}\big[\mathrm{Tr}_A[(\rho_A^T \otimes \mathbb{1}_B)\mathcal{C}_{\mathscr{E}}]\big] \\ &= \mathrm{Tr}\big[\mathrm{Tr}_B[(\rho_A^T \otimes \mathbb{1}_B)\mathcal{C}_{\mathscr{E}}]\big] \\ \mathrm{Tr}[\rho] &= \mathrm{Tr}\big[\rho_A^T \mathrm{Tr}_B[\mathcal{C}_{\mathscr{E}}]\big] \end{aligned} \tag{5.3-17}$$

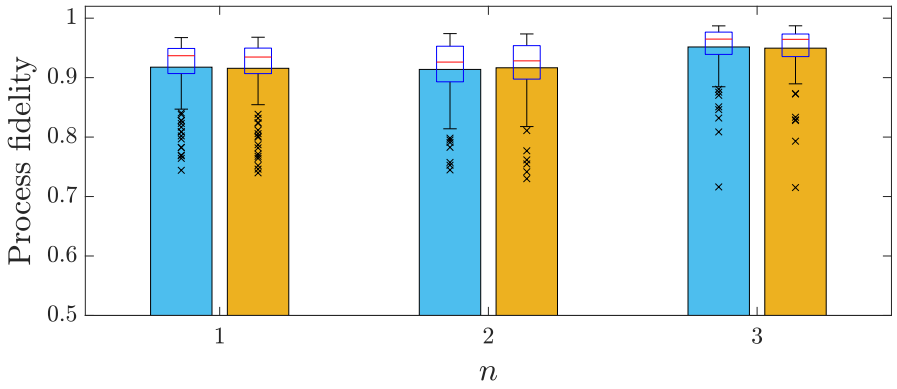


Figure 5.29: **Average process fidelities.** The barplot shows the average process fidelity for the three values of n, using two different reconstructions of the quantum process. The blue bars indicate the fidelities obtained when fitting the χ matrix, without enforcing the constraint that the quantum channel be trace preserving, while the yellow bars show the fidelities that result from fitting a Choi matrix that does respect this constraint. In this instance, the difference between the two approaches is negligible. Both sets of process fidelities, however, are slightly lower than the equivalent state fidelities shown in Fig. 5.20. This is not entirely surprising since the quantum process tomography involves fitting more data than the simpler qubit state tomography, and the data for the different input states might not be perfectly compatible with a quantum channel.

This implies that:

$$\mathrm{Tr}_B\left[\mathcal{C}_{\mathcal{E}}\right] = \mathbb{1}_A, \tag{5.3-18}$$

and this is both a necessary and sufficient condition for $\mathcal{E}$ to be trace preserving. Rather than parametrising $\mathcal{E}$ in a way that respects this constraint, one can fit the Choi matrix using semidefinite programming (SDP) methods, which is a type of convex optimization for linear objective functions, in which the optimization variable (matrix) is positive semidefinite, and possibly subject to additional constraints [409]. These optimization methods are widely used within quantum information [410,411]. The objective function (5.3-11) cannot be recast as a linear objective function due to the optimization variable appearing in the denominator, however the following related least-squares problem can:

$$\sum_{kl} \frac{\left(\mathrm{Tr}\left[|P_k\rangle\langle P_k|\,\mathcal{E}(|\Psi_l\rangle\langle\Psi_l|)\right] - M_{kl}\right)^2}{M_{kl}}. \tag{5.3-19}$$

While this objective function does not have interpretation of maximizing the likelihood that the minimized variable generated the observed data, it is in practice almost identical to (5.3-11) since $\mathrm{Tr}\left[|P_k\rangle\langle P_k|\,\mathcal{E}(|\Psi_l\rangle\langle\Psi_l|)\right] \to M_{kl}$ during the optimization. The corresponding SDP was solved using YALMIP and MOSEK in MATLAB as part of a Monte-Carlo simulation identical to the ones previously described.

As process fidelity one typically uses gate fidelity (1.2-10):

$$\mathcal{F}_{\text{process}} = \left\langle \mathrm{Tr}\left[\sqrt{\sqrt{\mathcal{E}(\rho)}U(\rho)\sqrt{\mathcal{E}(\rho)}}\right]\right\rangle_{\rho}, \tag{5.3-20}$$

restated here for convenience. The process fidelity was evaluated with respect to the

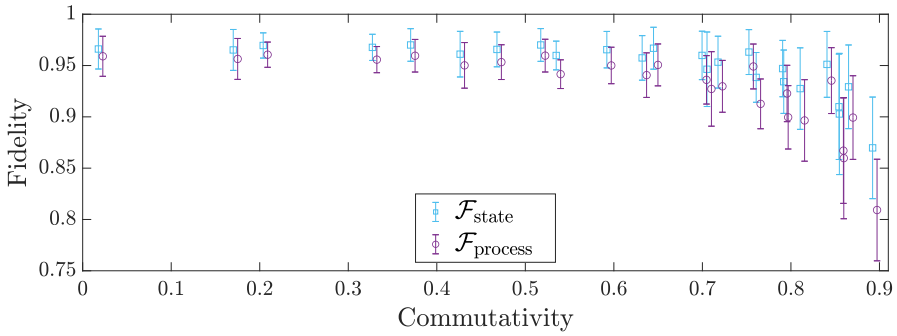


Figure 5.30: **State and process fidelities.** As would be expected, the state fidelities and process fidelities show a strong correlation. In the figure above these fidelities are plotted as a function of the commutativity N_c. The two datasets are symmetrically offset relative to each other on the x-axis for visual clarity. A higher commutativity corresponds to a lower success probability, and the decrease in fidelity as a function of the commutativity is explained by the lower photon count rates. Additionally the data for low success probabilities corresponds to measuring a nearly dark port of the interferometer, and is therefore more susceptible to for example imperfect interferometric visibility. The state fidelities consist of averages over 72 different measurements, while the data points for the process fidelities are averaged over 18 different Choi matrices.

post-selected rewinding protocol, which is a unitary channel:

$$U^{-n} \iff \mathcal{C} = |U^{-n}\rangle\!\rangle\langle\!\langle U^{-n}| = (\mathbb{1} \otimes U^{-n})\,|\Omega\rangle\langle\Omega|\,. \tag{5.3-21}$$

Since the Choi matrix $\mathcal{C}$ corresponds to a pure quantum state, one might be tempted to make an analogy to the expression for the fidelity between two quantum states (5.3-3), that holds when one of them is pure, and define a process fidelity:

$$\mathcal{F}_{\text{Choi}} = \frac{1}{d^2}\text{Tr}\big[\mathcal{C}\mathcal{C}_{\text{exp}}\big]. \tag{5.3-22}$$

This fidelity is not equivalent to (5.3-20), however the two are related by a simple affine transformation [412, 413]:

$$\begin{aligned}\mathcal{F}_{\text{process}} &= \left\langle \text{Tr}\left[\sqrt{\sqrt{\mathcal{E}(\rho)}\,U(\rho)\sqrt{\mathcal{E}(\rho)}}\,\right]\right\rangle_{\rho} \\ &= \int_{|\psi\rangle} \mathcal{F}\big(\mathcal{E}(|\psi\rangle\langle\psi|), U\,|\psi\rangle\langle\psi|\,U^{\dagger}\big)\,d\,|\psi\rangle \\ &= \frac{d}{d+1}\mathcal{F}_{\text{Choi}}\big(\mathcal{C}_{\mathcal{E}}, |U\rangle\!\rangle\langle\!\langle U|\big) + \frac{1}{d+1},\end{aligned} \tag{5.3-23}$$

where the average is assumed to be taken over the Haar measure. This formulation is convenient, because (5.3-23) gives a direct way to evaluate the process fidelity without actually having to perform the averaging over all states.

The process fidelity was evaluated both for the reconstructed Choi matrices as well as the χ matrices, and in the latter case the fidelity was averaged over 1000 states. The results of this are shown in Fig. 5.29. The two methods produce almost identical fidelities, because in practice the reconstructed χ matrix is, in most cases, only marginally

non-trace preserving. Both process fidelities are, however, consistently lower than the corresponding average state fidelities. This is to be expected, since the reconstruction of the process requires fitting more of the data at the same time than the state reconstruction. Measurement noise, as well as time dependent effects in the experiment, may cause the acquired data to represent an average of several different processes. A fit to this data will then not perfectly reproduce the measured states, let alone the ideal ones.

The state and process fidelities were nevertheless strongly correlated, and in Fig. 5.30 the two data sets are plotted together as a function of the commutativity N_c. The process fidelity is strictly smaller than the average state fidelity for every value of the commutativity, but the two fidelity metrics exhibit the same behaviour, decreasing as a function of the commutativity. This is explained by the fact that the success probability decreases as a function of the commutativity, and therefore so too does the photon count rate, increasing the impact of noise in the experiment.

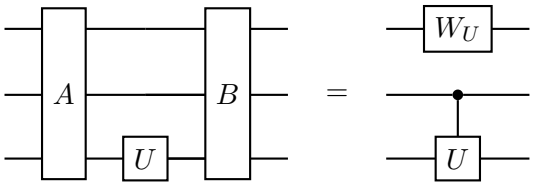
6

Experimental methods for "Experimental superposition of a quantum evolution with its time reverse"

This chapter will discuss some of the experimental details as well as the theoretical background of Publication 3. The experiment was motivated by a recent theoretical proposal by Chiribella and Liu [414], in which they describe a novel process that is not captured by the quantum circuit formalism. This process, named the quantum time flip, amounts to transforming a unitary operator into a coherent superposition of itself and its transpose. The name comes from the fact that the transpose of a unitary time evolution can be interpreted as its time reverse. It can be shown that deterministic unitary transposition of a single gate is forbidden in the quantum circuit formalism. However, by exploiting knowledge of the experimental apparatus, it is nevertheless possible to experimentally realise the time-flip process. Building on the results from Chapter 3 we devise a polarization gadget that simultaneously realises both a unitary and its transpose, thereby allowing for the two to be placed in superposition.

6.1 Non-quantum-circuit-model processes

The quantum circuit model is the most commonly used framework for describing quantum information processing tasks [15]. It is, however, important to recognize that this framework does not capture all of quantum mechanics, or indeed all quantum information processing protocols. A reliance on this formalism can therefore, in some cases, lead to a tension between what is allowed in the abstract model, and what is actually possible in an experiment. A canonical example of this is coherent control of a unitary with only a single use of the gate. In terms of quantum circuits the problem of coherent unitary control can be stated as finding two unitaries A and B such that circuit equality



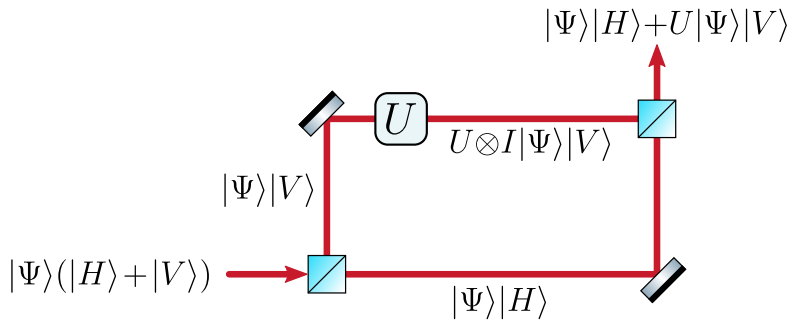


Figure 6.1: **Coherent unitary control.** Using only linear optics an arbitrary unitary U can be turned into a controlled version of itself. This is accomplished by coupling the control degree of freedom, in this example the photon polarizaiton, to the photon path, and then placing the device implementing the unitary U in one of the two photon paths. This unitary acts on the target state $|\Psi\rangle$, encoded in some other interngal degree of freedom of the photon.

is satisfied. Here W_U is some gate, possibly depending on U, acting on an ancillary state. In Ref. [415] it was proven that there is no assignment of A and B that satisfies the equation, even when only controlling U up to a global phase. In the same paper, the authors also pointed out that there is a simple way to perform coherent unitary control in a photonic setting, and coherent unitary control has indeed been demonstrated experimentally [416].

The process for doing this is illustrated in Fig. 6.1, in which a controlled operation between a qubit encoded in the photon polarization, and a second qubit encoded in a different internal degree of freedom, is performed by coupling the polarization to the photon path. The device implementing the single-qubit operation U is only present in one of the photon paths, and a controlled operation between the path and an internal degree of freedom is therefore realised. This method circumvents the no-go theorem of the circuit model by exploiting knowledge about the position of the physical system implementing the gate, and any device implementing a unitary transformation in a finite spatial region is in a sense already a controlled operation between a spatial and an internal degree of freedom of a particle.

Another example of a process that cannot be described in the quantum circuit model, but that nevertheless has a physical realisation, is the quantum switch [55]. As discussed in Chapter 5, this is a process that for qubits transforms a pair of unitaries (U, V) into a superposition of the gates being applied in different orders, and more generally transforms N unitaries into a superposition of them being applied in every possible permutation. This can be thought of as a controlled superposition of quantum circuits, for instance in the case of qubits:

Since a circuit consists of a fixed sequence of gates, the above process cannot be described in the quantum circuit formalism. The quantum switch was first introduced in the context of studying causality in quantum mechanics [54, 55], since if one associates the application of the gates U and V with events in spacetime, then the order of these

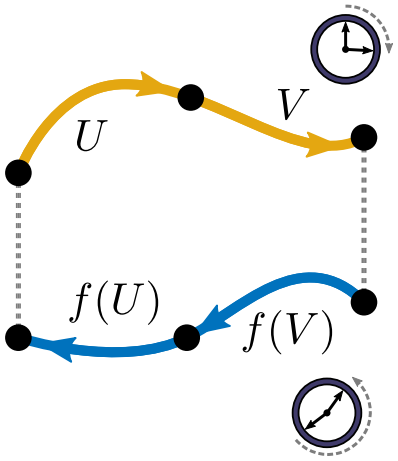


Figure 6.2: **Forwards and backwards time evolutions.** Any time-reversal operator in quantum mechanics has to have the property of being order reversing. This is illustrated here by decomposing a forwards (yellow) and a backwards (blue) time evolution into two steps. The total forwards time evolution is given by VU, and clearly the backwards evolution is $f(U)f(V)$, where f is the time-reversal operator. This shows that $f(VU) = f(U)f(V)$, meaning that f is order reversing. There are only two non-trivial choices of such operators, the transpose and the adjoint [414], and for two-level systems they are unitarily equivalent.

events becomes indefinite. The process matrix formalism outlined in Section 1.2.8 was developed to describe these kinds of processes with more general causal structures. Photonic implementations of the quantum switch bypass the limitations of the quantum circuit model much in the same way they do in the case of coherent unitary control, by using knowledge of the photon path to control the order of two or more gates, instead of controlling a single gate [11, 56, 406].

More general causally-indefinite models of quantum computation that make use of the quantum switch have been shown to allow for a speed-up in certain computational tasks [406, 417, 418, 419, 420, 421, 422]. This gives a motivation for studying these non-circuit-model processes, both theoretically and experimentally, that goes beyond the context of quantum foundations. An example of a computational task where the quantum switch provides an advantage over causally ordered circuits is the problem of deciding whether a pair of unitaries commute or anti-commute. This problem cannot be solved deterministically given a single use of the gates, whereas the quantum switch can solve it trivially, since as shown in Chapter 5 it can be used to create a controlled superposition of the commutator and anti-commutator.

6.1.1 The quantum time flip

The non-circuit-model processes described above are all captured by the process matrix formalism, but it turns out that quantum mechanics allows for even more general processes that cannot be described using this framework. An example of such a process, called the quantum time flip, was introduced in Ref. [414]. This process takes a unitary operator and transforms it into a superposition of itself and its time-reversed version:

$$U \mapsto U \otimes |0\rangle\langle 0|_C + \theta(U) \otimes |1\rangle\langle 1|_C \,. \tag{6.1-1}$$

Here θ is a time-reversal operator. Superpositions of processes in different time directions had previously been studied in Ref. [423]. In that work the authors considered superpositions of entropy increasing and decreasing processes, which can be associated with forwards and backwards arrows of time, respectively. In contrast, the quantum time flip is only well defined for processes that conserve entropy [414]. The relation of the quantum time fip to more general causally indefinite processes will be discussed in the next section. The choice of time-reversal operator θ in (6.1-1) is not unique; however, as shown in Fig. 6.2 the operator has to be order reversing:

$$\theta(UV) = \theta(V)\theta(U) \tag{6.1-2}$$

For qubits, two obvious candidates for such an operation are the transpose and the inverse:

$$(UV)^T = V^T U^T, \qquad (UV)^\dagger = V^\dagger U^\dagger. \tag{6.1-3}$$

In the case of qubit systems these operations are unitarily equivalent:

$$YU^T Y = U^\dagger, \tag{6.1-4}$$

and the choice of operation is therefore a matter of taste.[1] In particular, it does not affect the relation of the quantum time flip to causally ordered circuits, or causally indefinite processes. Deterministic unitary transposition, or inversion, is not possible in the circuit model when given access to a single use of a gate [424]. One way to understand this is by using the fact that a quantum supermap realising unitary transposition can be shown to be unique [425], which is a consequence of the constraint of positivity on such maps. Since, as shown in 5.1.1, quantum teleportation can be used to perform unitary transposition, the existence of such a supermap would imply that teleportation could be performed deterministically without adaptive corrections, which would in turn allow for superluminal communication.

For Publication 3 the transpose was chosen as the time-reversal operator:

$$\theta(U) = U^T. \tag{6.1-5}$$

One motivation for this is that when working with the Choi matrix of a unitary channel, transposing the operator corresponds to exchanging the input and output systems:

$$\mathbb{1} \otimes U^T |\Omega\rangle\langle\Omega| = U \otimes \mathbb{1} |\Omega\rangle\langle\Omega| . \tag{6.1-6}$$

This naturally leads to the concept of bidirectional quantum devices, for which one direction corresponds to the forwards time direction, and the other to the backwards time direction, as shown in Fig. 6.3. Introducing a control degree of freedom that determines the direction in which such a bidirectional device is accessed then generates the quantum time flip. This is illustrated in Fig. 6.4. These bidirectional devices furthermore map directly onto the experimental realisation of the quantum time flip.

[1] In the case of bipartite systems the transpose and adjoint exhibit a difference, as the latter is not guaranteed to generate CPTP maps when acting locally on one of the subsystems. See [414] for a discussion on this.

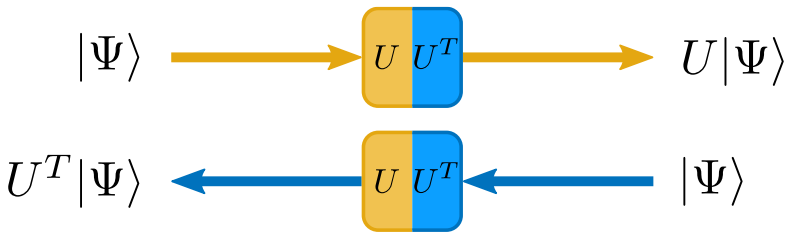


Figure 6.3: **Bidirectional devices.** In quantum information it is common to consider 'black boxes' with an input and an output port that implement some, possibly unknown, transformation on the input state. Time reversal in quantum mechanics can be thought of as exchanging the roles of the input and output ports of such a box, and it therefore natural to consider bidirectional boxes, where either port can act as an input or an output. As shown above, the direction in which such a box is accessed determines the time direction of the channel implemented by the box. Here the yellow path gives the forwards time evolution U, while the blue path gives the time-reversed evolution U^T.

The choice of time-reversal operator can also be motivated in a more formal way, as discussed in Ref. [414]. One begins by noting that the time-reversal operator has to be either unitary or anti-unitary in order to conserve probability. More concretely, let

$$\Theta\,|\psi(t)\rangle = |\widetilde{\psi}(\tau - t)\rangle\,, \tag{6.1-7}$$

be the time-reversed version of $|\psi(t)\rangle$, where Θ is a time-reversal operator and $\tau/2$ is a reference time with respect to which the time-reversal is taken. Then

$$\begin{aligned}\langle\psi(t)|\psi(t)\rangle &= \langle\widetilde{\psi}(\tau-t)|\widetilde{\psi}(\tau-t)\rangle \\ &= \langle\psi(t)|\Theta^\dagger\Theta|\psi(t)\rangle \implies \Theta^\dagger\Theta = 1.\end{aligned} \tag{6.1-8}$$

However, Θ cannot be unitary. This is because the position and momentum operators $\hat{x}$ and $\hat{p}$ transform as

$$\hat{x} \mapsto \hat{x} \tag{6.1-9}$$

$$\hat{p} \mapsto -\hat{p} \tag{6.1-10}$$

under time reversal. If the time-reversal operator is unitary then the canonical commutation relation for position and momentum is not preserved under time reversal:

$$[\hat{x},\hat{p}] = i\hbar \mapsto \Theta[\hat{x},\hat{p}]\Theta^\dagger = \Theta i\hbar\Theta^\dagger = i\hbar \neq [\hat{x},-\hat{p}]. \tag{6.1-11}$$

Hence Θ is anti-unitary. Next, let $U_{\text{fw}}(t,\tau)$ be the forward time evolution operator from time $T=\tau$ to $T=t$. Then

$$\begin{aligned}&U_{\text{fw}}(t,\tau)\,|\psi(\tau)\rangle = |\psi(t)\rangle \\ \iff\ &\Theta U_{\text{fw}}(t,\tau)\,|\psi(\tau)\rangle = |\widetilde{\psi}(\tau-t)\rangle\,.\end{aligned} \tag{6.1-12}$$

Using $\Theta^\dagger\Theta = \mathbb{1}$ this can be rewritten as

$$\begin{aligned}&\Theta U_{\text{fw}}(t,\tau)\Theta^\dagger\Theta\,|\psi(\tau)\rangle = |\widetilde{\psi}(\tau-t)\rangle \\ \iff\quad &\Theta U_{\text{fw}}(t,\tau)\Theta^\dagger\,|\widetilde{\psi}(0)\rangle = |\widetilde{\psi}(\tau-t)\rangle = U_{\text{bw}}(0,\tau-t)\,|\widetilde{\psi}(0)\rangle\,.\end{aligned} \tag{6.1-13}$$

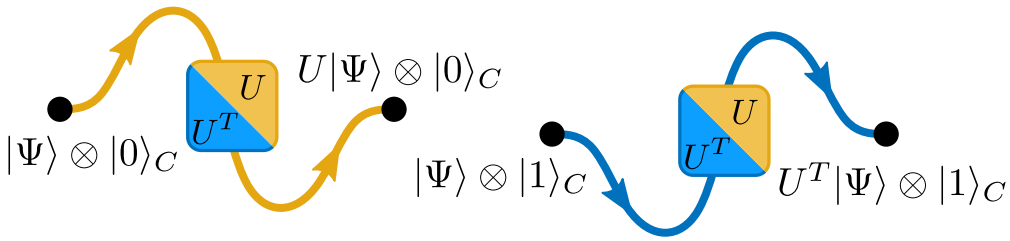


Figure 6.4: **Coherent control of time direction.** By adding a control degree of freedom that determines which direction a bidirectional device is accessed in, one gains control over the direction of the time evolution. By placing the control qubit in a superposition state the two time directions become coherently superposed, and this is a realisation of the quantum time flip.

Since this equation holds for any state $|\psi\rangle$ it follows that:

$$\Theta U_{\text{fw}}(t,\tau)\Theta^\dagger = U_{\text{bw}}(\tau - t, 0). \tag{6.1-14}$$

By taking the inverse, this can be rewritten as

$$\Theta U_{\text{fw}}^\dagger(t,\tau)\Theta^\dagger = U_{\text{bw}}(0, \tau - t), \tag{6.1-15}$$

or, by making the initial and final times implicit, as

$$\Theta U_{\text{fw}}^\dagger \Theta^\dagger = U_{\text{bw}}. \tag{6.1-16}$$

In the context of thermodynamics this equation is known as the microreversibility relation, which in the case of a time-dependent Hamiltonian requires a more careful derivation [426]. The interpretation of (6.1-16) is that if one (i) lets a state evolve in time under the forward time evolution, (ii) time-reverses the state, (iii) lets the resulting state evolve under the time-reversed evolution, (iv) time-reverses the state a second time, then one recovers the initial state.

Finally, we note that any anti-unitary operator can be decomposed in terms of a unitary operator W and the complex-conjugation operator $\mathcal{K}$:

$$\Theta = W\mathcal{K}, \tag{6.1-17}$$

and this lets (6.1-16) be expressed as

$$\begin{aligned} U_{\text{bw}} &= W\mathcal{K}U_{\text{fw}}^\dagger\mathcal{K}^\dagger W^\dagger \\ &= W(U_{\text{fw}}^\dagger)^* W^\dagger \\ &= W U_{\text{fw}}^T W^\dagger, \end{aligned} \tag{6.1-18}$$

showing, once again, that the time-reversal operator $\theta : U_{\text{fw}} \mapsto U_{\text{bw}}$ is unitarily equivalent to the transpose.

6.1.2 Advantage of the quantum time flip in a channel discrimination task

In analogy to how the quantum switch can outperform causally ordered quantum circuits in the task of determining whether two unitary channels commute or anti-commute,

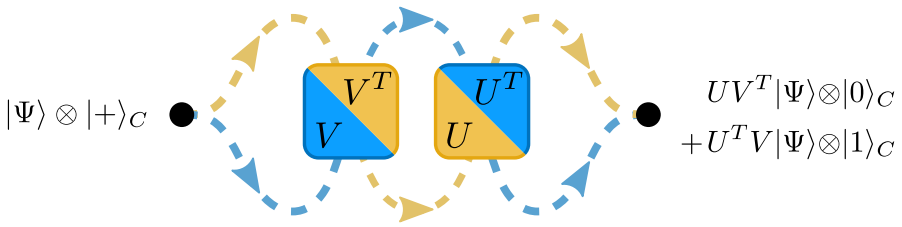


Figure 6.5: **Quantum time flip strategy.** The quantum time flip can be used to deterministically perform the channel discrimination task described in the main text, and method for doing this is illustrated above. The strategy uses two time flips, that have opposite actions with respect to the control qubit. This leads to an output superposition state where the two time evolutions act in the same order in both terms, but the relative time direction of the two evolutions is reversed. Projecting the control qubit of the output state in the $\{|+\rangle, |-\rangle\}$ basis reveals which set the pair of unitaries were picked from.

there exist a channel discrimination task that can only be deterministically accomplished by using the quantum time flip. More specifically, consider the following two sets of pairs of unitaries:

$$\mathcal{M}_+ := \Big\{(U,V) : UV^T = +U^TV\Big\} \tag{6.1-19}$$

$$\mathcal{M}_- := \Big\{(U,V) : UV^T = -U^TV\Big\}. \tag{6.1-20}$$

Given a pair of unitaries picked from either $\mathcal{M}_+$ or $\mathcal{M}_-$, the task is to determine from which set they were picked while only accessing each gate once. This task can be performed deterministically by concatenating two quantum time flips acting opposite to each other; the first one on V and the second one on U:

$$\begin{aligned}(U,V) &\mapsto V \otimes |0\rangle\langle 0|_C + V^T \otimes |1\rangle\langle 1|_C \\ &\mapsto U^TV \otimes |0\rangle\langle 0|_C + UV^T \otimes |1\rangle\langle 1|_C\,.\end{aligned} \tag{6.1-21}$$

One then lets this process act on an initial state for which the control qubit is in a balanced superposition:

$$|\Psi\rangle_T \otimes |+\rangle_C = |\Psi\rangle_T \otimes \frac{|0\rangle_C + |1\rangle_C}{\sqrt{2}}. \tag{6.1-22}$$

This results in the state:

$$\begin{aligned}&\frac{1}{\sqrt{2}}(U^TV\,|\Psi\rangle_T \otimes |0\rangle_C + UV^T\,|\Psi\rangle_T \otimes |1\rangle_C) \\ &= \frac{1}{2}(UV^T + U^TV)\,|\Psi\rangle_T \otimes |+\rangle_C + \frac{1}{2}(UV^T - U^TV)\,|\Psi\rangle_T \otimes |-\rangle_C\,.\end{aligned} \tag{6.1-23}$$

A measurement of the control qubit in the diagonal basis then reveals from which set the unitaries were picked. This strategy is illustrated in Fig. 6.5. In order to compare the time flip strategy with ones that do not employ the quantum time flip one needs to

specify the sets (6.1-19) and (6.1-20), since the optimal strategies can be tailored to the sets. For the experiment, the following sets, first introduced in Ref. [414], were chosen:

$$\begin{aligned}\mathcal{M}_+ := \Bigg\{ & (I,I),(I,X),(I,Z),(X,I),(X,X),(X,Z),(Z,I),(Z,X),(Z,Z), \\ & \left(\frac{X-Y}{\sqrt{2}},\frac{X+Y}{\sqrt{2}}\right),\left(\frac{X+Y}{\sqrt{2}},\frac{X-Y}{\sqrt{2}}\right), \\ & \left(\frac{Z-Y}{\sqrt{2}},\frac{Z+Y}{\sqrt{2}}\right),\left(\frac{Z+Y}{\sqrt{2}},\frac{Z-Y}{\sqrt{2}}\right)\Bigg\}\end{aligned} \tag{6.1-24}$$

$$\begin{aligned}\mathcal{M}_- := \Bigg\{ & (Y,I),(Y,X),(Y,Z),(I,Y),(X,Y),(Z,Y), \\ & \left(\frac{I+iY}{\sqrt{2}},\frac{I-iY}{\sqrt{2}}\right),\left(\frac{I-iY}{\sqrt{2}},\frac{I+iY}{\sqrt{2}}\right)\Bigg\}.\end{aligned} \tag{6.1-25}$$

One can then consider three different classes of strategies for the channel discrimination task, as illustrated in Fig. 6.6. The first of these strategies uses the gates in parallel, meaning that the experimenter prepares a general quantum state and then lets the two gates U and V act on different subsystems of this state, after which a generalized measurement may be performed. The second type of strategy involves using the gates sequentially, in a definite causal order, with an additional quantum channel applied between the two gates. The final, and most general, strategy is one described by a process matrix. This includes causally indefinite strategies using for example the quantum SWTCH or the Oreshkov-Costa-Brukner process [54]. This last class contains the most general strategies possible that only access the gates in a single time direction. Note that convex combinations of strategies in a given class are also valid strategies of the same class, i.e. an agent may probabilistically choose different strategies of the same class. All the strategies above correspond to processes that can be said to have a definite time direction, since they only access the gates in one time direction. Conversely, processes which cannot be written as a convex combination of processes that only access the gates in one time direction can therefore be said to have an indefinite time direction, in analogy to processes with an indefinite causal structure.

The three above strategies were analysed using the tester formalism [427, 428, 429]. Formally, a tester is a set of linear, positive-semidefinite operators acting on the joint input-output space of a channel:

$$T = \{T_i\}, \quad T_i \in \mathcal{L}(\mathcal{H}_I \otimes \mathcal{H}_O) \tag{6.1-26}$$

$$T_i \geq 0. \tag{6.1-27}$$

These operators furthermore have the property that they sum to a valid process matrix:

$$\sum_i T_i = W. \tag{6.1-28}$$

Testers encompass the preparation of the states sent through the channels, the manner in which the channels are accessed, and the measurements on the resulting output states.

They therefore describe a general scenario for measuring quantum channels, and can be thought of as POVMs for quantum processes. By imposing constraints on the process matrix that the testers add up to, one can restrict the analysis to certain classes of strategies. For instance, in the case of a parallel process the process matrix of the tester respects:

$$W^{\text{par}} = \frac{\mathbb{1}_O}{d_O}\text{Tr}_O\big[W^{\text{par}}\big], \tag{6.1-29}$$

where d_O is the dimension of the output Hilbert space. In this formalism, the measurement outcomes associated with each process-POVM element T_i is given by a generalization of Born's rule:

$$P(i) = \text{Tr}\big[T_i\mathcal{C}_{\mathcal{E}}\big], \tag{6.1-30}$$

where $\mathcal{C}_{\mathcal{E}}$ is the Choi matrix of the channel. For the two-channel task that is the focus of this section, this can be expressed as

$$P\big(k|(U_i,V_i)\big) = \text{Tr}\big[T_k|U_i\otimes V_i\rangle\!\rangle\langle\!\langle U_i\otimes V_i|\big]. \tag{6.1-31}$$

In the case of a parallel strategy this is equivalent to the more explicit expression:

$$P\big(k|(U_i,V_i)\big) = \text{Tr}\big[M_k(U_i\otimes V_i\otimes\mathbb{1})\rho(U_i^\dagger\otimes V_i^\dagger\otimes\mathbb{1})\big], \tag{6.1-32}$$

where M_k are POVM elements defining the measurement performed by the tester, and ρ is the input state used to probe the process. The advantage of the tester formalism is that the state preparation, the accessing of the channel and the final measurements are represented by a single object. This allows for the problem of determining the optimal success probability for each class of strategy to be expressed as a convex optimization problem that can be solved using semidefinite programming methods. For the channel discrimination task realised in the experiment the tester consists of two elements, corresponding two the two binary outcomes of the channels belonging to $\mathcal{M}_+$ or $\mathcal{M}_-$. The optimal strategy for a given class is the one that maximizes the success probability:

$$\begin{aligned}\max\frac{1}{N}\Bigg[&\sum_{(U_i,V_i)\in\mathcal{M}_+}\text{Tr}\big[T_+|U_i\otimes V_i\rangle\!\rangle\langle\!\langle U_i\otimes V_i|\big]\\ &+\sum_{(U_i,V_i)\in\mathcal{M}_-}\text{Tr}\big[T_-|U_i\otimes V_i\rangle\!\rangle\langle\!\langle U_i\otimes V_i|\big]\Bigg],\end{aligned} \tag{6.1-33}$$

where N is the total number of channels summed over. Maximizing (6.1-33) provides a lower bound on the maximum success probability. In order to instead upper bound the success probability one can make use of the so-called dual problem, which is a reformulation of (6.1-33) as a minimization problem. The upper and lower bounds were calculated using the computer assisted proof methods from [427], which ensure that the SDP constraints are satisfied exactly despite the finite precision in the floating-point arithmetic used by the numerical solver. The theoretical analysis of the three different strategies was performed by Marco Túlio Quintino.

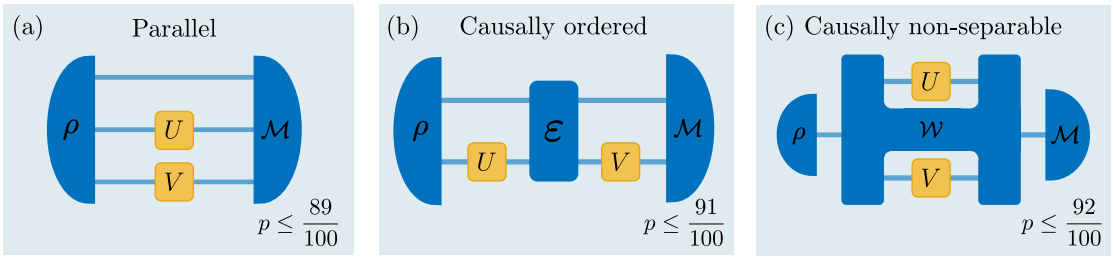


Figure 6.6: **Strategies with definite time direction** The strategies for the channel discrimination game that have a definite time direction can be divided into three categories. These strategies form a hierarchy, with each subsequent one being strictly better than the preceding ones. The first of these is the class of parallel processes, shown in (a), in which the gates U and V are probed in parallel with different parts of a potentially entangled state ρ, after which a generalized measurement $\mathcal{M}$ is performed. The next strategy is a casually ordered one, shown in (b). In strategies of this type the two gates are accessed one after another, and may be separated by another quantum channel $\mathcal{E}$. (c) The third and most general strategy with a definite time direction is one which is described by a process matrix $\mathcal{W}$, without any additional constraints. Such a strategy may therefore exhibit an indefinite causal structure, and can even violate causal inequalities.

For the strategy that uses the gates in parallel the success probability was found to be bounded by

$$\frac{88}{100} \leq p_{\text{par}} \leq \frac{89}{100}, \tag{6.1-34}$$

and this is the strategy with the lowest success probability. The second strategy is the ones which uses the gates sequentially, and the bound for this strategy was evaluated to be

$$\frac{90}{100} \leq p_{\text{seq}} \leq \frac{91}{100}. \tag{6.1-35}$$

Finally, the success probability of a player employing a strategy with an indefinite causal order has a success probability bounded by

$$\frac{91}{100} < p_{\text{i.c.}} \leq \frac{92}{100}. \tag{6.1-36}$$

Each subsequent strategy is strictly superior to the previous one, but even the most general strategy with a definite time direction cannot perform the discrimination task deteministically. Observing a success probability greater than the upper bound in (6.1-36) therefore certifies that the process in question has an indefinite time direction. It is also worth noting that the bounds above do not make any assumptions about the specific measurements used, and can therefore be used to certify an indefinite time direction in a semi-device independent way. Specifically, in a way that is independent of the measurement device, as will be discussed in Section 6.3.1.

6.2 Photonic implementation of the quantum time flip

While the quantum time flip cannot be realised within the quantum circuit model, similarly to the other non-quantum-circuit-model processes discussed in Section 6.1, it

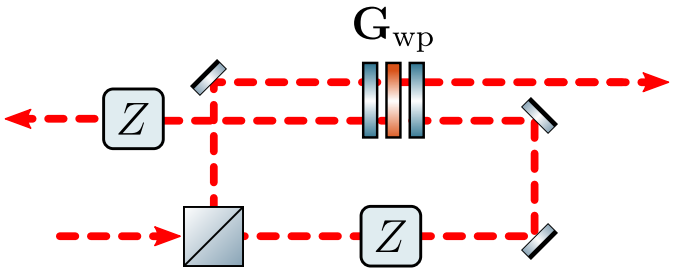


Figure 6.7: **Correcting a gadget transformation.** In the typically chosen convention for polarization states a wave-plate polarization gadget $\mathbf{G}_{\mathrm{wp}}$ transforms as $Z\mathbf{G}_{\mathrm{wp}}Z$ under counterpropagation. By performing a Z operation before and after the gadget one can recover the transpose, however when coherently superimposing two propagation directions through the gadget this requires the ability to act differentially on the two propagation directions.

still admits experimental realisations. More concretely, one can realise the bidirectional boxes pictured in Fig. 6.3 as physical devices, and use a control degree of freedom to set the direction in which they are accessed. In Chapter 3 it was shown that these boxes can realised as operations acting on the polarization degree of freedom of single photons, with the control degree of freedom being the propagation direction through the polarization gadget.

6.2.1 Bidirectional polarization gadget

As discussed in Section 6.1.1, the time-reversal operator was chosen to be the transpose. To realise the quantum time flip on the polarization degree of freedom one therefore needs a polarization gadget that tranforms as the transpose under counterpropagation, so that the implemented unitary in one propagation direction is the transpose of the unitary in the other direction. From the discussion in Chapter 3 we know that a polarization gadget consisting of linear retarders transforms as

$$\mathbf{G}_{\mathrm{wp}} = U \mapsto Z(\pi)U^T Z(-\pi) = ZU^T Z, \tag{6.2-1}$$

under counterpropagation. The unitary in the backwards direction is related to the transpose, but is sandwiched by two Pauli operators. This can be corrected by performing additional operations conditional on the control qubit:

$$\mathbf{G}_{\mathrm{wp}} \otimes |0\rangle\langle 0|_C + Z\Theta\big[\mathbf{G}_{\mathrm{wp}}\big]Z \otimes |1\rangle\langle 1|_C\,, \tag{6.2-2}$$

as shown in Fig. 6.7. This approach works, and the only downside is conceptual: the polarization devices in the two propagation directions are now physically distinct, and can in principle be fully decoupled from one another. Another option is to use the transpose gadget described in Chapter 3:

$$\begin{aligned}\mathbf{G}_T &= \mathbf{Z}^S(\alpha)\mathbf{Y}^A(\beta)\mathbf{Z}^S(\gamma)\\ &= \mathbf{Q}(\pi/4)\mathbf{H}(\alpha/4-\pi/4)\mathbf{Q}(\pi/4)\mathbf{Q}(-\pi/4)\\ &\times \mathbf{F}_-\mathbf{Q}(\pi/2)\mathbf{H}(\beta/4)\mathbf{Q}(\pi/2)\mathbf{F}_+\\ &\times \mathbf{Q}(\pi/4)\mathbf{Q}(\pi/4)\mathbf{H}(\gamma/4-\pi/4)\mathbf{Q}(\pi/4).\end{aligned} \tag{6.2-3}$$

This gadget has the correct symmetry properties, and therefore realises the transpose without any conditional correction operations. The obvious downside, however, is the increased number of optical components. While the waveplates on either side of the Faraday rotators can be compressed to single quater- half-wave plate pairs, a total of nine optical elements is still required. In practice, a higher number of elements typically results in lower fidelity operations, since the imperfections present in every element compound. A gadget with fewer components would therefore be desirable.

One approach to achieve this is to not undo the Z operations sandwiching the unitary using controlled operations that only act on one path degree of freedom, but to instead perform a global transformation that generates the correct symmetry. This can be done with the following basis transformation:

$$\begin{aligned} U_{\text{fw}} &\mapsto Z(\pi/2)U_{\text{fw}}\, Z(-\pi/2) = V_{\text{fw}} \\ U_{\text{bw}} &\mapsto Z(\pi/2)U_{\text{bw}}Z(-\pi/2) = V_{\text{bw}}. \end{aligned} \tag{6.2-4}$$

The basis transformation changes the unitaries, however this can be corrected for by simply choosing the pre-transformed unitary as

$$U_{\text{fw}} = Z(-\pi/2)UZ(\pi/2), \tag{6.2-5}$$

where U is the unitary one wants to preform the time flip on. It is straightforward to verify that the basis transformation above leads to the correct symmetry properties:

$$\begin{aligned} V_{\text{bw}} &= Z(\pi/2)U_{\text{bw}}Z(-\pi/2) \\ &= Z(\pi/2)\big[Z(\pi)U_{\text{fw}}^{T}Z(-\pi)\big]Z(-\pi/2) \\ &= Z(3\pi/2)U_{\text{fw}}^{T}Z(-3\pi/2) \\ &= \big[Z(\pi/2)U_{\text{fw}}Z(-\pi/2)\big]^{T} \\ &= V_{\text{fw}}^{T}. \end{aligned} \tag{6.2-6}$$

The basis transformation is essentially equivalent to swapping the X and Y eigenstates. This hints at the fact that it is not actually necessary to perform the basis transformation. One can instead simply redefine the polarization states by picking a different convention for the Pauli matrices and Stokes parameters. In a polarization gadget consisting of linear retarders it is always the part of the unitary acting on the diagonal polarization components that is anti-symmetric under counterpropagation:

$$\big[\theta \quad n_x \quad n_y \quad n_y\big] = \big[\theta \quad \nearrow \quad \circlearrowleft \quad \rightarrow\big] \rightarrow \big[\theta \quad \nwarrow \quad \circlearrowleft \quad \rightarrow\big] = \big[\theta \quad -\nearrow \quad \circlearrowleft \quad \rightarrow\big]. \tag{6.2-7}$$

Here the symbols on the right-hand side represent the polarization states that are the eigenstates of the corresponding component of the unitary. To get the correct transformation under counterpropagation one can simply choose to associate diagonal polarization states with the eigenstates of Y:

$$(S_1, S_2, S_3) = (\rightarrow, \nearrow, \circlearrowleft) \leftrightarrow (Z, Y, X). \tag{6.2-8}$$

This particular convention, which was initially picked for the experiment, is slightly inconvenient because it corresponds to a left-handed coordinate system. It was therefore later, after the data acquisition, changed to:

$$(S_1, S_2, S_3) = (\rightarrow, \nearrow, \circlearrowleft) \leftrightarrow (-Z, -Y, -X), \tag{6.2-9}$$

which when going from (6.2-8) is equivalent to adding a minus sign to the $\mathbb{1}$ term in a given unitary. In practice this simply means that the last two unitaries in the set $\mathcal{M}_-$ get exchanged with one another, while the others remain unaffected. One can also verify through direction calculation that this choice of conventions causes linearly birefringent gadgets transform as the transpose. Under the convention (6.2-9) a single linearly birefringent element with retardance η at an angle ϕ to the vertical axis can be written as

$$\mathbf{W}_\eta(\phi) = X(2\phi)\mathbf{W}_\eta(0)X(-2\phi), \tag{6.2-10}$$

since the Z and Y eigenstates now represent linear polarization states. The transformation of a single-element gadget under counterpropagation can then be calculated:

$$\begin{aligned}\Theta\big[\mathbf{W}_\eta(\phi)\big] &= \mathbf{W}_\eta(-\phi) \\ &= X(-2\phi)\mathbf{W}_\eta(0)X(2\phi) \\ &= \big[(X(2\phi)\mathbf{W}_\eta(0)X(-2\phi)\big]^T \\ &= \mathbf{W}_\eta(\phi)^T,\end{aligned} \tag{6.2-11}$$

and we see that it indeed transforms as the transpose. From this it immediately follows that sequences of linear retarders also transform as the transpose:

$$\begin{aligned}\Theta\big[\mathbf{G}_{\text{linear}}\big] &= \Theta\big[\mathbf{W}_{\eta_1}(\phi_1)\mathbf{W}_{\eta_2}(\phi_2)\cdots\mathbf{W}_{\eta_N}(\phi_N)\big] \\ &= \mathbf{W}_{\eta_N}(-\phi_N)\cdots\mathbf{W}_{\eta_2}(-\phi_2)\mathbf{W}_{\eta_1}(-\phi_1) \\ &= \big[\mathbf{W}_{\eta_1}(\phi_1)\mathbf{W}_{\eta_2}(\phi_2)\cdots\mathbf{W}_{\eta_N}(\phi_N)\big]^T \\ &= (\mathbf{G}_{\text{linear}})^T.\end{aligned} \tag{6.2-12}$$

To verify that the above gadget actually works, the unitaries in the sets (6.1-19) and (6.1-20) were implemented using two gadgets, and were then reconstructed using quantum process tomography restricted to unitary channels. The first (second) unitary U (V) in each pair was implemented with the gadget used to implement the same unitary in the experimental channel discrimination task. All 21 unitaries for each gadget were first measured in one propagation direction, and then independently in the other propagation direction. This was done for practical reasons, as moving the tomography setup between each measurement would have been too time consuming, but had the downside that the unitaries measured in the two measurement sets could be slightly different due to small differences in the wave-plate angles.

The unitaries were reconstructed from the measurement data using a maximum likelihood estimation, and were then used to calculate the fidelity between the unitary in the forwards propagation direction and the transpose of the unitary in the backwards propagation direction:

$$\mathcal{F}(U_{\text{fw}}, U_{\text{bw}}^T) = \big\langle (U_{\text{fw}}\,|\Psi\rangle)^\dagger U_{\text{bw}}^T\,|\Psi\rangle\big\rangle_{|\Psi\rangle}. \tag{6.2-13}$$

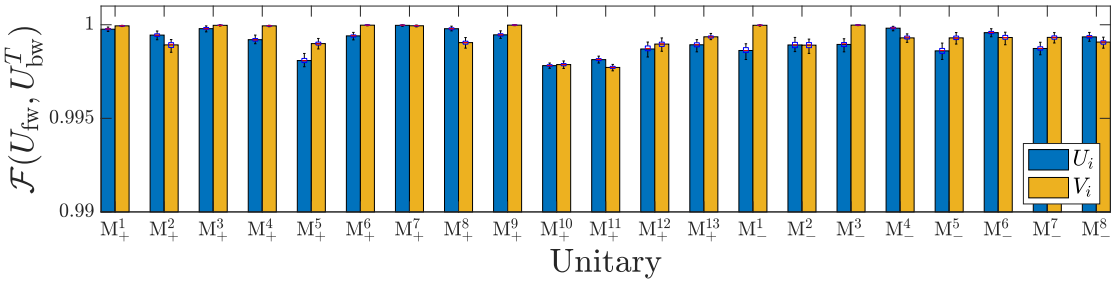


Figure 6.8: **Transpose fidelity.** Fidelities between the reconstructed unitary in the forwards direction and the transpose of the reconstructed unitary in the backwards direction, for all unitaries in the sets $\mathcal{M}_+$ and $\mathcal{M}_-$. The superimposed box plot shows the uncertainties in the fidelities, calculated using a Monte-Carlo simulation. The mean fidelity is 0.9992 ± 0.00064, indicating that the gadgets indeed have the correct symmetry properties.

This average was taken over 1000 Haar random states, and was evaluated as part of a Monte-Carlo simulation that accounted for wave-plate angle errors in the tomography. The distribution of the fidelities in the simulation was used as an estimate of the uncertainty in the reconstructed fidelities, as the wave-plate-motor repeatability was assumed to be the main error source. Tomography was also performed on the polarizers used to prepare the two input states $|H\rangle$ and $|+\rangle$ in the process tomography, and the corresponding reconstructed states were used in the maximum likelihood estimation of the unitaries.

The result of the tomographies is shown in Fig. 6.8, and both gadgets exhibit high fidelities. Note that these fidelities do not reflect how accurately the gadgets implement the unitaries U_i and V_i, only the symmetry properties of the gadgets under counterpropagation. This property is independent of errors in the retardance and angles of the wave plate. However, since as previously mentioned the two directions were measured independently, differences in the gadget wave-plate angles between the two measurements could have caused the fidelity to drop.

6.2.2 Interferometer design for the channel discrimination task

The bidirectional polarization gadgets described in the previous section were placed inside an interferometric setup designed to implement the quantum time flip strategy for the channel discrimination task presented in Section 6.1.2. The goal of the experiment was to demonstrate that the time direction of the resulting process was indefinite, by achieving a success probability for the channel discrimination strictly greater than that of any strategy with a definite time direction.

The fact that the quantum time flip can be realised by exploiting two different propagation directions through a single polarization gadget opens up the possibility of novel optical geometries. Our implementation made use of fiber-optic circulators to realise a partially common-mode interferometer, that used the non-reciprocal properties of the circulators to separate two overlapping and counterpropagating beams. The input-output relations of these devices is shown in Fig. 6.9. Overlapping the beams in the two propagation directions through the polarization gadgets ensures that the polarization transformation in one direction is exactly the transpose of the transformation in the

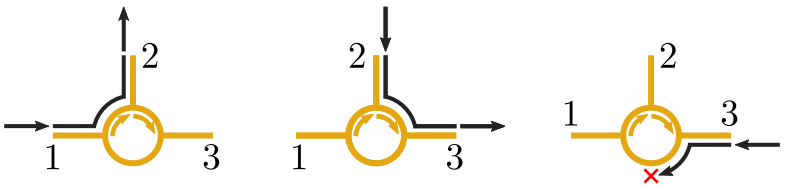


Figure 6.9: **Fiber circulator input-output relations.** The experimental setup made use of fiber circulators to superpose two different propagation directions, and to then separate these counterpropagating modes. These devices work as shown above. Light input in mode 1 gets routed to mode 2, while light entering in mode 2 is sent to mode 3. Finally, light entering in mode 3 is rejected, though this feature was not utilized in the experiment.

other direction, since the two beams sample the same spots on the wave plates. The mode overlap between the two beams was ensured by fiber coupling the light before and after each gate.

A sketch of the full experimental setup is shown in Fig. 6.10. The main part of the interferometer consists of a mixed bulk- and fiber-optic interferometer that implements the quantum time flip strategy for the channel discrimination task. The input beamsplitter in the top left of the figure prepares the state:

$$\frac{1}{\sqrt{2}} |\Psi\rangle_T (|0\rangle_C + |1\rangle_C), \tag{6.2-14}$$

where $|\Psi\rangle_T$ is a target polarization state, and the computational basis states of the control qubit are encoded in the photon path. In practice the beamsplitter does not prepare the state $|+\rangle_C$ since the reflectivity will in general not be perfectly 50 : 50, and typically also depends on the optical alignment. Additionally, the two control qubit states may experience differential loss throughout the setup, which can introduce additional imbalance between the two terms. Finally, the output beamsplitter will also exhibit some degree of imbalance in its splitting ratio. By choosing a bulk beamsplitter for the input of the interferometer the control state superposition can be balanced by adjusting the relative coupling efficiency into the single-mode fibers. This, in turns, allows the interferometric visibility to be greatly improved, compared to a setup without control over the non-common-mode loss.

After being initialized in a spatial superposition by the beamsplitter, two fiber circulators were used to send the target state through the gadget implementing the unitaries V_i in two different directions, thereby generating the join target-control state

$$\frac{1}{\sqrt{2}}\big(V |\Psi\rangle_T |0\rangle_C + V^T |\Psi\rangle_T |1\rangle_C\big). \tag{6.2-15}$$

Entering the circulators from what was previously the output port, the two photon paths were then routed to a second free-space path containing the gadget implementing the U_i gates. The two counterpropagating paths through this gadget were aligned with opposite relative directions with respect to the V gadget, meaning that the transposed unitary was applied on the term with the $|0\rangle_C$ component, giving the state

$$\frac{1}{\sqrt{2}}\big(U^T V |\Psi\rangle_T |0\rangle_C + U V^T |\Psi\rangle_T |1\rangle_C\big). \tag{6.2-16}$$

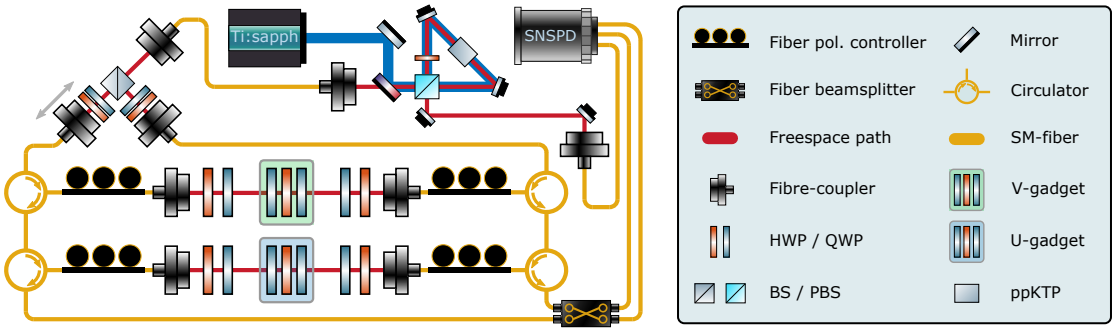


Figure 6.10: **Experimental setup.** The double-time-flip strategy for the channel discrimination task was implemented using an interferometer consisting of both bulk and fiber optics. In order to balance the two interferometer arms as well as control the interferometer phase, a bulk optics beamsplitter was used at the input, and the fiber coupler collecting the transmission port of the beamsplitter was placed on a linear translation stage. Adjusting the relative coupling of the two fiber couplers at the input also allowed for differential loss in the two interferometer arms to be balanced. The polarization gadgets implementing the unitary channels U and V were placed in two separate free-space paths, and four fiber circulators allowed the photons to be superposed in two different propagation directions through these two paths. A spontaneous parametric down-conversion source was used to generate single-photon pairs, and non-polarization-resolving detection was carried out with superconducting nanowire single-photon detectors (SNSPDs).

This is illustrated in Fig. 6.11, which shows the two photon paths through the interferometer. Finally, a second beamsplitter was used to perform a Hadamard operation on the control qubit, thereby correlating its computational basis states with the defining properties of the sets $\mathcal{M}_{\pm}$:

$$\begin{aligned} &\frac{1}{2}U^T V\left|\Psi\right\rangle_T\left(\left|0\right\rangle_C+\left|1\right\rangle_C\right)+\frac{1}{2}UV^T\left|\Psi\right\rangle_T\left(\left|0\right\rangle_C-\left|1\right\rangle_C\right)\\ =&\frac{1}{2}\left(U^T V+UV^T\right)\left|\Psi\right\rangle_T\left|0\right\rangle_C+\frac{1}{2}\left(U^T V-UV^T\right)\left|\Psi\right\rangle_T\left|1\right\rangle_C \end{aligned} \tag{6.2-17}$$

This Hadamard transformation at the output of the interferometer was implemented in fiber in order to simplify the spatial mode matching, and to thereby help facilitate a high interferometric visibility. The two output ports of the interferometer were connected to single-photon detectors and no polarization resolving measurements were performed. Formally, the target state was therefore traced out. The superconducting single-photon nanowire detectors (SNSPDs) used in the experiment did have a slight polarization dependence, however, as will be discussed later this did not have to be accounted for.

6.2.3 Path-length balancing

In addition to balancing the interferometer loss, the bulk beamsplitter at the input served the purpose of allowing for control over the relative path length of the two optical paths through the setup. This was achieved by placing the fiber couplers collecting the light in the transmitted port of the beamsplitter on a linear translation stage. Since the range of the translation stage was relatively limited, around 20 mm, the lengths of the fibers in the non-common-path sections of the interferometer were carefully measured, and these fibers were connected in a configuration that minimized the relative

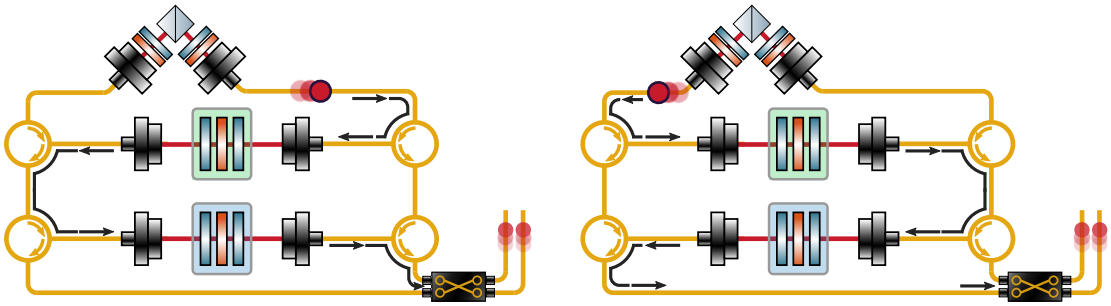
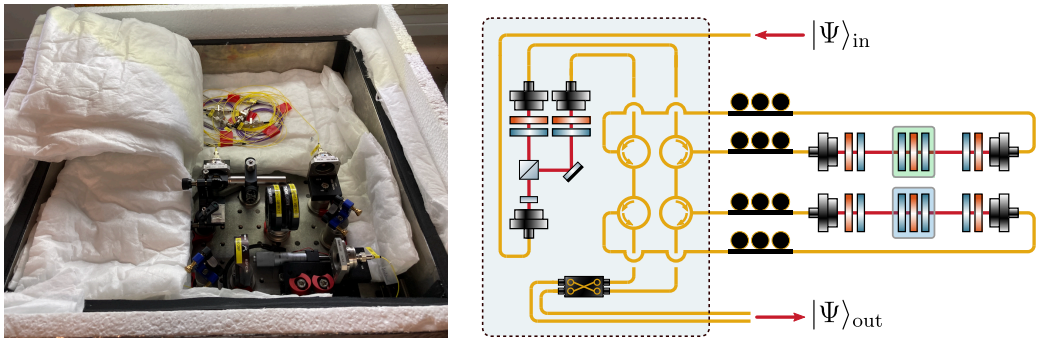


Figure 6.11: **Photon paths.** The reflected and transmitted ports of the input beamsplitter of the interferometer propagate along two different paths that partially overlap. The reflected port (left) enters the upper-right fiber circulator in port 1 and propagates through the first polarization gadget from right to left, yielding the transformation V^T. After this, the two left-hand circulators route it through the second polarization gadget in the opposite direction, which applies U. Conversely, the transmitted port of the input beamsplitter is sent to the upper-left circulator, and propagates through the first gadget left to right, which causes it to apply the unitary V. The two right-hand circulators then route photons in this path through the second gadget backwards, and consequently U^T is applied.

path-length difference. The residual path-length difference was pre-adjusted by placing the fiber coupler for the reflected port of the beamsplitter at the appropriate distance, estimated from the measured fiber lengths. After this, the path length was coarsely minimized by using a diode (ThorLabs L1550P5DFB) with a linewidth of approximately 0.1 nm, corresponding to a coherence length of about 7.5 mm, slightly exceeding that of the single photons. The minimization was done by manually scanning the relative optical path length while observing the interferometric visibility by sweeping the interferometric phase with a piezo-electric actuator, also connected to the linear translation stage.

While the coherence length of the diode laser was in principle short enough to measure the path-length difference accurately enough to ensure high visibility between the single photons, the resolution of the measurement was limited by back reflections from the various optical components as well as fiber connectors. While all the free-space optics in the experiment were anti-reflection coated individual components can still have reflectivities on the order of tens of a percent. When attempting to resolve a very low interferometric minimum these small signals are sufficient to ruin the measurement. Due to the relatively short coherence length of the diode laser, one would expect most backreflections to simply form an incoherent background that could be subtracted, but instead a highly erratic signal near the interferometric minimum was observed. One explanation of the interference exhibited by the backreflections could be signals resulting from two double backreflections from the same components in the common-path parts of the setup, as these would preserve the relative path-length difference. Since these signals involve two back reflections though, they should be even smaller in magnitude.

The underlying cause of this behaviour was never investigated in detail, since it is challenging to observe the back reflections in isolation when using continuous wave light, and ultimately it was not necessary to perform the full path-length balancing using the diode laser. Instead, after sufficiently high visibility was achieved the remaining



(a) Photograph of opened box (b) Distribution of optical components

Figure 6.12: **Interferometer box.** The interferometer was stabilized using two nested boxes, as shown in the photograph in (a). The inner box contained the bulk beamsplitter used at the input, the associated fiber couplers, a translation stage used to control the path-length difference between the interferometer arms and to set the interferometer phase, a polarizer and wave plates for polarization compensation (one HWP not pictured), as well as the fiber ciruclators and output fiber beamsplitter. The fiber components were placed between four sheets of Thinsulate in order to avoid direct thermal contact with the metal box. Since the fibers were not rigidly secured, slight polarization disturbances could be observed when opening and closing the box, but these were not enough to need addressing. (b) shows which optical components were placed inside the box (light blue) for stabilization, and which ones were placed outside for easier access.

path-length imbalance was corrected by observing the single-photon interference in the interferometer, while conditioning the signal on the heralding photon from the SPDC source. Since the back reflections will take a different optical path than the main signal, and therefore have a different time delay with respect to the heralding photon, this time filtering eliminates any back reflections arriving outside a time window defined by the jitter of the detectors. The SNSPDs used in the experiment had a jitter of approximately 1 ns. This corresponds to an optical path length of 30 cm, and the separation between the fiber couplers surrounding the polarization gates was about 60 cm. This means that any back reflection traversing a free-space path twice, or not at all, will be filtered out. Back reflections occurring in the middle of the free-space path would not get filtered out. However, to the extent that these were present they did not significantly degrade the measurements, and single-photon interference visibilities of around 0.998 were observed.

6.2.4 Interferometer stabilization

The interferometric setup described in the previous section is not fully common path, and is therefore sensitive to path-length fluctuations that translate into phase noise. Optical fibers, in particular, add a significant amount of instability to interferometers unless actively or passively stabilized. Active stabilization of the interferometer would present the problem that interference condition of the locking light is affected by the polarization transformations in the free-space parts. This is not ideal, because the laser light is in effect performing a classical measurement on the gates in the channel, while the channel discrimination task is formulated such that only a single use of each gate is allowed. A passive stabilization approach was therefore pursued instead.

The passive thermal and acoustic stabilization was done using two nested boxes.

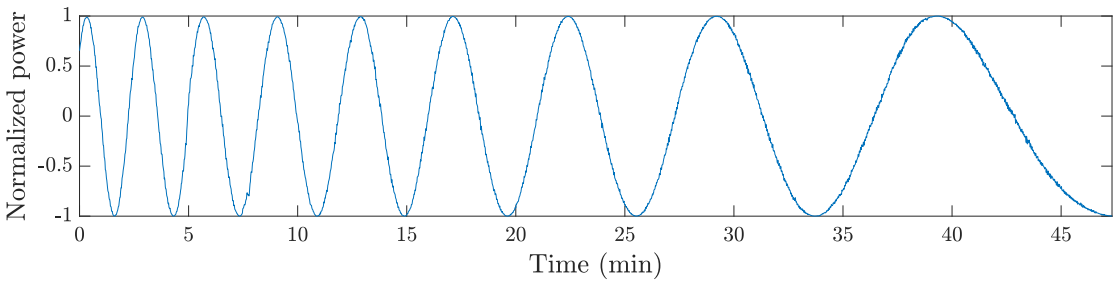


Figure 6.13: **Interferometer thermalization.** Immediately after closing the box used to stabilize the phase-sensitive parts of the interferometer the observed passive phase drift is very large. However, this behaviour is transient and decays exponentially, until the interferometer settles. In the figure above the transient manifests as a frequency chirp in the sine curve.

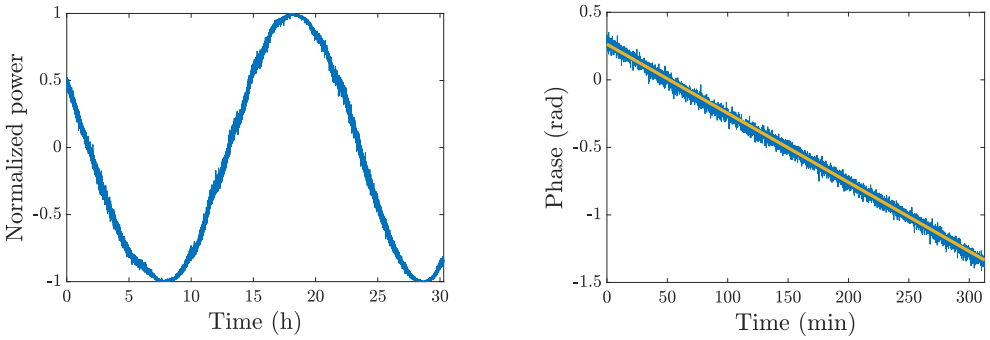
Optics were housed inside an inner box consisting of a steel optical breadboard, and aluminium walls and lid lined with sorbothane. This box was placed on top of a set of sorbothane vibration-absorbing feet. A larger polystyrene box was placed around the inner box, and Thinsulate was used to fill the air gap between the two boxes. Thinsulate was additionally used as padding inside the inner box, as well as to cover the outside of the polystyrene box. In Fig. 6.12 a picture of the opened box is shown.

One downside of housing the experiment inside an insulated box is that the optical components cannot be readjusted without having to open the box, and then subsequently waiting for it to re-thermalize. Fortunately, due to the partially common-path geometry of the interferometer many of the optical components could actually be placed outside the box without compromising the phase stability. As shown in Fig. 6.12b only the two beamsplitters, as well as the four fiber circulators were placed inside the box, while the four fiber polarization controllers and the two free-space optical paths containing the polarization gadgets were placed outside the box.

The performance of the passive stabilization is illustrated in Fig. 6.13 and Fig. 6.14, which show the initial transient behaviour of the box, and the steady state behaviour, respectively. Given sufficiently long time to thermalize, the observed drift rate was as low as $5\,\mathrm{mrad\,min^{-1}}$, which was more than sufficient for the experiment. For the actual single-photon measurements though, the box was not allowed to thermalize for this long. This was due to the fact that a very slight decrease in the polarization contrast over time was observed. Albeit small, this effect was nevertheless sufficient to measurably influence the success probabilities in the channel discrimination task negatively. Since, as will be discussed in Section 6.2.5, the polarization compensation procedure required the box to be opened, the compensation could not be tuned up after the box had been allowed to thermalize. Instead, a compromise between the stability of the interferometer and the polarization contrast was found, and the time spent with the box open was minimized.

6.2.5 Polarization compensation

The basic principle of the polarization compensation in the experiment is the same as described in Section 5.2.3, and will not be restated here. Due to the slightly unconventional geometry of the interferometer, however, a few extra considerations had to be made. In particular, since the polarization compensation was performed using coher-



(a) Power drift (b) Linear phase drift

Figure 6.14: **Interferometer phase drift.** After thermalizing, the interferometer displayed significantly lower phase drift than in the initial transient period. As can be seen in (a), a drift of less than 2π radians was observed a period of 30 h. In (b) the drift is quantified with a linear fit to a 300 min subset of the data in (a), shown in yellow, and the drift is found to be 5 mrad min^{-1}. This drift rate only quantifies the DC noise, resulting from temperature gradients between the inside and outside of the box. The full power spectral density of the thermal noise was not measured, and the passive temperature stabilization does not reduce the intrinsic thermal noise in the fibers.

ent light, the presence of back reflections was found to limit the resolvable polarization contrast. Consider, for example, compensating the path $A \rightarrow M$ in Fig. 6.15. Light propagating along the path $A \rightarrow B \rightarrow M$ will interfere with the light in the path to be compensated, and this light could come from for example back reflections in the path $B \rightarrow A$, or light entering the circulator adjacent to A from the other port. Without addressing these effects, the polarization contrast was limited to less than 30 dB, and in some cases much lower, due to a mixed polarization state at the measurement polarizer.

In order to mitigate the effects of the back reflections the state-preparation polarizers were supplemented with an additional polarizer as well as a Faraday rotator, in order to construct a makeshift Faraday isolator. In addition to this, it was observed that the linear polarizers used (ThorLabs LPNIR050) reflected most of the rejected light, and in some cases this resulted in double back reflections, for example from the polarizer and the fiber coupler in front of which it was placed, strong enough to significantly degrade the achievable polarization contrast. The polarizers were therefore placed as far away from the fiber couplers as possible, and were slightly tilted in order to misalign the back reflection.

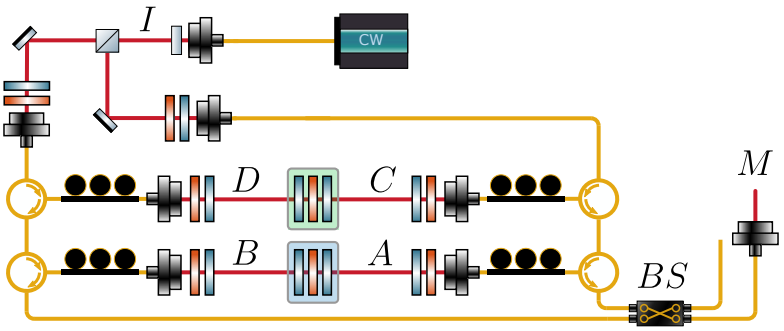


Figure 6.15: **Polarization compensation.** In the experiment the polarization of the photons was not resolved as part of the photodetection, and therefore it was technically speaking not necessary to compensate the polarization at the output of the interferometer. It was, however, still necessary to compensate the polarization at the second fiber beamsplitter (BS) where interference takes place. Since the polarization at the beamsplitter cannot be directly measured one of the output fibers was connected to a fiber coupler, at the point M in the figure. While compensating the fibers in the paths $A \to M$ and $B \to M$ does not ensure that the paths $A \to BS$ and $B \to BS$ realise the identity transformation, it does guarantee that the polarization transformations of the two paths are identical. When compensating for example the path $A \to M$ it was possible, due to the partially common-path geometry, for back reflections of the form $A \to B \to M$ or $C \to A \to B \to M$ to interfere with the measurement signal, and this significantly limited the resolvable polarization contrast. In order to combat this problem makeshift optical isolators, in place of linear polarizers, were used to prepare the input polarization states. The probe light for the compensation was injected using a CW laser at the input side of the interferometer. The compensation was carried out in a backwards order, starting with the path $A \to M$ ($B \to M$) using a combination of the wave plates and paddles at A (B). After this $C \to A$ ($D \to B$) was compensated the same way, and finally $I \to C$ ($I \to D$). A linear polarizer at I was used to prepare the input state for both the compensation and the experiment, and the paths connecting the input to the first gadget therefore only had to be compensated in one basis. This single-basis compensation was performed using the two wave plates in front of each input coupler.

6.3 Data acquisition and analysis

For each of the 21 pairs of unitaries in $\mathcal{M}_{\pm}$ the wave-plate gadgets were first set to implement the identity operation so that the phase of the interferometer could be monitored, and if needed reset. After zeroing the interferometer phase, the gadgets were set to implement unitaries U (V) and data was acquired for 5 s. This short acquisition time was enabled by the relatively low experimental loss. After the data acquisition, time-ordered lists of two-fold coincidences between signal and trigger photons were calculated using code re-purposed from Publication 2. Due to the polarization dependence in the single-photon detectors the total number of photon counts for each setting varied by as much as a factor of two, and the data sets for all settings were truncated to the first 55 000 detection events.

Using this data set the channel discrimination game was played. In every round of the game a referee randomly picks one of the 21 pairs of channels, and the player attempts to guess from which set they were picked. The referee's choices were sampled uniformly randomly, and the player's guesses were picked from the measured data. No measurement outcomes were used more than once, and in each round the player's outcome was given by the first unused detection event in the list of events corresponding

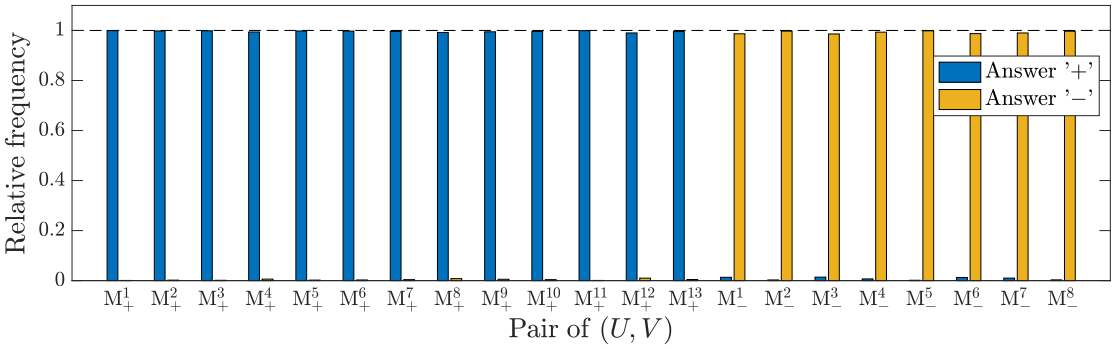


Figure 6.16: **Relative frequencies.** The figure shows the relative frequency of observed outcomes, '+' or '−', for the 21 pairs of channels the referee picked from in the game. For the set $\mathcal{M}_+$ ($\mathcal{M}_-$) obtaining the outcome '+' ('−') wins the game, and it can be seen that the wining frequency is close to 1 for every choice of channels. The average winning frequency observed in the game was 0.9945, far exceeding the bound of 0.92 for strategies with a definite time direction, while the highest and lowest winning frequencies were 0.9993 and 0.9860, respectively.

to the pair of channels that the referee picked. The exact procedure for game was the following:

Let k be the number of rounds played, n_i the number of times that the referee has picked the channels (U_i, V_i) in the previous rounds and $\mathcal{O}_i^m$ the m-th entry in the time-ordered list of detection events corresponding to the channels (U_i, V_i). Then the outcome of the k-th round of the game was determined by retrieving the event $\mathcal{O}_i^{n_i}$.

While this game could also have been implemented by randomly sampling from the unitaries during the experiment and implementing a different pair of channels for every round of the game, the choice of instead playing the game using the acquired data was taken in order to not have the data acquisition time be limited by the time spent rotating the wave plates.

One million rounds of the game were played using the procedure described above. This number was chosen to be high, while still ensuring that there would be a sufficient number of detection events for every choice of unitaries even if the referee's random picks resulted in a slightly non-uniform distribution. During the game the player outputs an answer every round, and we are interested in the relative frequency with which the player gives the two possible answers. Let N_k^+ (N_k^-) be the number of times that the player output the answer + (−) in a round when the referee picked the pair of unitaries (U_k, V_k), and N_k the total amount of times that those unitaries were picked, then the relative frequencies of the two outcomes '+' and '−' are:

$$f_k^+ = \frac{N_k^+}{N^k}, \quad f_k^- = \frac{N_k^-}{N^k}. \tag{6.3-1}$$

In Fig. 6.16 the observed relative frequencies for all of the 21 possible choices of unitary channels by the referee are shown. It can be seen that all the relative success frequencies far exceed the bound of 0.92 on the success probability for processes with definite time directions, with the highest (lowest) observed relative win frequency being 0.9993 (0.9860).

The relative frequencies do not have an uncertainty associated with them, since they are given directly by how many times each outcome was observed. One way to statistically analyse the result would be to interpret the relative frequencies as empirical estimates of the underlying success probabilities for every choice of channels by the referee. These estimates would have an uncertainty given by Poissonian counting statistics, and this uncertainty could be propagated to find an uncertainty in the mean success probability. However, a more suitable way to quantify the statistical significance of the results is to calculate the p-value for the process having an indefinite time direction, which is the probability of a process with a definite time direction having generated at least as many wins as observed. This probability can be expressed as

$$P = \sum_{k=v}^{N} \binom{N}{k} p^k (1-p)^{N-k}. \tag{6.3-2}$$

Here $p = 0.92$ is the highest success probability for a strategy with a definite time direction, and $v = 994,512$ is the number of experimentally observed wins in the game. The binomial coefficient in the expression above counts all the different ways in which the k wins can be distributed over the N rounds. Due to the large number of rounds played the binomial coefficients become extremely large, while the exponentiated probabilities tend to zero, and this makes the expression difficult to evaluate numerically. Instead of calculating the exact p-value, one can use a Chernoff bound to calculate an upper bound on the p-value [430]. This bound can be expressed as

$$P \leq \exp\Big[-N D\Big(\frac{v}{N}\Big\|p\Big)\Big], \tag{6.3-3}$$

where

$$D\Big(\frac{v}{N}\Big\|p\Big) = \frac{v}{N}\ln\Big(\frac{v}{Np}\Big) + \Big(1-\frac{v}{N}\Big)\ln\Big(\frac{1-v/N}{1-p}\Big), \tag{6.3-4}$$

is known as the Kullback-Leibler divergence. Evaluating the expression above gives

$$D\left(\frac{994,512}{10^6}\middle\|0.92\right) \approx 0.06275, \tag{6.3-5}$$

and hence

$$P \leq e^{-6\cdot 10^4}. \tag{6.3-6}$$

This is a very small number, and one can thus conclude that a process with a definite time direction could not achieved as many wins as the experimentally realised process, which must therefore have an indefinite time direction. As a complement to the above bound, the p-value was also calculated explicitly for smaller numbers of rounds and then extrapolated. This was done by exploiting the fact that the natural logarithm of a binomial coefficient can be expressed as

$$\ln\binom{N}{k} = \ln\Gamma(N+1) - \ln\Gamma(k+1) - \ln\Gamma(N-k+1), \tag{6.3-7}$$

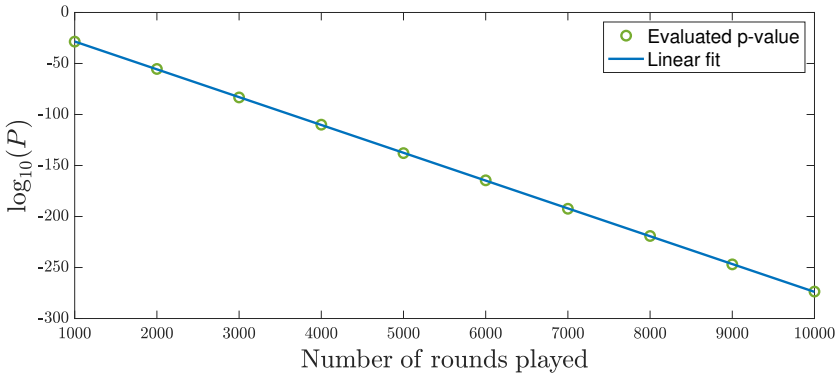


Figure 6.17: **P-values.** Since the p-value could not be directly calculated for the $N = 10^6$ rounds that were played, it was instead evaluated for lower numbers of rounds and then extrapolated. This was done for round numbers up to 10,000, and as shown in the figure the p-value decrease exponentially with the round number. The data points in the figure were calculated using a number of wins given by the average success probability 0.994512, not by sampling number of wins for a given number of rounds. Extrapolating the linear fit to $N = 10^6$ gives a p-value of $P = \exp[-6.279 \cdot 10^4]$.

where $\Gamma(\cdot)$ is the gamma function. The p-value can then be evaluated for relatively large round numbers as

$$P = \sum_{k=v}^{N} \exp\left[\ln\binom{N}{k} + k\ln p + (N-k)\ln(1-p)\right]. \tag{6.3-8}$$

The results of this calculation are shown in Fig. 6.17 for up to 10,000 rounds. Extrapolating from a linear fit to the logarithm of the p-values one finds a p-value of $\exp[-6.279 \cdot 10^4]$ for $N = 10^6$, suggesting that the bound (6.3-6) is quite tight.

6.3.1 Semi-device independence

The data analysis outlined above did not account for imbalanced or polarization-state dependent detection efficiencies, which were known to be present in the setup. The reason for this is that the certification method for the indefinite time direction was semi-device independent; specifically, it was independent of the measurement device. This device independence is a consequence of the theoretical methods used to device the bounds on the success probabilities for strategies that do not have access to coherent unitary transposition. More concretely, since the tester formalism encompasses the measurement performed by the player, the bounds derived using this formalism represent an optimization over all possible measurements the player can perform. As a result, there is no possible measurement that a player using a strategy from a given class could perform that would enable them to violate the bound derived for that class. This in turn means that if a winning rate exceeding the bound for a class of strategies is observed, then then one does not need to know what measurement was performed in order to conclude that the experimentally implemented process does not belong to the class in question. This notion of semi-device independence is taken from [431], where it was applied to the certification of indefinite causal orders, leading to the first semi-device independent demonstration of an indefinite causal order [432].

The contrast between such a demonstration and previous device dependent ones is best illustrated through an example. In Ref. [11] a causal witness was used for the first certification of an indefinite causal order. This was achieved by constructing a witness operator $\hat{S}$ that had the property that

$$\mathrm{Tr}\big[\hat{S}W^{\mathrm{sep}}\big] < 0, \tag{6.3-9}$$

for any causally separable process matrix W^{sep}. Experimentally evaluating $\mathrm{Tr}\big[\hat{S}W^{\mathrm{exp}}\big]$ and recording a value greater than zero therefore certifies that the experimental process W^{exp} is causally indefinite. In practice this is done by decomposing $\hat{S}$ in terms of an experimentally realisable set of observables,[2] and then experimentally evaluating the expectation values of these observables. However, the estimates of the expectation values carry with them uncertainties resulting from measurement imperfections, thereby introducing a corresponding uncertainty in the estimate of $\mathrm{Tr}\big[\hat{S}W^{\mathrm{exp}}\big]$. In order to observe a violation of (6.3-9) that is statistically significant, it is therefore imperative to have well characterised measurement devices.

Returning to the semi-device independent demonstration of an indefinite time direction in Publication 3, we note that the imbalanced detection efficiencies present in the SNSPDs can be modelled as generalized measurements. For example, a biased measurement of the control qubit could look like:

$$\begin{aligned} E_1 &= p_1 \left|0\right\rangle\!\left\langle 0\right|_C + p_2 \left|1\right\rangle\!\left\langle 1\right|_C, \\ E_2 &= (1-p_1) \left|0\right\rangle\!\left\langle 0\right|_C + (1-p_2) \left|1\right\rangle\!\left\langle 1\right|_C, \end{aligned} \tag{6.3-10}$$

where $p_1, p_2 \in [0,1]$ represent the detector bias. This description gives a more concrete understanding of how such measurement imperfections are considered in the derivation of the success probability bounds. Accounting for measurement bias could likely have yielded a larger violation, or higher rate of wins in the game, however this was not necessary since the recorded data already certified the indefinite time-direction, and modelling the detection device would have led to a less elegant result.

6.3.2 Discussion on number of gate uses

The advantage of the quantum time flip in the channel discrimination task hinges on the restriction that the player can only access each gate once. When being given access to more uses of the gates new strategies open up. For example, when given access to four uses of the same unitary, there exists a circuit to deterministically transform this gate into its transpose [433], and when a high number of gate uses is allowed a player could simply perform quantum process tomography to determine what channels the referee picked. In the photonic implementation one would intuitively expect that the gates are only used once, since the two photon paths only pass through the gates once, however it is not straightforward to actually verify this. This issue was first discussed in the context of the query complexity of causally indefinite processes, and in Ref. [419] the authors proposed using an additional system as a counter of the number of gate uses, however

[2] In the case of the quantum switch these observables consist of an input state, the operations inside the quantum switch, and the actual measurements performed on the output system.

this approach has not yet been rigorously formulated. A different approach, based on quantum process tomography of so-called time-delocalized subsystems, was put forth in Ref. [434].

The point of such discussions is to elucidate whether or not the physical realisation of causally or temporally indefinite processes actually carry greater information processing power than physical realisations of quantum-circuit-model processes. One can then instead ask what the physically relevant resources are, and how they are consumed. Such an analysis was carried out in Ref. [435] by studying the energy cost of a process. The authors found that implementations of the quantum switch using a single realisation of each operation compared favourably to 'simulations' of the quantum switch using two realisations of each operation. However, the quantum time flip discussed in this chapter was applied to single qubit gates. These gates, by definition, only affect a single, two-level quantum system, and therefore implicitly require an interaction with a classical environment on which there is no back action. This in turn means that this classical resource is inexhaustible. Indeed, the transfer of angular momentum from a single photon to a polarization gadget implementing a single-qubit gate can be completely neglected, and one could just as easily apply the same transformation to 10^{15} photons in a single optical pulse in parallel. In contrast, the study performed in Ref. [435] considered single-qubit gates mediated by other quantum systems, and it was noted that such operations will in general be slightly entangling and therefore introduce non-unitary reduced dynamics on the target system.

To establish the existence of practically useful advantages of non-circuit-model processes one ultimately needs to identify a more experimentally relevant difference in terms of resource consumption. This would likely involve understanding how the requirements on the physical implementations of these processes scale when extrapolated to more complicated tasks on larger systems.

7

Background for "Demonstration of a Quantum Switch in a Sagnac Configuration"

This chapter will discuss Publication 4, "Demonstration of a quantum switch in a Sagnac configuration", but unlike the preceding chapters will not go into as great a detail about the experimental methods as many of the experimental methods have been covered in previous chapters. The purpose of the experiment was to test the theoretical results presented in Chapter 3 and to demonstrate the applicability of one of the polarization gadgets to a concrete task. The quantum SWTICH was a natural candidate for such a task, since it is usually realised in Mach-Zehnder geometries that are not amenable to active phase stabilization. At the end of the chapter it is discussed how the intrinsic phase stability can be extended to a quantum switch with more than two parties.

7.1 Optical implementations of the quantum switch

As discussed in Chapter 5 the quantum switch is a process that maps two gates, U and V, into a superposition of the two gates being applied in opposite orders, conditioned on the state of a control system C

$$(U, V) \mapsto UV \otimes |0\rangle\langle 0|_C + VU \otimes |1\rangle\langle 1|_C \,. \tag{7.1-1}$$

The relative ease with which single photons can be placed in spatial superpositions has made photonics the simplest platform for realising the quantum switch, and to date all realisations of the process have relied on different single-photon encodings of quantum information [436].

7.1.1 Path-polarization quantum switch

The first, and most common, type of realisation of the photonic quantum switch is one in which the polarization degree of freedom is used to encode the target system, and the path degree of freedom is used to encode the control degree of freedom [11, 406, 437, 438, 439, 440]. In practice this is typically done by using polarization optics placed inside a Mach-Zehnder interferometer, aligned in such a way that the photons propagate through the polarization gadgets in different orders in the two different interferometer arms, as shown in Fig. 5.3. The advantage of this type of implementation is that high fidelity single-qubit polarization transformations can be realised with relative ease, while

the typically fixed control-qubit operation is realised using the harder to control path degree of freedom.

The main challenge with this type of implementation is that the interferometer phase needs to be stabilized. Passive stabilization has the problem of limiting the duty cycle of the setup, as the phase needs to be periodically checked and reset [11], and this also presents a trade-off between the available data collection time and the fidelity of the device operation. On the other hand, active phase locking of interferometers, through for example Pound-Drever-Hall locking [441, 442], is a well established technique that can enable very high phase stability. While there is no fundamental problem preventing the use of these methods in the realisation of photonic quantum switches, there is a practical problem in that the locking light cannot propagate through the polarization gadgets inside the inferferometer. This is because the polarization transformations affect the apparent phase of the interferometer. For light passing through the gadgets, the output of the interferometer is given by applying a Hadamard operation on the control qubit in (7.1-1):

$$\begin{aligned}
&\frac{1}{\sqrt{2}}\left[UV \otimes H\,|0\rangle\langle 0|_C + e^{i\varphi} VU \otimes H\,|1\rangle\langle 1|_C\right] \\
&= UV \otimes \frac{|0\rangle\langle 0|_C + |1\rangle\langle 1|_C}{2} + e^{i\varphi} VU \otimes \frac{|0\rangle\langle 0|_C - |1\rangle\langle 1|_C}{2} \\
&= \frac{1}{\sqrt{2}}\left[(UV + e^{i\varphi} VU) \otimes |0\rangle\langle 0|_C + (UV - e^{i\varphi} VU) \otimes |1\rangle\langle 1|_C\right]
\end{aligned} \tag{7.1-2}$$

For the phase measurement one would like $U = V = \mathbb{1}$, which gives

$$\frac{1}{2}\left[(\mathbb{1} + e^{i\varphi}\mathbb{1}) \otimes |0\rangle\langle 0|_C + (\mathbb{1} - e^{i\varphi}\mathbb{1}) \otimes |1\rangle\langle 1|_C\right]. \tag{7.1-3}$$

The measurement probabilities on the control qubit are given by the magnitude of the corresponding transformed polarization state:

$$\mathrm{P}_0 = \frac{1}{4}\langle\Psi|(\mathbb{1} + e^{i\varphi}\mathbb{1})^\dagger(\mathbb{1} + e^{i\varphi}\mathbb{1})|\Psi\rangle = \left|\frac{(1 + e^{i\varphi})}{2}\right|^2 = \cos^2\frac{\varphi}{2} \tag{7.1-4}$$

$$\mathrm{P}_1 = \frac{1}{4}\langle\Psi|(\mathbb{1} - e^{i\varphi}\mathbb{1})^\dagger(\mathbb{1} - e^{i\varphi}\mathbb{1})|\Psi\rangle = \left|\frac{(1 - e^{i\varphi})}{2}\right|^2 = \sin^2\frac{\varphi}{2}, \tag{7.1-5}$$

where $|\Psi\rangle$ is the polarization state injected into interferometer. As can be seen above, when the polarization transformations are the identity operation simply monitoring the relative power in the two output ports of the interferometer constitutes a direct measurement of the phase φ. However, if the locking light passes through the polarization gadgets when they are not implementing the identity transformation the measurement probabilities above become modified by the magnitudes of

$$\frac{1}{4}\langle\Psi|(UV \pm e^{i\varphi} VU)^\dagger(UV \pm e^{i\varphi} VU)|\Psi\rangle, \tag{7.1-6}$$

which in general depend both on the polarization state $|\Psi\rangle$ and the polarization transformations U and V, and the phase information cannot easily be extracted. The above

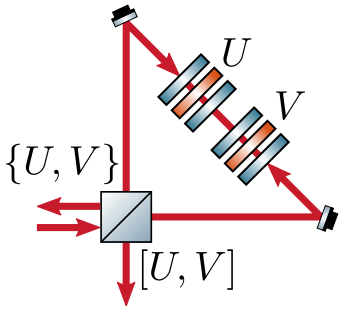


Figure 7.1: **Sagnac quantum switch.** A path-polarization quantum switch can be realised in a Sagnac geometry using standard linearly birefringent polarization gadgets to implement the unitary transformations U and V. However, in this case only rotations in the two-parameter subset $\{U = \exp[-i\alpha(\cos\theta Y + \sin\theta Z)|\alpha, \theta \in [0, 2\pi]\}$ of SU(2) can be implemented, restricting the use of this approach to special cases only requiring a limited gate set.

calculation was done for single photon states, which for obvious reasons would not be used to lock the interferometer. However, as discussed in Appendix B, it is also valid for coherent input states.

The output probabilities of the quantum switch can also be calculated using the process matrix formalism outlined in Section 1.2.8. The process matrix for the quantum switch can be expressed as [11, 53]

$$W_{\text{switch}} = |\mathrm{w}\rangle\langle\mathrm{w}|\,, \tag{7.1-7}$$

where

$$|\mathrm{w}\rangle = \frac{1}{\sqrt{2}}\big(|\mathrm{w}^{A\to B}\rangle\,|0\rangle_C + |\mathrm{w}^{B\to A}\rangle\,|1\rangle_C\big), \tag{7.1-8}$$

and

$$\begin{aligned} |\mathrm{w}^{A\to B}\rangle &= |\mathbb{1}\rangle\!\rangle_{T_{\text{in}}A_{\text{in}}}|\mathbb{1}\rangle\!\rangle_{A_{\text{out}}B_{\text{in}}}|\mathbb{1}\rangle\!\rangle_{B_{\text{out}}T_{\text{out}}} \\ |\mathrm{w}^{B\to A}\rangle &= |\mathbb{1}\rangle\!\rangle_{T_{\text{in}}B_{\text{in}}}|\mathbb{1}\rangle\!\rangle_{B_{\text{out}}A_{\text{in}}}|\mathbb{1}\rangle\!\rangle_{A_{\text{out}}T_{\text{out}}}. \end{aligned} \tag{7.1-9}$$

The actual probabilities are then found as

$$P_{\pm} = \mathrm{Tr}\big[SW_{\text{switch}}\big], \tag{7.1-10}$$

with [443]

$$S = \rho^T_{T_{\text{in}}} \otimes \big(|U\rangle\!\rangle\langle\!\langle U|_A\big)^T \otimes \big(|V\rangle\!\rangle\langle\!\langle V|_B\big)^T \otimes \mathbb{1}_{T_{\text{out}}} \otimes M_C, \tag{7.1-11}$$

where

$$\rho_{T_{\text{in}}} = |\Psi\rangle\langle\Psi|_{T_{\text{in}}} \tag{7.1-12}$$

is the polarization input state and

$$M_C = R_z(\varphi)\,|\pm\rangle\langle\pm|_C\,R_z(-\varphi) \tag{7.1-13}$$

is the measurement projector at the output beamsplitter, defined by the interferometric phase φ. Here the outcome '+' ('−') corresponds to the outcome '0' ('1') in (7.1-4).

Since during the experiment the polarization gadgets will typically not implement the identity transformation, the locking light needs to be aligned so that it does not pass through the gadgets. While this is of course possible in principle, it is harder to do in practice as wave-plate mounts, and in particular motorized ones, tend to be quite bulky. This necessitates very large beam separations or mounting solutions specifically engineered to let a secondary beam pass unaffected. Therefore, while active stabilization of a quantum switch has been demonstrated [432], it is typically not done.

An alternative is to use a passively stable Sagnac geometry, as shown in Fig. 7.1. This was done in Publication 2, but due to the nonreciprocal nature of wave-plate-only gadgets this type of implementation is limited to a two-parameter subset of SU(2) generated by the two Pauli operators that are symmetric under counterpropagation.

7.1.2 Alternative realisations

There have also been demonstrations of quantum switches using other encodings than the path-polarization one, and some of these have employed passively phase-stable geometries. The experiments presented in Ref. [56, 444] used the polarization degree of freedom as a control system, while the target qubit system was encoded in the transverse modes of the light; specifically, the two Hermite-Gaussian modes HG_{01} and HG_{10} [106]. As shown in Fig. 7.2 a PBS was used to map the polarization degree of freedom to the photon path in a polarization-path interferometer. However, through the use of a looped geometry both photon trajectories, corresponding to the control qubit basis states, traverse both arms of the interferometer, and do therefore not acquire any relative phase from low-frequency path-length fluctuations. Additionally, the optical geometry ensures that photons in the two different paths through the quantum switch hit the devices implementing the two unitaries on the same spot, unlike the typical path-polarization geometry that relies on having two offset beams through the same gadget.

The main drawback with this approach is the high complexity associated with the manipulation of the Hermite-Gaussian modes. Transformations on the modes were realised using a sequence of rotating prisms and cylindrical lenses, requiring a relatively high number of optical components. Furthermore, since the transverse modes form a countably infinite set, imperfect transformations will typically take the state out of the $\{\mathrm{HG}_{01}, \mathrm{HG}_{10}\}$ qubit subspace. Finally, state initialization was done using spatial filtering, and was therefore post-selective. These experimental challenges are likely the reason why the transverse-mode encoding has not seen more widespread use.

Another realisation of a quantum switch without the need for phase stabilization is the demonstration of exponential advantage in communication complexity presented in Ref. [421]. The experiment used a time-bin encoding for the target states, consisting of $d = 2^{n+1}$ time-bins, with $n = 12$. The implemented unitary transformations consisted of shifts on the time-bins:

$$X_d \left|\tau\right\rangle = \left|\tau + 1 \bmod d\right\rangle, \tag{7.1-14}$$

and time-bin dependent phase-shifts:

$$Z_d \left|\tau\right\rangle = e^{2\pi i \tau/d} \left|\tau\right\rangle. \tag{7.1-15}$$

The different time-shift operations in the experiment were realised by two sets of n fiber patch-cable pairs that were connected in specific combinations to yield the appropriate

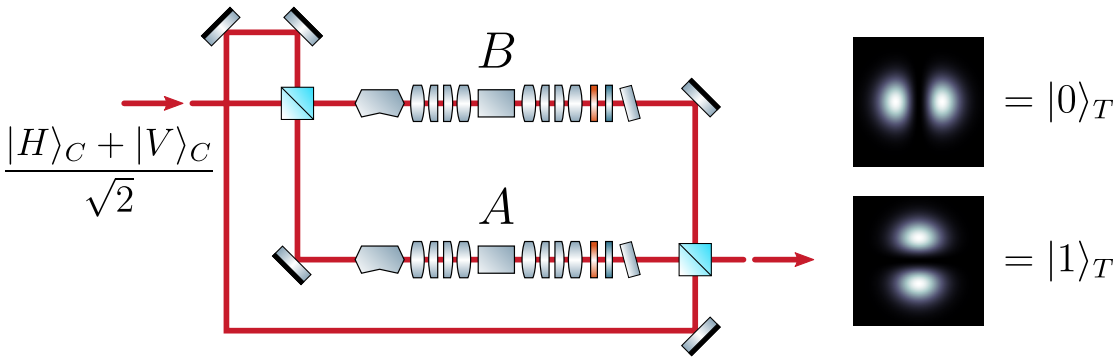


Figure 7.2: **Polarization-TEM quantum switch.** By encoding both the target and control systems of the quantum switch in internal degrees of freedom of a photon one can achieve passively phase-stable operation. This was demonstrated in Ref. [56] using the setup shown above. The control qubit was encoded in the polarization state, while the target qubit was encoded in the Hermite-Gaussian transverse electromagnetic modes (TEM) shown on the right. Using a PBS the polarization state is mapped to the photon path and the transformations A and B are applied conditional on the state of the control qubit using a combinations of lenses, prisms and wave plates. After this, the photons paths are recombined on a second PBS, and routed through the polarization-path interferometer via the second input of the first PBS. Since the photons always traverse both arms of the interferometer any relative phase between these arms only manifest as a global phase.

time delays. This meant that to implement different X_d shifts the setup had to be manually reconfigured. Furthermore, the Z_d phase shifts required the use of ultra-fast phase modulators, with modulation bandwidths in excess of 500 MHz. Finally, the restriction to time translations and phase shifts meant that the setup could not realise fully general transformations on the time-bins, or even prepare superpositions of the computational basis states. Therefore, while the experimental setup was cleverly designed to demonstrate the communication advantage presented in Ref. [420], the experimental methods employed do not readily generalise to other uses of the quantum switch.

7.1.3 Common-path path-polarization quantum switch

To maintain the advantage of the high fidelity and easily reconfigurable unitary transformations enabled by a polarization encoding for the target state, while at the same time making use of a passively phase-stable common-path geometry, one can build a quantum switch using the reciprocal polarization gadget described in Chapter 3:

$$\mathbf{G}_{\mathrm{R}} = \mathbf{Q}(\theta)\mathbf{H}(\phi)\mathbf{X}^S(\gamma)\mathbf{H}(-\phi)\mathbf{Q}(-\theta). \tag{7.1-16}$$

The reciprocity of the gadget follows immediately from the palindromic order, and the fact that the middle X-gadget

$$\begin{aligned}\mathbf{X}^S(\gamma) &= \mathbf{H}(\pi/8)\mathbf{Z}^A(\gamma-2\pi)\mathbf{H}(\pi/8)\\ &= \mathbf{H}(\pi/8)\mathbf{F}_-\mathbf{X}^A(\gamma-2\pi)\mathbf{F}_+\mathbf{H}(\pi/8)\\ &= \mathbf{H}(\pi/8)\mathbf{F}_-\mathbf{Q}(\pi/2)\mathbf{H}(\gamma/4-\pi/2)\mathbf{Q}(\pi/2)\mathbf{F}_+\mathbf{H}(\pi/8)\end{aligned} \tag{7.1-17}$$

is symmetric. While the symmetry of this gadget was shown mathematically in Chapter 3, it can also be understood visually. As illustrated in Fig. 7.3, the key principle

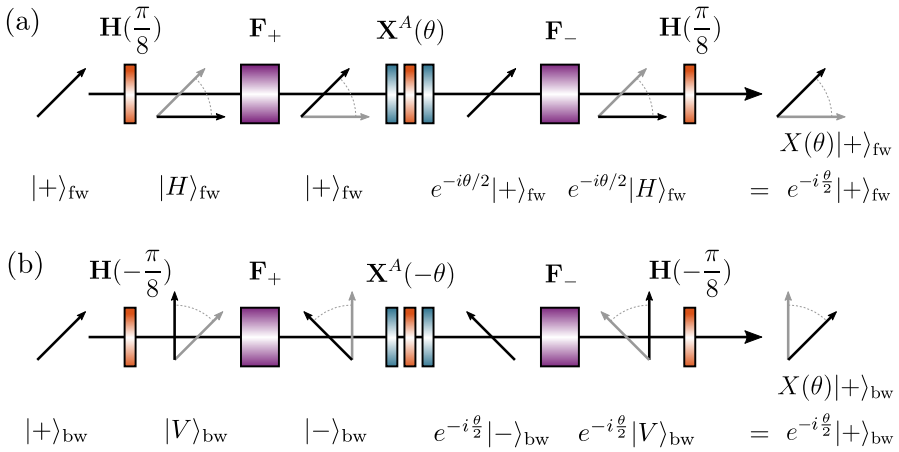


Figure 7.3: **Reciprocity of the $\mathbf{X}^S$ gadget**. The reciprocal universal gadget uses a reciprocal X-gadget as a key building block, and the symmetry properties of this latter gadget has an intuitive explanation. In both sub-figures the evolution of the eigenstate $|+\rangle$ is shown in the co-moving frame of the light. The gadget consists of a Simon–Mukunda gadget configured to implement an X-rotation, sandwiched by two fixed Faraday rotators and two half-wave plates. As shown in (a), in the forward direction the first half-wave plate, set at an angle of $\pi/8$, rotates an initial $|+\rangle$ eigenstate in a clockwise direction to $|H\rangle$, and the $\mathbf{F}_+$ Faraday rotator then applies a counter-clockwise rotation, cancelling the action of the HWP and taking the state back to $|+\rangle$. After receiving a phase shift from the central gadget, the process is repeated by the second HWP and Faraday-rotator pair. In the counterpropagating direction, shown in (b) the HWP angles are mirrored, causing the rotations of the HWPs and Faraday rotators to add, turning the input state $|+\rangle$ into $|-\rangle$. Due to the anti-symmetric nature of the central gadget it applies the same phase to the anti-diagonal state in the backwards direction as to the diagonal state in the forwards direction. Finally, the last Faraday rotator and HWP turn the state back to $|+\rangle$, and we see that this eigenstate acquires the same phase in both directions, thereby showing the symmetry of the gadget. Note that additional phases inside the gadget that ultimately cancel have been omitted for the sake of clarity.

of the gadget is that both the Faraday rotators $\mathbf{F}_\pm$ and the half-wave plates $\mathbf{H}(\pm\pi/8)$ rotate a diagonally polarized state $|\pm\rangle$ by $\pi/4$, and in the forward propagation direction these rotations cancel, while in the backwards propagation direction the rotations add to a $\pi/2$ rotation that has the effect of swapping the eigenstates. Specifically, in the forward direction:

$$\mathbf{H}(\pi/8)\mathbf{F}_- = \mathbf{F}_+\mathbf{H}(\pi/8) = -iX, \tag{7.1-18}$$

the rotations cancel, preserving the states $|\pm\rangle$, and in the backwards direction:

$$\mathbf{H}(-\pi/8)\mathbf{F}_- = \mathbf{F}_+\mathbf{H}(-\pi/8) = -iZ, \tag{7.1-19}$$

the rotations add, swapping the states $|\pm\rangle \mapsto |\mp\rangle$. This exchange of eigenstates on both sides of the central gadget cancels the with the sign change in the rotation angle to

produce the same transformation in both directions:

$$\begin{aligned}\Theta\big[\mathbf{X}^S(\gamma)\big] &= \Theta\big[\mathbf{H}(\pi/8)\mathbf{F}_-\mathbf{X}^A(\gamma-2\pi)\mathbf{F}_+\mathbf{H}(\pi/8)\big]\\ &= \mathbf{H}(-\pi/8)\mathbf{F}_-\mathbf{X}^A(-\gamma+2\pi)\mathbf{F}_+\mathbf{H}(-\pi/8)\\ &= i^2 ZX(-\gamma+2\pi)Z\,|\pm\rangle\\ &= ZX(-\gamma)Z\,|\pm\rangle\\ &= ZX(-\gamma)\,|\mp\rangle\\ &= X(\gamma)Z\,|\mp\rangle\\ &= X(\gamma)\,|\pm\rangle\,.\end{aligned} \tag{7.1-20}$$

Given the reciprocal gadget (7.1-16) construction of a quantum switch becomes fairly straightforward.

7.2 Advantage of the quantum switch in a channel discrimination task

The aim of the experimental part of our work was to certify our design of the quantum switch, and to this end we chose to demonstrate a task at which the quantum switch is known to outperform all causally ordered processes. This task, similar to the one presented in Chapter 6, consists of determining whether a given pair of unitaries U and V commute or anti-commute, while only given access to a single use of these unitaries. Our specific version of the task is adapted from [53], where it was first presented.

We begin by defining the following set of unitaries:

$$\mathcal{G} := \left\{\mathbb{1}, X, Y, Z, \frac{X\pm Y}{\sqrt{2}}, \frac{X\pm Z}{\sqrt{2}}, \frac{Y\pm Z}{\sqrt{2}}\right\}. \tag{7.2-1}$$

Using this set, we then define a two new sets consisting of pairs of unitaries (U_i, V_j) from (7.2-1) that either commute or anti-commute:

$$\mathcal{G}_\pm := \left\{(U_i, V_j)\,\middle|\, U_i, V_j \in \mathcal{G},\ U_iV_j = \pm V_jU_i\right\}. \tag{7.2-2}$$

These two sets contain 28 commuting and 24 anti-commuting pairs unitaries respectively, for a total of 52 pairs. It is well known that the quantum switch can determine from which set a given pair was picked with unity success probability by creating a superposition of the two gates such that the commutator or the anti-commutator is applied conditional on the state of a control qubit [406], in the way previously described in Chapter 5. In short, one prepares the joint target-control input state $|\Psi\rangle_T \otimes |+\rangle_C$ and lets the quantum switch generate a superposition of the gates U and V being applied in different orders, giving the state

$$\begin{aligned}&\big(UV\otimes|0\rangle\langle 0|_C + VU\otimes|1\rangle\langle 1|_C\big)\,|\Psi\rangle_T\otimes|+\rangle_C\\ &= \frac{1}{2}\{U,V\}\,|\Psi\rangle_T\otimes|+\rangle_C + \frac{1}{2}[U,V]\,|\Psi\rangle_T\otimes|-\rangle_C\,.\end{aligned} \tag{7.2-3}$$

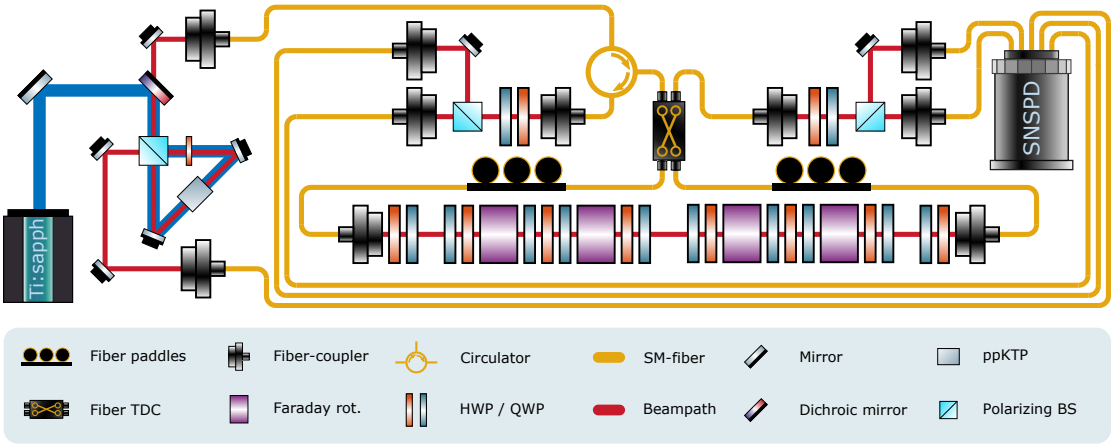


Figure 7.4: **Experimental setup.** Similar to the other experiments presented in this thesis, a spontaneous parametric down-conversion source with a ppKTP crystal in a Sagnac configuration was used to generate single-photon pairs. Since polarization entanglement between the signal-idler pair was not needed the source was pumped in a single direction. The signal photon was used as a trigger for the idler photon, which was sent into the Sagnac quantum switch through a fiber circulator. Using a tunable directional coupler (TDC) in lieu of a regular fiber beamsplitter allowed the splitting ratio to be fine tuned, thereby increasing the interferometric visibility. The free-space part of the interferometer contained two pairs of half- and quarter-wave plates used together with the fiber paddles for polarization control, as well as two reciprocal polarization gadgets implementing the two unitaries U and V. After being split on the TDC the single photons propagate through the two gadgets in a superposition of directions, and therefore a superposition of the orders in which the two gates can be applied. Upon re-interfering on the TDC the photons exit in one of the two ports depending on whether U and V commute or anti-commute. The fiber circulator picks off the backwards propagating photons in the input port, and two polarizing beamsplitters allow for polarization dependent detection efficiencies in the superconducting nanowire single-photon detectors (SNSPDs) to be corrected for.

Tracing out the target qubit and measuring the state of the control qubit then reveals from which set U and V were picked.

In contrast to casually indefinite processes such as the quantum switch, causally separable processes exhibit a maximum success probability strictly less than 1. Using the computer assisted proof methods from [427] we derive bounds on the minimum and average success probabilities for causally separable strategies, and these are

$$\min\big(p_{\mathrm{s}}(i,j)\big) \leq 0.841 \tag{7.2-4}$$

$$\langle p_{\mathrm{s}}\rangle = \frac{1}{N}\sum_{i,j} p_{\mathrm{s}}(i,j) \leq 0.904. \tag{7.2-5}$$

For the average success probability it is assumed that the unitaries (U_i, V_j) are picked from $\mathcal{G}_{\pm}$ using a uniformly random distribution, and in this case the task can equivalently be formulated as a causal witness, similar to the approach taken in [11, 53]. Violating these bounds shows that the experimentally realised process is causally indefinite, and this would in turn validate our design.

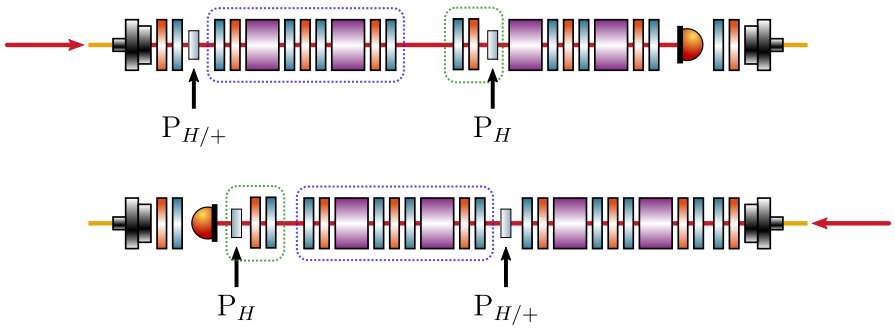


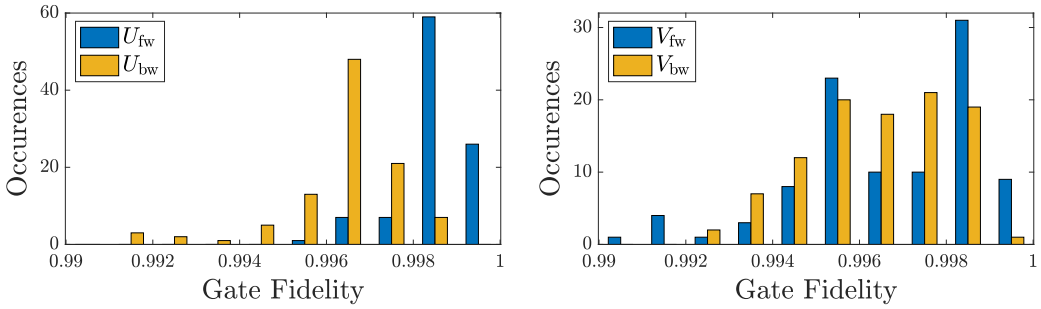
Figure 7.5: **Gadget process tomography setup.** Tomography was performed on each of the two gadgets in both propagation directions. In one of these directions the second gadget could be used to set the measurement basis for the tomography, while in the other propagation direction two additional wave plates were inserted. The blue outline indicates the gadget being measured, while the green outline marks the measurement apparatus. The polarizer used for the projective measurement was fixed at $|H\rangle\langle H|$, while the state-preparation polarizer was set to either $|H\rangle\langle H|$ or $|+\rangle\langle +|$ depending on the input state.

7.2.1 Interferometer and experiment design

The inherent phase stability of the Sagnac geometry means that the use of fiber-optic elements inside the interferometer does not pose a problem, and our experimental setup therefore utilized a tunable directional coupler (TDC) instead of a bulk-optic beamsplitter. This choice was firstly made because the near unity mode overlap inside the TDC allows for higher interferometric visibilities than those typically achieved in bulk interferometers. Secondly, the tunable nature of the TDC allows for further optimization of the interferometric visibility, by adjusting the splitting ratio while monitoring the dark port of the interferometer. Finally, injecting the light into fibers also simplifies the construction of the free-space part of the interferometer housing the polarization gadgets, since this part can be limited to a straight line punctuated by two fiber collimators. This is particularly convenient when using Faraday rotators, because of their typically small clear apertures of 5 mm or less.

Due to the two, essentially perfectly, overlapping beams inside the interferometer, the setup becomes quite susceptible to back reflections. This is not a major issue when using heralded single photons with short coherence length, as the back reflections will not contribute coincident two-photon detections. However, for characterisation measurements using CW light, back reflections can contribute significant noise and lower the apparent visibility of the interferometer. For this reason, anti-reflection coated APC fibers were used to minimize reflections when transitioning from fiber to free-space. This allowed high values of interferometric visibility to be recorded, with typical values in excess of 0.9995. These visibility measurement were done by using the polarization transformations to change the interference condition.

A full sketch of the experimental setup, including single-photon generation, polarization compensation optics and single-photon detection is shown in Fig. 7.4. In the input port of the interferometer a fiber-circulator was used to separate the backwards propagating light so that it could be detected. Polarization resolving detection stages were placed in both output ports in order to compensate for polarization-dependent de-



(a) U-gadget (b) V-gadget

Figure 7.6: **Gadget fidelities.** The two reciprocal gadgets were characterised by performing quantum state tomography on 100 randomly chosen unitaries. The fidelities of the reconstructed unitaries with respect to the ideal ones are shown for the U- and V-gadgets in (a) and (b), respectively. The U-gadget achieves consistently higher fidelities than the V-gadget, but exhibits an asymmetry in the two propagation directions. This behaviour is quantified in Fig. 7.7.

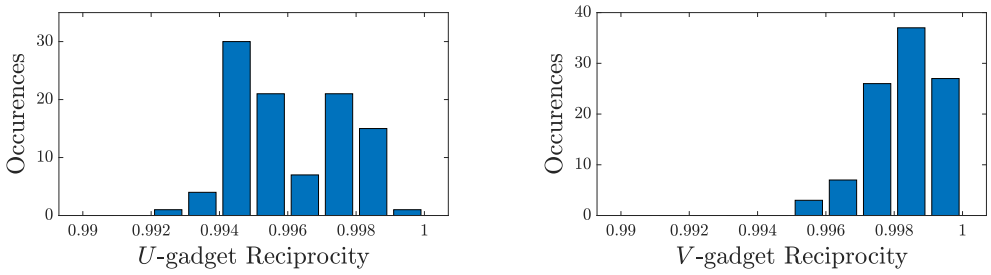
tection efficiencies in the single-photon detectors. This was done by splitting the H- and V-components of the light, and then performing single-basis polarization compensation on the fibers connected to the detectors to maximize the detection efficiency.

The common-path geometry employed in the experiment can potentially complicate the main polarization compensation procedure, since selectively blocking light in one arm of the interferometer requires an optical isolator. However, as will be discussed in Section 7.2, the implemented measurements involved tracing out the target qubit encoded in the photon polarization, and it was therefore only necessary to compensate the polarization in the two fibers internal to the interferometer. Doing this ensures that the input state to the first unitary is the same in both propagation directions, and likewise that the rotated polarization basis inside the fiber beamsplitter is the same in both fibers when the photons re-interfere with themselves as they exit the interferometer.

7.2.2 Gadget characterisation

Before performing the channel discrimination task described in Section 7.2, it was first necessary to characterise the two reciprocal gadgets, and verify that they worked as intended. This was done by performing quantum process tomography on the devices for 100 randomly picked SU(2) unitaries. While it would have been possible to perform these measurements using the polarization resolving measurement stages used in the experiment, unwanted polarization rotations in the fibers could potentially have introduced excess noise. The gadgets were therefore characterised fully in free-space, inside the interferometer itself. In order to quantify the reciprocity of the gadgets tomography was performed on each gadget in both propagation directions. As illustrated in Fig. 7.5, in one direction the second gadget could be used to perform the tomography of the first, while in the opposite direction two additional wave plates were inserted.

The tomography was performed by probing the gadgets with horizontally and diagonally polarized CW light, and reconstructing the resulting output polarization states. These were then used to fit the unitaries using a maximum likelihood estimation. Fi-



(a) U-gadget (b) V-gadget

Figure 7.7: **Gadget reciprocity.** The reciprocity of the polarization gadgets was quantified by measuring the fidelity of a given unitary transformation in one propagation direction with respect to the transformation in the opposite propagation direction. This analysis was done using the same dataset of 100 randomly chosen unitaries as in Fig. 7.6. Here we see that the V-gadget achieves a higher measure of reciprocity than the U-gadget, despite having an, on average, lower fidelity with respect to the ideal unitaries. Since most forms of uncorrelated noise will induce a slight symmetry breaking in the device, one would expect the fidelities and reciprocities to be correlated. The absence of this correlation indicates that imperfections in the measurement itself was likely a major contributor to the fidelity reductions.

nally, the fidelity of the unitaries were calculated as the average of the fidelities of the mapped states:

$$\mathcal{F}(W_i, W^d_{i,\text{exp}}) = \frac{1}{2}\big(\langle H|W_i^\dagger W_i^d|H\rangle + \langle +|W^\dagger W_i^d|+\rangle\big). \tag{7.2-6}$$

Here $d \in \{\text{fw}, \text{bw}\}$ denotes the propagation direction in which the experimentally reconstructed unitary was probed. The average fidelity over the 100 unitaries was found to be 0.9975 ± 0.0008 for the U-gadget, and 0.9966 ± 0.0013 for the V-gadget. Histograms of the fidelities for each measurement set are shown in Fig. 7.6, and it can be seen that the gadgets display similar behaviour in both directions. This was quantified further by calculating the reciprocity of the gadgets, which we define as

$$\mathcal{R}(W_i^{\text{fw}}, W_i^{\text{bw}}) = \frac{1}{2}\big(\langle H|(W_i^{\text{fw}})^\dagger W_i^{\text{bw}}|H\rangle + \langle +|(W_i^{\text{fw}})^\dagger W_i^{\text{bw}}|+\rangle\big). \tag{7.2-7}$$

This quantity was calculated using the same measurement data as the fidelities, and for the U gadget the reciprocity was found to be 0.9960 ± 0.0016, while for the V gadget it was 0.99834 ± 0.00099. The histograms in Fig. 7.7 show the distribution of the reciprocities over the 100 measured unitaries. Since the measured reciprocity can be seen to be strictly greater than 0.99 for all measurements, one can therefore be justified in calling the gadgets reciprocal. The uncertainties in both the fidelity and reciprocity were calculated as the standard deviations of the corresponding 100 measurement results.

The small deviations from the ideal behaviour of the gadgets could have been caused by imperfect circular retardances in the Faraday rotators, imperfect linear retardance in particular in the half-wave plates adjacent to the Faraday rotators, or errors in the angles of these wave plates. Additionally, it is worth pointing out that during the tomography all unitaries were first measured in a single direction, and then afterwards in the opposite direction. Consequently, finite repeatability in the wave-plate motors

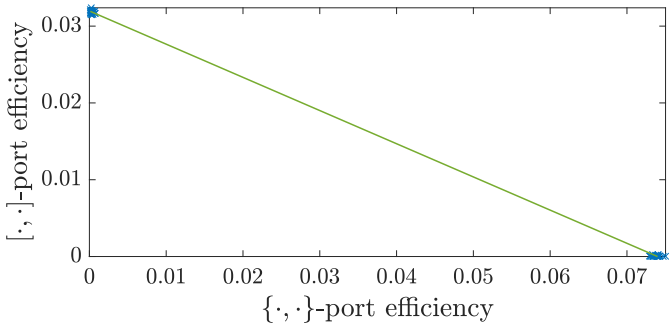


Figure 7.8: **Detection efficiencies.** The two output ports of the interferometer had different efficiencies, and it was therefore necessary to characterise and correct for this in order to accurately estimate the probability for a photon to exit in a given port. This was done by exploiting the fact that the total photon-detection event rate should be constant even as the interference condition is changed. In an ideal experiment, where all efficiencies are equal, changing the interference condition only redistributes the photon events between the two output ports, and if one plots the event rate in one detector as a function of the rate in the other, one would then find that the points lie of a line with slope -1. In the case of unequal detection efficiencies one instead observes a different line, whose slope gives the relative detection efficiency. Above, such a line is shown fitted to the experimental data in the third run of the experiment, instead of scanning the phase continuously. The slope of the fitted line is -0.432, indicating that one port had just over twice as much overall loss. The axes of the figure show the photon event rate normalized to the heralding photon rate, and not the raw detection rate.

could have led to slightly different wave-plate angles for the two measurement directions. Finally, as shown in Fig. 7.5, different wave plates were used for the measurements in the two directions, which might also have contributed measurement errors.

Comparing Fig. 7.6 and Fig. 7.7 one sees that the U-gadget showed marginally higher fidelities, while the V-gadget achieved higher reciprocity. Generally, uncorrelated errors in the gadgets should contribute to a symmetry breaking in the device, and one would therefore expect the fidelity and reciprocity to be somewhat correlated. Imperfections that do not break the reciprocity are limited to symmetric errors in the outer wave-plate pairs, and errors in the central Simon–Mukunda gadget that do not introduce a Z-component to the corresponding unitary. While such errors are unlikely to have occurred systematically, the observed results could have been caused by imperfections in the measurements rather than the gadgets. This is because observing a high reciprocity necessitates that the measurements in the two directions be highly correlated. In total three different sets of wave plates were used in the tomography; one moveable set, and the two inner waveaplates of each gadget. One possible explanation is thus that the inner wave plates of the U-gadget, used to measure the V-gadget in one direction, matched the moveable wave plates more closely than those of the V-gadget used to measure the U-gadget did.

7.2.3 Experimental channel discrimination

To certify the design of the quantum switch, the channel discrimination task presented in Section 7.2 was performed experimentally, with the intent to violate the bounds

imposed on causally separable processes. As previously explained the quantum switch transforms the unitaries (U, V) into a superposition of being applied in two different orders correlated with the state of the control qubit. Measuring this latter state in the X-basis then reveals whether the pair of unitaries was commuting or anti-commuting, and experimentally this is equivalent to measuring the path of the photon as it exits the interferometer.

As illustrated in Fig. 7.8 the detection efficiency in one output port of the interferometer was approximately double that of the other output port. This was partially due to extra loss introduced by the fiber circulator, but the excess loss also had multitude of other contributing factors, such as higher fiber mating loss, lower SNSPD detection efficiencies and diminished free-space to fiber coupling efficiencies in the polarization resolving detection stage. In order to more accurately estimate the probability for a photon to exit in a given port, the detected photon rates were therefore normalized to the observed detection efficiencies.

Each of the 52 unitaries was implemented six independent times, and two-photon events were recorded for 60 s for each setting and run, giving a total measurement time of 5.2 h, not including the time spent rotating the wave plates. The duty cycle of the experiment was in excess of 90 %, and could have been higher for longer individual data acquisition times. This is in contrast to passively stabilized realisations of the quantum switch, which require the interferometric phase to be periodically reset.

The results of the detection-efficiency-corrected measurements are shown in Fig. 7.9, where it is apparent that the $\min\big(p_{\mathrm{s}}(i,j)\big) \leq 0.841$ and $\langle p_{\mathrm{s}} \rangle =\leq 0.904$ bounds on the on the minimum and average success probabilities, respectively, are violated by a significant margin. More quantitatively, the observed minimum success probability was 0.9895, while the average success probability was 0.99639 ± 0.00007. Since the goal of the design was to achieve a more robust and performant setup, it is worth comparing the success fidelities to the experiment in which an analogous channel disctrimination task was first implemented, namely the experiment by Procopoio et al. In their work [406], mean success probabilities of 0.973 and 0.976 for Pauli and random gates, respectively, were achieved. This lower performance can be attributed to challenges associated with phase drifts, imperfect mode matching on the bulk beamsplitter, and spatial-nonuniformity of the wave plates, all of which are overcome in our design.

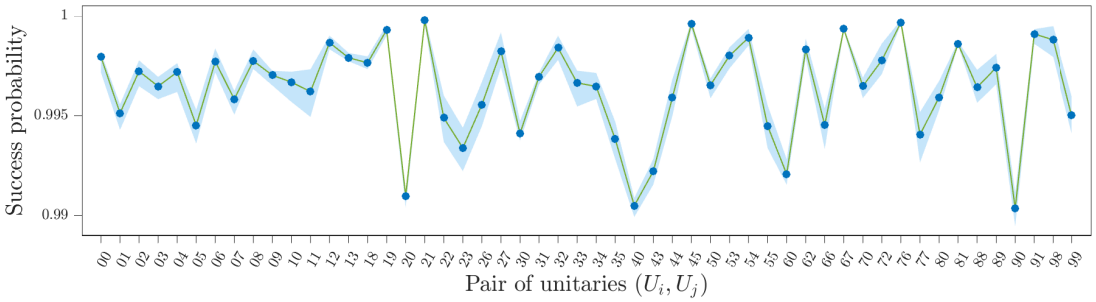


Figure 7.9: **Channel discrimination success probabilities.** The channel discrimination task was performed experimentally for 52 different pairs of unitaries, and the observed probabilities of successfully determining from which set the unitaries were picked are shown above. Each pair of unitaries was implemented in six independent runs; the solid dots indicate the average success probability across these runs, and the blue shaded area shows the minimum and maximum probabilities. The mean success probability was $\langle p_s^{\text{exp}} \rangle = 0.99639 \pm 0.00007$, with maximum and minimum probabilities of $\min(p_s^{\text{exp}}) = 0.9895$, $\max(p_s^{\text{exp}}) = 0.99997$, clearly violating the bounds for causally separable processes. The high repeatability between the six runs demonstrates the stability of the setup, and the lower success probabilities for certain settings are likely a result of systematic errors, such as wave-plate imperfections.

7.3 Extensions beyond the two-party quantum switch

Realisations of casually indefinite orders have generally been restricted to superpositions of two gates, due to the challenge of building the complex interferometers needed to realise larger superpositions. For example, creating a superposition of the $3! = 6$ permutations of a set of three gates using a path-polarization encoding would require aligning six different beams through the same polarization optics, and stabilizing the relative phase of these beams. In Ref. [445], the only experimental realisation of a superposition of $N > 2$ gate orders, this challenge was tackled using multi-core fibers that, after free-space focusing, allowed multiple orthogonal modes to be overlapped in the far field. This implementation, however, still required direct control over the phase in the non-common-path geometry.

Interestingly, the methods applied to construct a passively stable two-party quantum switch in this chapter, specifically the reciprocal polarization gadget, can be extended to three parties as well, and could for example be used to realise the Hadamard promise problems presented in Ref. [422]. To begin, one needs to realise a beamsplitter transformation that prepares an appropriate superposition of control qubit states, in accordance with the gate superpositions to be realised. In the case of the full 6 permutations for three single-qubit gates, this could for example be the 6-dimensional quantum Fourier transform:

$$U_{BS}^{6} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \frac{1+i\sqrt{3}}{2} & \frac{-1+i\sqrt{3}}{2} & -1 & \frac{-1-i\sqrt{3}}{2} & \frac{1-i\sqrt{3}}{2} \\ 1 & \frac{-1+i\sqrt{3}}{2} & \frac{-1-i\sqrt{3}}{2} & 1 & \frac{-1+i\sqrt{3}}{2} & \frac{-1-i\sqrt{3}}{2} \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & \frac{-1-i\sqrt{3}}{2} & \frac{-1+i\sqrt{3}}{2} & 1 & \frac{-1-i\sqrt{3}}{2} & \frac{-1+i\sqrt{3}}{2} \\ 1 & \frac{1-i\sqrt{3}}{2} & \frac{-1-i\sqrt{3}}{2} & -1 & \frac{-1+i\sqrt{3}}{2} & \frac{1+i\sqrt{3}}{2} \end{bmatrix}. \tag{7.3-1}$$

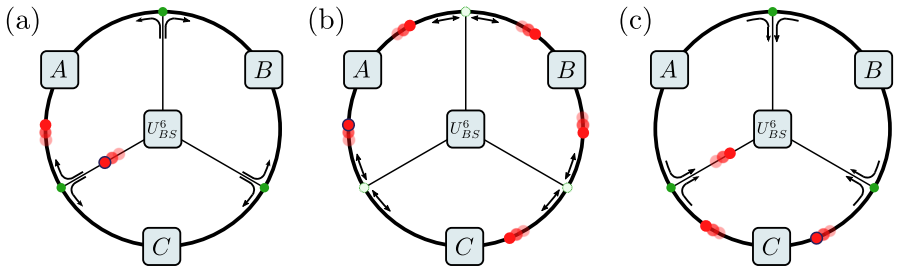


Figure 7.10: **Passively phase stable three-party quantum switch.** The three-party quantum switch is a process that transforms three gates A, B, C into a coherent superposition of every possible order of them. These orders can be realised in a circular geometry with two propagation directions and three input/output points. The center of the ring contains a device implementing a generalized beamsplitter transformation on six modes. These modes are then injected into the ring in two different propagation directions, travelling through three reciprocal polarization gadgets. (a) The subfigure shows a photon exiting in one output port of the generalized beamsplitter. The photon's initial position is indicated by a contour, and the other circles show the photon at subsequent times. An active switch, indicated by a green circle, injects the photon into the loop where it first travels to the gadget implementing the unitary operator A. (b) After the photon has entered the loop, the active switches change state to trap the photon in the loop, allowing it to propagate through the remaining two polarization gadgets. **(c)** After the photon has completed one full revolution, the switches once again change state and send the photon back to the central beamsplitter. The path in the figure generates the gate sequence CBA, but any other permutation of the three gates can be generated by considering different input/output points and/or a different propagation direction.

This transformation, when acting on the photon path, could be implemented in several different ways, for example using a Reck or Clements encoding in a waveguide, or using beam-displacers.

The central idea of the proposed realisation is to let the photon propagate through the three gates along a circular path. By injecting the photon along different points of this circle, and in different propagation directions, according to the state of the control qubit, all six permutations of the gates can be realised. This is illustrated in Fig. 7.10. After completing one full revolution the photon would propagate along the same spoke it was initially injected along. Since the paths corresponding to the different control qubit states all traverse the circular path exactly once, any phase fluctuations in this path fully cancel. Similar to the Sagnac geometry a reciprocal polarization gadget is required, since both clockwise and counter-clockwise propagating paths are present.

If the photons in the loop propagate in the same spatial mode, then the points where they enter the loop, marked in green in Fig. 7.10, need to be actively switched on and off. However, this switching can be accomplished with a single physical device using the optical geometry shown in Fig. 7.11. After exiting the generalized beamsplitter, the photons are injected in one of two different ports of a Mach-Zehnder interferometer, depending on whether they will propagate clockwise or counter-clockwise through the loop. The phase φ of the interferometer is controlled by a fast electro-optical modulator, and is initially set to $\varphi = 0$, injecting the photons into the loop. Setting the phase to $\varphi = \pi$ traps the photons, and resetting the phase after three passses through the MZI directs them back to the beamsplitter. Errors in the MZI phase would not contribute to

phase errors in the actual quantum switch, but would instead simply manifest as photon loss.

As seen in Fig. 7.11, the colour- coded photon paths do not retrace themselves as they enter the beamsplitter a second time, but instead paths with the same color swap. Therefore, these photon trajectories are relatively phase stable. In Fig. 7.10 they correspond to two counterpropagating photon paths injected along different spokes, that overlap in the first gadget they pass through. Paths marked with different colours can in principle acquire relative phases through path length fluctuations. However, similarly to displaced Sagnac interferometers [154, 446], one would in practice expect the geometry to be stable. Finally it is worth noting that the multi-core fibers employed in Ref. [445] could be used to circumvent the need for an active phase modulator, since, if the photons propagate through the loop in six orthogonal spatial modes, the spatial transformations in the injection and extraction points at the end of each spoke could be statically configured without violating unitarity.

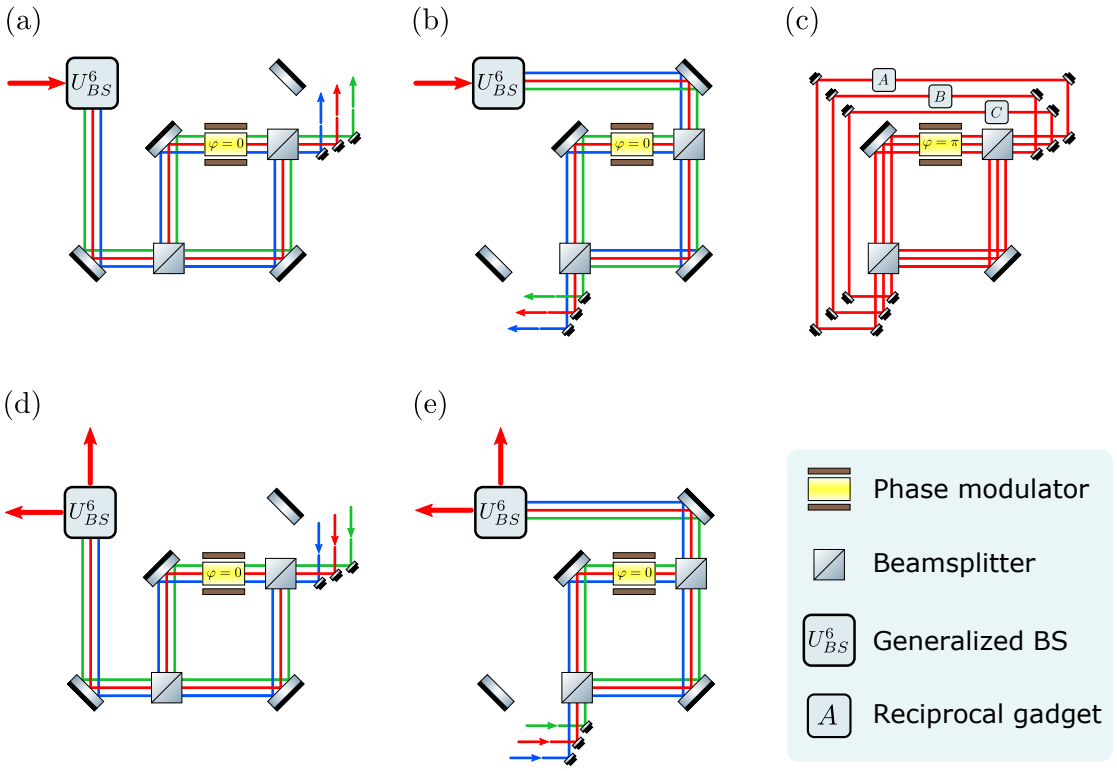


Figure 7.11: **Proposed optical setup for the three-party quantum switch.** The figure shows how a single Mach-Zehnder interferometer containing an ultra-fast phase modulator can be used to inject a photon into the loop. Note that the phase modulator has to be polarization independent to preserve the polarization state. This can be achieved by for example using two crossed EOMs working in parallel. (a) Since two propagation directions overlap inside the Mach-Zehnder, only the bottom three output ports of the generalized beamsplitter are shown. When the phase of the modulator is $\varphi = 0$ the photons entering the Mach-Zehnder interferometer will exit in the top right port, and travel through the loop in a counter-clockwise direction. (b) In analogy to the lower photon paths, the right output modes of the generalized beamsplitter will exit the Mach-Zehnder in the bottom port, and travel through the loop in a clockwise fashion. **(c)** When a photon has entered the loop the phase modulator changes state to $\varphi = \pi$ causing the photons to be trapped. The paths are aligned such that a photon will always propagate through all three polarization gadgets after three revolutions. **(d-e)** After the photons have passed through the gadgets implementing A, B, C exactly once, the phase modulator changes its state back to $\varphi = 0$ and the photons are released from the loop. Note that the photons in (a), that initially took the lower path, return along the upper path in (c), and likewise for (b) and (d). This means that any path-length fluctuations in this part of the setup are common mode and do not induce a relative phase shift on any of the colour-coded photon amplitudes pairs.

8

Reprint of "Trace-free counterfactual communication with a nanphotonic processor"

8.1 Contribution statement

For this work I assembled and characterised the experimental setup with help from Nicholas C. Harris, and performed the measurements and collected the data presented in the publication. Chiara Greganti and Valeria Saggio assisted with the single-photon generation, and Irati Alonso Calafell and Lee Rozema with the superconducting nanowire single-photon detectors. The data analysis and manuscript writing was carried out jointly with Lee Rozema and Irati Alonso Calafell, with supervision from Philip Walther. David R. M. Arvidsson-Shukur and Crispin H. W. Barnes developed the theoretical ideas. Irati Alonso Calafell, Lee Rozema and David R. M. Arvidsson-Shukur and Philip Walther conceived the experiment. Nicholas C. Harris, Mihika Prabhu, Jacques Carolan, Michael Hochberg, Tom Baehr-Jones and Dirk Englund designed and produced, or facilitated the production of the photonic waveguide. Nicholas C. Harris developed the control electronics and software.

npj | Quantum Information

www.nature.com/npjqi

ARTICLE OPEN

Trace-free counterfactual communication with a nanophotonic processor

I. Alonso Calafell[1], T. Strömberg[1], D. R. M. Arvidsson-Shukur[2,3], L. A. Rozema[1], V. Saggio[1], C. Greganti[1], N. C. Harris[4], M. Prabhu[4], J. Carolan[4], M. Hochberg[5], T. Baehr-Jones[5], D. Englund[4], C. H. W. Barnes[2] and P. Walther[1]

In standard communication information is carried by particles or waves. Counterintuitively, in counterfactual communication particles and information can travel in opposite directions. The quantum Zeno effect allows Bob to transmit a message to Alice by encoding information in particles he never interacts with. A first remarkable protocol for counterfactual communication relied on thousands of ideal optical operations for high success rate performance. Experimental realizations of that protocol have thus employed post-selection to demonstrate counterfactuality. This post-selection, together with arguments concerning a so-called "weak trace" of the particles traveling from Bob to Alice, have led to a discussion regarding the counterfactual nature of the protocol. Here we circumvent these controversies, implementing a new, and fundamentally different, protocol in a programmable nanophotonic processor, based on reconfigurable silicon-on-insulator waveguides that operate at telecom wavelengths. This, together with our telecom single-photon source and highly efficient superconducting nanowire single-photon detectors, provides a versatile and stable platform for a high-fidelity implementation of counterfactual communication with single photons, allowing us to actively tune the number of steps in the Zeno measurement, and achieve a bit error probability below 1%, without post-selection and with a vanishing weak trace. Our demonstration shows how our programmable nanophotonic processor could be applied to more complex counterfactual tasks and quantum information protocols.

npj Quantum Information (2019)5:61 ; https://doi.org/10.1038/s41534-019-0179-2

INTRODUCTION

Interaction-free measurements allow one to measure whether or not an object is present without ever interacting with it.[1] This is made clear in Elitzur and Vaidman's well-known bomb-testing gedanken experiment.[2] In this experiment, a single photon used in a Mach-Zehnder interferometer (MZI) sometimes reveals whether or not an absorbing object (e.g., a bomb) had been placed in one of the interferometer arms, without any interaction between the photon and the bomb. It was later shown that the quantum Zeno effect, wherein repeated observations prevent the system from evolving,[3,4] can be used to bring the success probability of this protocol arbitrarily close to unity.[3–6] Such protocols are often referred to as "counterfactual", and have now been applied to quantum computing,[7] quantum key distribution[8–10] and communication.[11,12] Here, we experimentally implement a counterfactual communication (CFC) protocol where information can propagate without being carried by physical particles.

The first suggested protocol for CFC was developed by Salih et al., and it is based on a chain of nested MZIs.[11,13] Following its publication, this fascinating protocol has been subject to both intense criticism and vigorous defense. There are four main points of discussion: (1) Achieving a high success probability (say > 95%) requires thousands of optical elements.[11,12,14] (2) An analysis of the Fisher information flow indicates that to retain counterfactuality in Salih's protocol, perfect quantum channels are needed.[15] (3) If one performs a weak measurement in Bob's lab, one can detect the presence of photons that are later found in Alice's laboratory. Some authors have argued that the presence of the "weak trace" renders the counterfactuality of the protocol invalid,[16–19] but others have dismissed the weak trace as a consequence of the unwanted weak measurement's disturbance.[20–22] (4) Unless operated in the theoretical limit of infinite optical operations, this scheme requires post-selection to remove the CFC violations.[13,21,23] It has recently been shown that also a classical communication protocol can be counterfactual if post-selection is allowed.[14]

To circumvent these issues, we implement a novel CFC protocol[12] that does not need post-selection and requires orders of magnitude fewer optical elements than nested MZI protocols. In our scheme single photons travel from Alice to Bob but information from Bob to Alice; this has been dubbed type-II CFC, in contrast to type-I schemes, where the photon should remain with Alice throughout the protocol.[15] In both types the particles and the information never co-propagate, thereby making the communication counterfactual. Note that the very recent proposals[23,24] discussing means of making the Salih scheme trace-free still require the post-selected removal of non-counterfactual events, as well as thousands of ideal optical operations.

[1]Faculty of Physics, Vienna Center for Quantum Science and Technology (VCQ), University of Vienna, Boltzmanngasse 5, A-1090 Vienna, Austria; [2]Department of Physics, Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, UK; [3]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; [4]Quantum Photonics Group, RLE, Massachusetts Institute of Technology, Cambridge, MA 02139, USA and [5]Elenion Technologies, New York, NY 10016, USA
Correspondence: I. Alonso Calafell (irati.alonso.calafell@univie.ac.at) or T. Strömberg (teodor.stroemberg@univie.ac.at)
These authors contributed equally: I. Alonso Calafell, T. Strömberg

Received: 4 October 2018 Accepted: 27 June 2019
Published online: 23 July 2019

Published in partnership with The University of New South Wales

npj nature partner journals

RESULTS

We perform our experiment using telecom single-photons in a state-of-the-art programmable nanophotonic processor (PNP),[25] which is orders of magnitude more precise and stable than previous bulk-optic approaches.[5,6] Our PNP also provides unprecedented tunability, which we use to investigate the scaling of the protocol by changing the number of chained interferometers. By combining the novel CFC protocol with our advanced photonic technology, we are able to implement counterfactual communication with a bit success probability above 99%, without post-selection. As in previous CFC protocols, the interferometer implementing the quantum Zeno effect is shared between Bob's laboratory and the fully passive transmission channel. In contrast to these protocols, ours protocol allows for Alice's laboratory to be situated outside the interferometer. As a result, even our proof-of-principle demonstration would allow for counterfactual communication over arbitrary distances, even if the region in which the non-local information transfer takes place is bounded by the size of the PNP.

Our protocol uses a series of N beamsplitters with reflectivity $R = \cos^2(\pi/2N)$, which, together with mirrors, form a circuit of $N-1$ chained MZIs. As shown in Fig. 1, the communication protocol begins with Alice injecting a single photon into her input port. If Bob wants to send a logic 0 he leaves his mirrors in place, causing the photon to self interfere such that it exits in D_B with unit probability (Fig. 1a). To send a logic 1 Bob *locally* modifies the circuit to have the upper paths open (Fig. 1b). In this case the photon will successfully reflect off of all the beamsplitters and exit in D_A with probability R^N. Removing the mirrors effectively collapses the wavefunction after every beamsplitter, suppressing interference and implementing the Zeno effect. The probability that the photon remains in the lower arm after N beamsplitters can be made arbitrarily high by increasing N (and changing the reflectivities accordingly).

Since any implementation is restricted to a finite number of beamsplitters, there will be a probability for a photon to exit the wrong port when Bob tries to send a logic 1. This error probability is a function that decreases with N as $P_{1,err} = 1 - R(N)^N$. In the non-ideal case, optical losses in the system will increase this probability further. The errors associated with Bob's attempt to transmit a logic 0 are of a different nature. In theory, he can always perfectly transmit a logic 0, independent of N; that is, $P_{0,err} = 0$. In practice, however, imperfections in the interferometers will lead to cases in which the photon re-enters Alice's laboratory and she incorrectly records a logic 1. This leads to a rare counterfactual violation, as the wavefunction "leaks" from Bob's to Alice's laboratory,[15] leaving a weak trace in Bob's lab, while the photon is detected in Alice's laboratory. The high-fidelity operations enabled by our PNP allows us to make the probability of such violations vanishingly small. Although they do not contribute to a counterfactual violation, dark counts in Alice's detector will also increase this error rate.

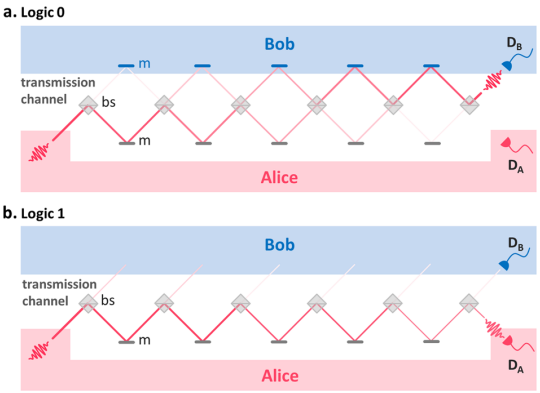


Fig. 1 Architecture of the chained MZI protocol. Alice inputs a photon into the transmission channel, consisting of a row of beamsplitters (BSs) and the lower row of mirrors (marked with an 'm'). **a** If Bob intends to send a logic 0, he places mirrors in his laboratory to form MZIs that span his lab and the transmission channel, creating constructive interference in Bob's port (D_B). **b** If he intends to send a logic 1, he removes the mirrors, causing the photons to arrive back in Alice's laboratory (D_A) with high probability

We can overcome the bit errors by encoding each logical bit into M single photons, at the cost of slightly increasing the CFC violation. If Alice sends M photons into the transmission channel without detecting any at D_A, she will record a logic 0. On the other hand, if she detects one or more photons in her laboratory, she will record a logic 1. Assuming messages with a balanced number of 0s and 1s, the average bit error probability is given by:

$$\bar{P}_{\mathrm{err}}(M) = \frac{1}{2}\left[(P_{1,\mathrm{err}})^M + MP_{0,\mathrm{err}}\right] \quad (1)$$

where the second term is an approximation of $1 - P_0^M$ valid for small values of $MP_{0,\mathrm{err}}$. By increasing M we can thus decrease the contributions of $P_{1,\mathrm{err}}$ exponentially while only increasing those of $P_{0,\mathrm{err}}$ linearly. The counterfactual violation probability for a random bit is given by

$$\bar{P}_{\mathrm{CFC}}(M) = \frac{1}{2\eta}MP_{0,\mathrm{err}}, \quad (2)$$

where η is the detector efficiency. We can thus find an M that minimizes the average bit error, while also maintaining a low counterfactual violation probability. In our experiment this expression slightly overestimates the violation probability, as it includes the detector dark counts.

As illustrated in Fig. 2, we implement a series of chained MZIs using a PNP. At the intersections of each of the modes shown in the figure there are smaller MZIs that act as beamsplitters with tunable reflectivities and phases. Since each of the MZIs is completely tunable, we were able to implement our CFC protocol using two to six concatenated beamsplitters on the same photonic chip. Given the layout of our chip, six is the maximum number of beamsplitter that we can concatenate. In addition, the high interferometric visibility of the PNP, which we measure to be 99.94% on average, allows us to keep the rate of counterfactual violations low, without post-selection. The single photons are generated in a spontaneous parametric down conversion process and detected using superconducting nanowire single-photon detectors with detection efficiencies $\eta \sim 90\%$ (see Methods).

DISCUSSION

To study the performance of this CFC protocol we measure the average bit error, as a function of the number of photons in which the bit is encoded, M, for five different values of N number of BSs. For the logic 0, we configure the MZIs in Bob's laboratory as mirrors (see Fig. 2), while for the logic 1 we let the MZIs in Bob's laboratory act as SWAP gates, routing the light out of the interferometer chain. Since Alice cannot access detector D_B, she assumes that a photon is injected in the transmission channel every time she detects a heralding photon in D_H. We thus run the measurement until we have M recorded single-photon events in D_H (typical rates were 1.1 MHz) and look for the coincidences that these events have with D_A within a set coincidence window $\Delta\tau =$ 2.5 ns that is shorter than the pulse separation. Our heralding efficiency was ~3% through the PNP.

Figure 3a shows the experimental average error probability of our CFC protocol as a function of M for different N. We also include a theoretical calculation of the expected error probabilities, which considers the heralding efficiency of the single photons and the success probability of the interferometer that is in good

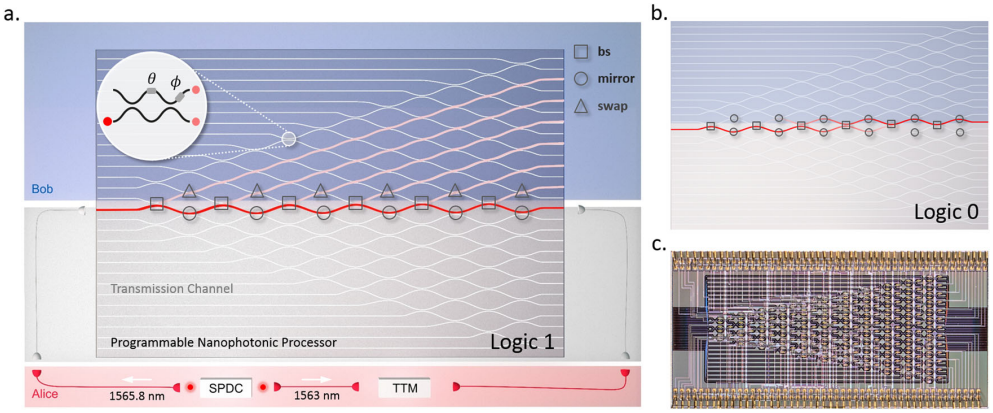


Fig. 2 Experimental setup. **a** Our experiment is implemented in a programmable nanophotonic processor (PNP), which is composed of 26 interconnected waveguides. The waveguides are coupled by 88 Mach-Zehnder interferometers (MZIs), as indicated by the top-left inset. Each MZI is equipped with a pair of thermo-optic phase shifters, which allows us to treat them as beamsplitters with fully tunable reflectivities (set via $\theta \in [0, 2\pi]$) and phases (($\phi \in [0, 2\pi]$). In our work, we set θ to π, 0 or $\pi/2N$, to implement mirrors (circles), SWAPs (triangles) or beamsplitters (squares), respectively. In Alice's laboratory (the pink shaded region) a spontaneous parametric down-conversion source creates a frequency non-degenerate photon pair at $\lambda_H = 1563$ nm and $\lambda_T = 1565.8$ nm. Detection of the λ_H photon in detector H heralds the λ_T photon that is injected into the transmission channel. This channel is comprised of the lower half of the PNP, in which MZIs are set to act as mirrors, as well as the MZIs that couple the upper and lower half of the waveguide. The latter of these MZIs are configured to act as beamsplitters, whose reflectivity varies with N (the number of beamsplitters used in the protocol) as $R(N) = \cos^2(\pi/2N)$. Bob's laboratory consists of the upper set of MZIs (blue shaded area), which he can set as mode swaps to send a logic 1 or **b** as mirrors to send a logic 0. Thus in total we used 48 MZIs: 6 to implement the tunable beamsplitters, 30 to implement the loss channels, 6 for the mirrors in the transmission channel, and 6 for the mode swaps/mirrors in Bob's laboratory. The photons are detected in Alice's laboratory by superconducting nanowire single-photon detectors with detection efficiencies of approximately 90%. Coincident detection events are recorded with a custom-made Time Tagging Module (TTM). **c** Micrograph of the PNP with dimensions 4.9 × 2.4 mm

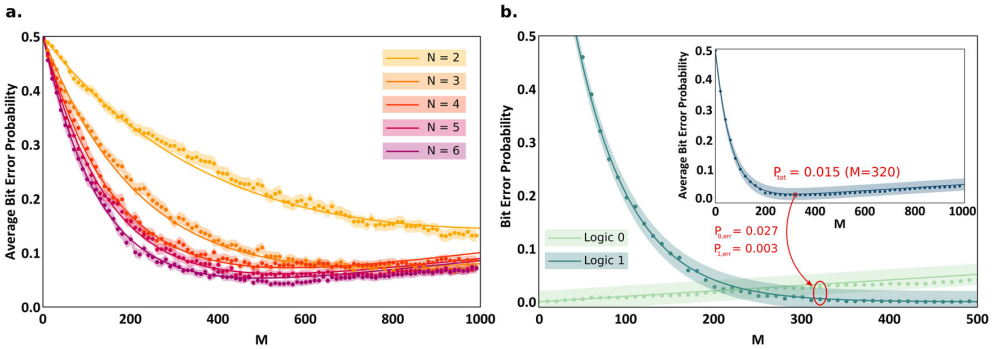


Fig. 3 Success probabilities of the CFC communication. The curves are theoretical models of our experiment with no free parameters, and the points are experimental data. **a** Measured average bit error (as defined in the main text) of the protocol for different number of beamsplitters (N) as a function of the number of photons (M) used to encode each bit. For small M the $\cos^{2N}(\pi/2N)$ dependence of the logic 1 error dominates the average error, making the latter decrease with M as expected. As M is increased more, the linearly growing error in the logic 0, caused by imperfect destructive interference in Alice's port (D_A), starts to dominate. **b** In the $N = 6$ case, the optimization of the interferometer fidelity and heralding efficiency leads to an average bit error rate of 1.5% for $M = 320$, where the average CFC violation probability is 2.4%

agreement with the experimental data. Note that these are not fits to the data, but rather models with no free parameters. As theoretically predicted, the error rate of the logic 1 decreases exponentially with increasing M and the error rate of the logic 0 increases linearly with M. We observe that higher N requires smaller M, and also results in lower bit error probabilities.

The success probability of this CFC scheme is highly sensitive to the fidelity of the interferometers and the overall heralding efficiency, which depends on the single-photon source and the coupling efficiency throughout the system. Hence, we optimized the setup for the $N = 6$ case. Figure 3b shows the corresponding error probability of the logic 1 and the logic 0. The inset in Fig. 3b shows the average error probability, where we find a minimum of 1.5% for $M = 320$, while the average counterfactual violation is kept at 2.4%. Owing to backscattering in Bob's laboratory (i.e., imperfect SWAP operations) small "amounts" of wavefunction amplitude leak back into the transmission line in the 1 bit process. Although these do not all lead to detection events in Alice's laboratory, the sum of their squares provides an upper bound on the probability of a counterfactual violation. We estimate that the probability for a photon to reflect off of a SWAP operation is at most 1%. Hence, in our experiment (Fig. 4) with $M = 320$ and $N =$

Fig. 4 Image sent from Bob to Alice. The bits are encoded in different numbers of single photons $M = \{10, 50, 320, 500\}$. The white and black pixels are defined to correspond to logic 1 and logic 0, respectively. The success probability increases with increasing M, reaching 99% for $M = 320$. The CFC violation probability (P_{CFC}) also increases with increasing M, but it remains as low as 0.6% for $M = 320$. Note that this CFC violation comes only from the logic 0 errors, which we can directly measure; the total CFC violation would include a small portion of successful logic 1 events, as discussed in the main text. Increasing M beyond 320 increases the success probability at the expense of increasing the CFC violation. As it can be observed, these probabilities are directly related to the transmission fidelity (F) of the white pixels, which increases with M, and the transmission fidelity of the black pixels, which decreases with M

6, the weak trace is vanishingly small and the contribution from the logic 1 to a CFC violation is less than 1.1%. Note that this violation probability decreases with N, even if the errors remain the same.

To demonstrate the performance of the communication protocol we proceed to analyze the quality of a message in the form of a black and white image, sent from Bob to Alice, for $N = 6$ and $M = \{10, 50, 320, 500\}$. We arbitrarily define the white and black pixels of the image as logic 1 and logic 0, respectively.

Figure 4 shows the message transmitted from Bob to Alice for different numbers of encoding photons. We define the image fidelity as

$$F = \sum_{i=1}^{T} \frac{1 + (-1)^{A_i + B_i}}{2T} \tag{3}$$

where B_i is the bit that Bob intended to send, A_i is the bit that Alice recorded, and T is the total number of bits in the image. In this case we define the CFC violation probability as the number of incorrectly transmitted logic 0 s (black pixels) over T. The encoding using $M = 10$ is clearly not enough to overcome the losses of the system, with a very low image fidelity of 31.77%. As we increase M, the success probability and legibility of the message increases (the individual fidelities are listed below each panel). The image fidelity reaches 99.09% at $M = 320$, at which point the CFC violation probability from 0 bit errors remains as low as 0.6%. For $M = 500$ the image fidelity does not noticeably change; however, the CFC violation increases slightly. If the CFC violation of the 1 bit (caused by on-chip beamsplitter imperfections) is accounted for, the CFC violation at $M = 320$ increases to 2.3%. Note that these values are lower than the value in Fig. 3b due to the unbalanced distribution of black and white pixels in the image.

Our high-fidelity implementation of a counterfactual communication protocol without post-selection was enabled by a programmable nano-photonic processor. The high (99.94%) average visibility of the individual integrated interferometers allowed bit error probabilities as low as 1.5%, while, at the same time, keeping the probability for the transmission of a single bit to result in a counterfactual violation below 2.4%. By combining our state-of-the-art photonic technology with a novel theoretical proposal we contradicted a crucial premise of communication theory:[26] that a message is carried by physical particles or waves. In fact, our work shows that "interaction-free non-locality", first described by Elitzur and Vaidman,[2] can be utilized to send information that is not necessarily bound to the trajectory of a wavefunction or to a physical particle. In addition to enabling further high-fidelity demonstrations of counterfactual protocols, our work highlights the important role that technological advancements can play in experimental investigations of fundamentals of quantum mechanics and information theory. We thus anticipate nanophotonic processors, such as ours, to be central to future photonic quantum information experiments all the way from the foundational level to commercialized products.

METHODS

Telecom photon source

We use a pulsed Ti:Sapphire laser with a repetition rate of 76 MHz, an average power of 0.2 W, a central wavelength of 782.2 nm, and a pulse duration of 2.1 ps. The repetition rate is doubled via a passive temporal multiplexing stage.[27,28] This beam pumps a periodically poled KTP crystal phase matched for collinear type-II spontaneous parametric down conversion, generating frequency non-degenerate photon pairs at $\lambda_H = 1563$ nm, $\lambda_T = 1565.8$ nm. Registering the shorter wavelength photon at the detector D_H heralds the presence of the longer wavelength one, which is sent to the waveguide.

Programmable nanophotonic processor

Our chained Mach-Zehnder interferometers (MZIs) are implemented in a silicon-on-insulator (SOI) programmable waveguide, developed by the Quantum Photonics Laboratory at the Massachusetts Institute of Technology.[25] The device consists of 88 MZIs, each accompanied by a pair of thermo-optic phase shifters that facilitate full control over the internal and external phases of the MZIs. The phase shifters are controlled by a 240-channel, 16-bit precision voltage supply, allowing for a phase precision higher than 250 µrad. The switching frequency of these phase shifters is 130 kHz. The coupling of the single photons in/out of the chip is performed using two Si_3N_4–SiO_2 waveguides manufactured by Lionix International, that adiabatically taper the 10×10 µm mode from the single mode fiber down to 2×2 µm, matching the mode field diameter of the programmable waveguide at the input facet. The total insertion loss per facet was measured as low as 3 dB.

Superconducting nanowire single-photon detectors

The photons are detected using superconducting nanowire single-photon detectors.[29,30] These detectors are produced by photonSpot and are optimized to reach detection efficiencies ~90% at telecom wavelengths.

DATA AVAILABILITY

The datasets generated and analyzed during the current study are available from the corresponding author if you ask nicely.

ACKNOWLEDGEMENTS

I.A.C. and T.S. acknowledge support from the University of Vienna via the Vienna Doctoral School. L.A.R. acknowledges support from the Templeton World Charity Foundation (fellowship no. TWCF0194). P.W. acknowledges support from the European Commission through ErBeSta (No. 800942), the Austrian Research Promotion Agency (FFG) through the QuantERA ERA-NET Cofund project HiPhoP, from the Austrian Science Fund (FWF) through CoQuS (W1210-N25), BeyondC (F7113-N38) and NaMuG (P30067-N36), the U.S. Air Force Office of Scientific Research

(FA2386-233 17-1-4011), and Red Bull GmbH. D.R.M.A.S. acknowledges support from the EPSRC, Hitachi Cambridge, Lars Hierta's Memorial Foundation and the Sweden-America Foundation. N.H. was supported in part by the Air Force Research Laboratory RITA program (FA8750-14-2-0120); Research program FA9550-16-1-0391, supervised by Gernot Pomrenke; and D.E. acknowledges partial support from the Office of Naval Research CONQUEST program. The authors would like to express their gratitude towards J. Zeuner for helpful discussions and T. Rögelsperger for the artistic input. They furthermore thank LioniX International BV for the manufacturing of the interposer waveguides used in the experiment.

AUTHOR CONTRIBUTIONS

I.A.C., T.S., D.R.M.A.-S., L.A.R. and P.W. designed the experiment, analyzed the results, and wrote the paper. I.A.C., T.S., and L.A.R. performed the measurements. D.R.M.A.-S. and C.H.W.B. provided theoretical support. V.S. and C.G. built the telecom photon pair source. N.C.H., J.C., M.H, T.B.-J., and D.E. fabricated and characterized the nanophotonic processor. All authors read and commented on the paper.

ADDITIONAL INFORMATION

Competing interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Dicke, R. H. Interaction-free quantum measurements: a paradox? *Am. J. Phys.* **49**, 925–930 (1981).
2. Elitzur, A. C. & Vaidman, L. Quantum mechanical interaction-free measurements. *Found. Phys.* **23**, 987–997 (1993).
3. Degasperis, A., Fonda, L. & Ghirardi, G. C. Does the lifetime of an unstable system depend on the measuring apparatus? *Il Nuovo Cim. A (1965–1970)* **21**, 471–484 (1974).
4. Misra, B. & Sudarshan, E. C. G. The zeno's paradox in quantum theory. *J. Math. Phys.* **18**, 756–763 (1977).
5. Kwiat, P., Weinfurter, H., Herzog, T., Zeilinger, A. & Kasevich, M. A. Interaction-free measurement. *Phys. Rev. Lett.* **74**, 4763–4766 (1995).
6. Kwiat, P. G. et al. High-efficiency quantum interrogation measurements via the quantum zeno effect. *Phys. Rev. Lett.* **83**, 4725–4728 (1999).
7. Hosten, O., Rakher, M. T., Barreiro, J. T., Peters, N. A. & Kwiat, P. G. Counterfactual quantum computation through quantum interrogation. *Nature* **439**, 949–952 (2006).
8. Noh, T.-G. Counterfactual quantum cryptography. *Phys. Rev. Lett.* **103**, 230501 (2009).
9. Yin, Z.-Q., Li, H.-W., Chen, W., Han, Z.-F. & Guo, G.-C. Security of counterfactual quantum cryptography. *Phys. Rev. A* **82**, 042335 (2010).
10. Liu, X. et al. Eavesdropping on counterfactual quantum key distribution with finite resources. *Phys. Rev. A* **90**, 022318 (2014).
11. Salih, H., Li, Z.-H., Al-Amri, M. & Zubairy, M. S. Protocol for direct counterfactual quantum communication. *Phys. Rev. Lett.* **110**, 170502 (2013).
12. Arvidsson-Shukur, D. R. M. & Barnes, C. H. W. Quantum counterfactual communication without a weak trace. *Phys. Rev. A* **94**, 062303 (2016).
13. Cao, Y. et al. Direct counterfactual communication via quantum zeno effect. *Proc. Natl Acad. Sci.* **114**, 4920–4924 (2017).
14. Arvidsson-Shukur, D. R. & Barnes, C. H. Postselection and counterfactual communication. *Phys. Rev. A* **99**, 060102 (2019).
15. Arvidsson-Shukur, D. R. M., Gottfries, A. N. O. & Barnes, C. H. W. Evaluation of counterfactuality in counterfactual communication protocols. *Phys. Rev. A* **96**, 062316 (2017).
16. Danan, A., Farfurnik, D., Bar-Ad, S. & Vaidman, L. Asking photons where they have been. *Phys. Rev. Lett.* **111**, 240402 (2013).
17. Vaidman, L. Past of a quantum particle. *Phys. Rev. A* **87**, 052104 (2013).
18. Vaidman, L. Reply to comment on 'past of a quantum particle'. *Phys. Rev. A* **88**, 046103 (2013).
19. Vaidman, L. Comment on protocol for direct counterfactual quantum communication. *Phys. Rev. Lett.* **112**, 208901 (2014).
20. Li, Z.-H., Al-Amri, M. & Zubairy, M. S. Comment on past of a quantum particle. *Phys. Rev. A* **88**, 046102 (2013).
21. Salih, H., Li, Z.-H., Al-Amri, M. & Zubairy, M. S. Salih et al. reply. *Phys. Rev. Lett.* **112**, 208902 (2014).
22. Li, Z.-H., Al-Amri, M. & Zubairy, M. S. Direct counterfactual transmission of a quantum state. *Phys. Rev. A* **92**, 052315 (2015).
23. Aharonov, Y. & Vaidman, L. Modification of counterfactual communication protocols that eliminates weak particle traces. *Phys. Rev. A* **99**, 010103 (2019).
24. Salih, H., McCutcheon, W. & Rarity, J. Do the laws of physics prohibit counterfactual communication? Preprint at arXiv: 1806.01257 (2018).
25. Harris, N. C. et al. Quantum transport simulations in a programmable nanophotonic processor. *Nat. Photonics* **11**, 447 (2017).
26. Shannon, C. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
27. Greganti, C. et al. Tuning single-photon sources for telecom multi-photon experiments. *Opt. Express* **26**, 3286–3302 (2018).
28. Broome, M. A., Almeida, M. P., Fedrizzi, A. & White, A. G. Reducing multi-photon rates in pulsed down-conversion by temporal multiplexing. *Opt. Express* **19**, 22698–22708 (2011).
29. Natarajan, C. M., Tanner, M. G. & Hadfield, R. H. Superconducting nanowire single-photon detectors: physics and applications. *IOPscience* **25**, 063001 (2012).
30. Marsili, F. et al. Detecting single infrared photons with 93% system efficiency. *Nat. Photon.* **7**, 210–214 (2013).

9 Reprint of "Demonstration of universal time-reversal for qubit processes"

9.1 Contribution statement

The experimental side of the work presented in this publication was carried out in equal measure by Peter Schiansky and myself. This included the planning of the experiment, the assembling of the experimental setup, the characterisation of optical elements, and the acquisition of data. The methods used in the data analysis were independently implemented by both Peter Schiansky and myself as a validation measure. Peter Schiansky took more responsibility for the automation of the experiment and the efficient pre-processing of measurement data, and I carried out simulations in support of the experiment. The manuscript was written by Peter Schiansky and myself, together with Miguel Navascues, David Trillo and Philip Walther. All figures were made by Peter Schiansky and myself, aside from the illustration of a dandelion, which was commissioned from an artist. David Trillo and Miguel Navascues developed the theoretical idea and the mathematical proofs. The experiment was conceived by Miguel Navascues, Peter Schiansky and Philip Walther.

OPTICA

Demonstration of universal time-reversal for qubit processes

P. Schiansky,[1,3,†] T. Strömberg,[1,†] D. Trillo,[2] V. Saggio,[1] B. Dive,[2] M. Navascués,[2] and P. Walther[1,4]

[1]University of Vienna, Faculty of Physics & Research Network Quantum & Aspects of Space Time (TURIS), Boltzmanngasse 5, 1090 Vienna, Austria
[2]Institute for Quantum Optics and Quantum Information, Boltzmanngasse 3, 1090 Vienna, Austria
[3]e-mail: peter.schiansky@univie.ac.at
[4]e-mail: philip.walther@univie.ac.at

Received 29 June 2022; revised 24 November 2022; accepted 20 December 2022; published 26 January 2023

In quantum mechanics, the unitary nature of time evolution makes it intrinsically reversible, given control over the system in question. Remarkably, there have been several recent demonstrations of protocols for reverting unknown unitaries in scenarios where even the interactions with the target system are unknown. These protocols are limited by their probabilistic nature, raising the fundamental question of whether time-reversal could be performed deterministically. Here we show that quantum physics indeed allows for this by exploiting the non-commuting nature of quantum operators, and demonstrate a recursive protocol for two-level quantum systems with an arbitrarily high probability of success. Using a photonic platform, we achieve an average rewinding fidelity of over 95%. Our protocol, requiring no knowledge of the quantum process to be rewound, is optimal in its running time, and brings quantum rewinding into a regime of practical relevance.

https://doi.org/10.1364/OPTICA.469109

1. INTRODUCTION

In the macroscopic world, there is an apparent unidirectionality of processes in time, which stands in contrast to the time-reversal symmetric nature of the underlying laws of physics. This tension was first pointed out by Eddington, who coined the term "arrow of time" to describe the asymmetry [1]. In classical physics, an arrow of time emerges through the second law of thermodynamics, giving rise to processes that cannot be reversed [2]. Due to the statistical nature of the law, and the determinism of classical physics, the irreversibility is not fundamental. Indeed, for classical wave mechanics, it is well known that the time evolution of a system can be reversed without any knowledge of the dynamics through a technique called phase conjugation [3,4]. In the microscopic quantum realm, however, the ability to perform phase conjugation becomes limited by fundamental quantum noise [5], due to the non-unitary nature of the process. It has therefore remained an open question whether or not the dynamics of quantum systems can be reversed in a universal manner.

Recently, there have been several works addressing this question, in which probabilistic protocols for "rewinding" quantum systems were presented [6,7] and demonstrated in a laboratory setting [8,9]. These protocols work independently of both the free Hamiltonian guiding the time evolution of the system in question, and the system's interaction with the experimental apparatus. A major drawback of the protocols in [7] is that they suffer from low success probabilities, typically of the order of 10^{-3}. The scheme in [6], on the other hand, allows for a form of error correction, whereby the protocol can be repeated when it fails. However, it is not known whether these feed-forward corrections can boost the success probability arbitrarily close to one. Moreover, the protocol cannot rewind a target system in "real time," instead taking three units of time for every one rewound.

The teleportation based protocol in [10,11], as well as more traditional methods to rewind a quantum system with an unknown free Hamiltonian, such as the refocusing techniques used in nuclear magnetic resonance [12], require the ability to implement controlled operations that are specifically tailored to the target quantum system, and are therefore not universal. The work of [13] combines both quantum theory and general relativity to devise a "time translator," capable of rewinding or fast-forwarding quantum systems. While this method can time-translate any quantum system, it has two drawbacks: (1) it works only approximately, and under a restriction on the free Hamiltonian of the target; (2) if we demand reasonable precision, the probability of success of the process becomes astronomically small.

In this paper, we demonstrate a novel universal time-reversal protocol (Fig. 1) for which the success probability can be made arbitrarily high, making it, in effect, deterministic. At its heart, the protocol is based on the non-commutativity of quantum operators, a core concept in quantum mechanics. This conceptual simplicity, which translates directly into a straight-forward implementation in the laboratory based on the recently developed

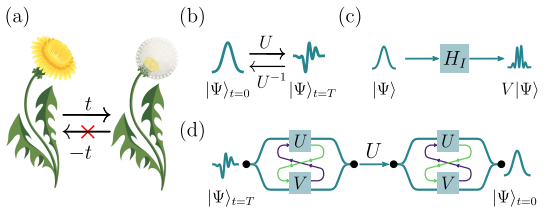


Fig. 1. (a) In the classical world, there is an unmistakable directionality to time, illustrated here through the process of aging, a process that cannot be reversed in practice despite its deterministic nature. In this paper, we show that these same limitations do not apply in the quantum realm. (b) The unitarity of quantum mechanics guarantees that an inverse of a given time evolution U always exists, even though it may be unknown. (c) By letting a target quantum system pass through an interaction region, a perturbed time evolution V can be realized. (d) A quantum switch makes the target system evolve in a superposition of its free evolution U and perturbed evolution V. This superposition of time evolutions can be used to "rewind" the system backwards in time, without requiring any knowledge about either U, V or the state $|\Psi\rangle$.

quantum switch [14,15], allows us to overcome the limitations of previous proposals. The quantum switch is a process that, in its simplest form, transforms two unitaries into a controlled superposition of the two gates being applied in different orders: $(U, V) \to UV \otimes |0\rangle\langle 0|_C + VU \otimes |1\rangle\langle 1|_C$.

In this work, the utilization of quantum switches allows us to time-translate the unknown internal degree of freedom of a target system by setting it on a superposition of different trajectories. For some of these trajectories, the free evolution U of the target's internal degree of freedom is perturbed by an unknown but repeatable interaction, which induces an evolution V on the target. This perturbation can be achieved by any physical interaction and thus can be applied to every possible quantum system. We make these trajectories sequentially interfere in such a way that the final state of the target's internal degree of freedom is propagated by U^{-n}, for some positive n, independently of the operators U, V. Each quantum switch requires a projection of the target system's path degree of freedom to induce the desired superposition of time evolutions. An advantage of our scheme is that even in the event that the projection fails, a simple and repeatable error-correction procedure can be applied, yielding an arbitrarily high success rate, as long as $[U, V] \neq 0$. It is also worth emphasizing that the protocol runs in real time, meaning that the time it takes to rewind the system is equal to the amount of time to be rewound, aside from a bounded overhead.

We demonstrate the universality of our protocol by running it on a large set of different time evolutions. Our demonstration utilizes a quantum photonics platform with control of path and polarization degrees of freedom of single photons. We generate a discrete time evolution of a single photon by implementing a "polarization Hamiltonian" using a combination of half- and quarter-wave plates. A superposition of time evolutions is achieved via an interferometric quantum switch in which the propagation direction defines the order of the evolutions U and V. Our setup uses two fast optical switches that allow the quantum switch to be accessed several times.

2. METHODS

A. Protocol

In this section, we will give a description of how the rewinding protocol works in a photonic setting, the basic steps of which are illustrated in Fig. 2. An alternative formulation using a scattering scenario is given in Supplement 1. A full description, as well as the accompanying proofs, can be found in [16]. Given an unknown target system $|\Psi\rangle$, whose time evolution is described by $U = e^{-i\Delta T H_0}$, where H_0 is an unknown Hamiltonian, our goal will be to rewind the system: $|\Psi(t = n\Delta T)\rangle \to |\Psi(t = 0)\rangle$, where n is the number of discrete time steps to be rewound. The basis of our protocol is the following identity [7]:

$$[U, V]U^n[U, V] \propto U^{-n}. \tag{1}$$

Here U, V are any 2×2 matrices, with U being invertible. When the matrix U describes the time evolution of a system, we see that an experimenter able to implement a commutator can reverse the time evolution, even if U is unknown. The basic protocol is thus as follows: apply the commutator between the time evolution operator U and any other 2×2 matrix V, let the system evolve freely for the amount of time to be rewound, then apply the commutator again. The matrix V represents, in the general setting, a time evolution that is perturbed by any repeatable means, for example, by bringing the target on a trajectory that leads it through some interaction region. This perturbed evolution can also remain unknown; however, the magnitude of the commutator $[U, V]$ affects the success probability of a single-step attempt to rewind the system.

In a photonic setting, a commutator can be realized using a quantum switch acting on two degrees of freedom of a single photon. The control qubit, defining the order of gate operations, is encoded in the photon's path, while the target qubit is encoded in the polarization. The two possible gate orders, UV and VU, are superposed by initializing the control qubit in the superposition state $(|0\rangle_C + |1\rangle_C)/\sqrt{2}$ and then applying a controlled operation between the control and target systems [15]:

$$|0\rangle_C \otimes |\Psi\rangle_T \to \frac{|0\rangle_C + |1\rangle_C}{\sqrt{2}} \otimes |\Psi\rangle_T \to$$

$$\frac{1}{\sqrt{2}}\Big[|0\rangle_C \otimes UV|\Psi\rangle_T + |1\rangle_C \otimes VU|\Psi\rangle_T\Big]. \tag{2}$$

By applying a Hadamard gate to the control qubit, one obtains the following state:

$$|0\rangle_C \otimes \frac{1}{2}\{U, V\}|\Psi\rangle_T + |1\rangle_C \otimes \frac{1}{2}[U, V]|\Psi\rangle_T. \tag{3}$$

A measurement of the control qubit now projects the target state onto either the commutator or the anticommutator, where the latter is denoted by $\{\cdot\}$. If the measurement outcome of the control qubit is $|0\rangle$, the anticommutator is applied by the quantum switch, but the protocol does not necessarily fail. Instead, the following matrix identities can be used to correct the error:

$$\{U, V\}^m[U, V]\{U, V\}^m \propto [U, V], \tag{4}$$

$$[U, V]^2 \propto \mathbb{1}. \tag{5}$$

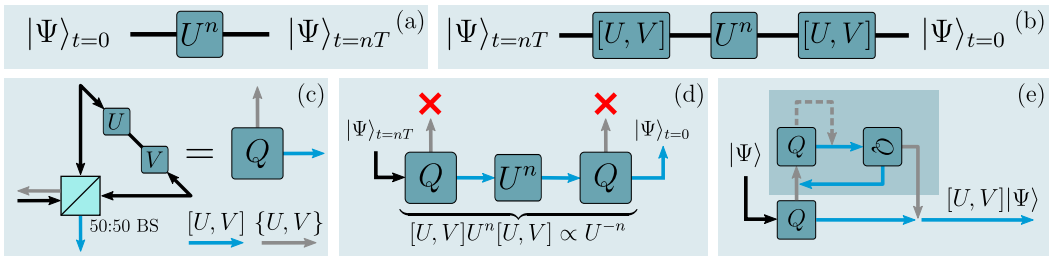


Fig. 2. Interferometric protocol. (a) Circuit diagram for the free evolution U^n of a target quantum from time $t = 0$ to $t = nT$. (b) Symbolic circuit diagram for the rewinding protocol. The target quantum system is made to propagate backwards in time by way of the identity: $U^{-1} \propto [U, V]U[U, V]$, where V is a perturbed evolution. (c) Q is a quantum switch, pictured here as a Sagnac interferometer, and acts as the basic building block of the interferometric scheme, probabilistically applying the commutator $[U, V]$. (d) A full interferometric implementation of the single-step protocol, which succeeds whenever the photon, in which the target state $|\Psi\rangle$ is encoded, exits both quantum switches in the commutator port (blue arrow). Detecting a photon in the anticommutator port (gray arrow) heralds a failure of the quantum switch. (e) Adaptive error correction for achieving an arbitrarily high probability of success. This entire diagram replaces a single quantum switch in (d). Instead of detecting the failure mode of the quantum switch, the photon is made to re-interfere with itself. Whenever it exits in the bottom right, the commutator $[U, V]$ will have been applied (see [16]) The dashed path represents recursive applications of the diagram, through which the success probability can be made arbitrarily high, while the darker shaded area indicates the additional quantum switches needed.

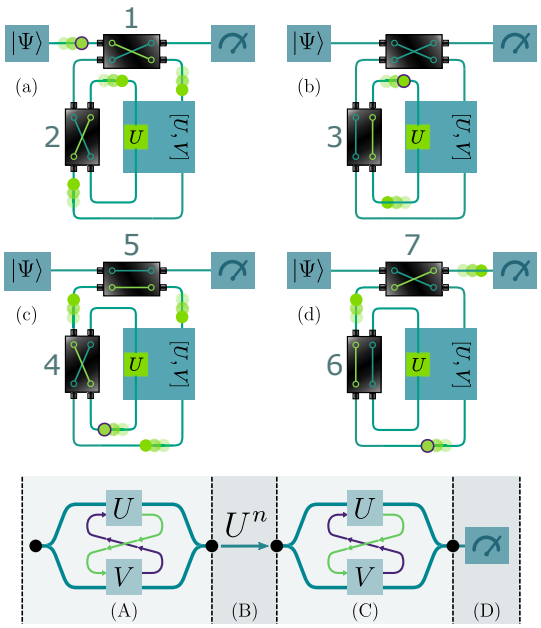


Fig. 3. Active photon routing. The use of electro-optical (EO) switches enables active routing of a single photon (green dot) encoding the target quantum state. The settings of the EO-switches determine whether the photon passes through the quantum switch, evolves freely, or is sent to a detector. Sub-diagrams (a)–(d) indicate the states of EO-switches at different steps of the protocol (A)–(D), illustrated at the bottom of the figure. The numbers index the order in which the EO-switches are traversed. In each sub-diagram, the photon's initial position is indicated by a contour; the subsequent dots represent the photon at a slightly later time, and the green trace shows the photon path through a given switch. (a) The photon passes through the quantum switch for the first time. (b) For $n \geq 2$, the photon is trapped in a loop until the free time-evolution operator U has been applied a total of n times. (c) The photon passes through the quantum switch a second time. (d) The EO-switches direct the photon to a quantum tomography stage.

Through recursive application of these identities, an anticommutator can always be turned into a commutator. This process can be described using a virtual road map, illustrated in Fig. 2(e). In [16], some of us prove that when U, V are unitary and $[U, V] \neq 0$, the protocol always terminates in a finite number of steps. Note that for random U, V, the probability of the commutator vanishing is zero. We also point out that Eqs. (1), (4), and (5) hold even for non-unitary matrices. Remarkably, the protocol can thus be used to rewind, for example, a two-level system undergoing a continuous decay governed by a non-Hermitian Hamiltonian.

From Eq. (1), it can be seen that to rewind a free evolution of time T, our protocol runs for $T + O(1)$ units of time, which is asymptotically optimal [7], and where the $O(1)$ term accounts for the constant overhead introduced by the adaptive error correction. In comparison, the protocol demonstrated in [8,9] takes $3T + O(1)$ units of time for the same task, making our protocol superior not only in terms of success probability, but also in terms of running time, at the cost of requiring coherent control over the time evolution.

The above description of the protocol involves placing the target quantum system in a spatial superposition; however, we note that the alternative, but equivalent, description of the protocol provided in Supplement 1 does not require this.

B. Experiment

The rewinding protocol described in the previous section is applied to a qubit state encoded in the polarization degree of freedom of a single photon, while the path degree of freedom of the same photon is used at two points to encode a second qubit that acts as a control system, thereby enabling the application of a commutator through a controlled unitary inside a quantum switch. The photons are generated using spontaneous parametric downconversion (SPDC). The SPDC process produces pairs of photons denoted signal and idler, the former of which is sent straight to a detector and is used to herald the presence of the idler photon. Upon such a heralding event, a trigger signal is transmitted in optical fiber to a field programmable gate array (FPGA) controlling two active electro-optical (EO) switches to permit the idler photon to pass through parts of the setup multiple times. The active routing of

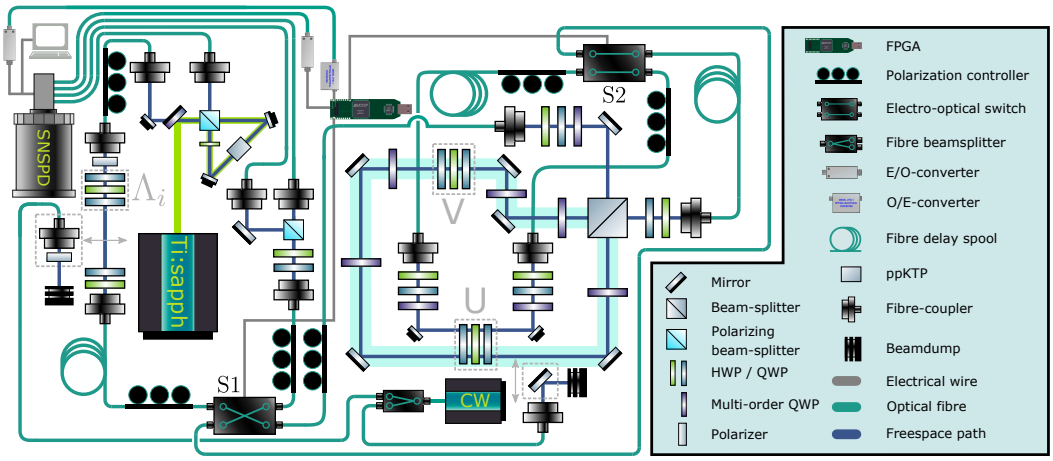


Fig. 4. Experimental setup. A pulsed Ti:sapphire laser pumps a spontaneous parametric downconversion source to generate pairs of single photons in a type-II process using a ppKTP crystal (top left). The signal photon is directed to the high efficiency superconducting nanowire single-photon detectors (SNSPDs), and a successful detection event triggers an FPGA to initiate a pulse sequence for the EO-switches (see Fig. 3). The approximately 400 ns rise time of the EO-switches is compensated for by three fiber spools, each around 100 m long, adding the needed optical delay. The target state, encoded in the idler photon, is initialized using a state-preparation stage after which it is sent to the quantum switch, realized using a free-space Sagnac interferometer (highlighted in blue). The unitaries U and V are implemented using a combination of half- and quarter-wave plates. Additional wave plates are used in conjunction with the fiber polarization controllers to compensate for unwanted polarization rotations induced by the fibers and mirrors. Two additional fiber couplers placed inside the Sagnac allow the photons to propagate through U separately. A tomography stage at the output of EO-switch S1 is used to measure the photons' polarization. The CW laser is used during the pre-measurement polarization compensation procedure.

the photons by the EO-switches is shown in Fig. 3, while a detailed schematic of the setup is displayed in Fig. 4.

The unitary Λ_i initializes the idler photon into the polarization state $|\Psi_i\rangle$ chosen from a tomographically complete set, after which an EO-switch (S1) routes the photon into the quantum switch. The unitaries V and U inside the quantum switch are implemented using two sets of three wave plates [17]. Note that there is only one physical realization of U and V, and they could thus in principle remain unknown without compromising the protocol. Depending on whether the photon exits in the backwards or forwards propagating port of the interferometer, either $[U, V]$ or $\{U, V\}$ is applied. The backpropagating port corresponding to $\{U, V\}$ [Fig. 2(c)] is disregarded in our implementation, and no adaptive error correction is applied, but photons exiting in this port could be used to increase the success probability of the protocol. Any photon leaving the interferometer in the forward propagating direction passes through a second EO-switch (S2), which traps the photon in a loop, allowing it to propagate through U a total of n times. Upon exiting the loop, the photon is directed back to S1, which sends the photon through the quantum switch a second time where $[U, V]$ is probabilistically applied once more. Finally, the photon is routed to a quantum tomography stage by S1, where its polarization is measured, post-selecting on succesful application of the commutators. These measurements are then used to reconstruct the density matrix ρ. In a successful run of the experiment, the state $U^{-n}|\Psi_i\rangle$ is recorded.

3. RESULTS

To demonstrate that the performance of the protocol is independent of the initial state $|\Psi_i\rangle$, the free evolution U, the perturbed evolution V, and the number of time steps n, a large set of combinations of these parameters was realized. More specifically, the unitary operators U and V were chosen from the set

$$U_p = e^{-i\arcsin(\alpha)\sigma_z}, \quad \alpha = \frac{p}{10}, \quad p = 1, \ldots, 10,$$

$$V_q = \cos(\theta)\sigma_z + \sin(\theta)\sigma_y, \quad \theta = \frac{q\pi}{11}, \quad q = 1, \ldots, 10.$$

Depending on the choice of p and q, the degree to which the operators U, V commute changes. As a measure of the commutativity, we define

$$N_c = 1 - \left(\frac{||[U, V]||_2}{2}\right)^2, \tag{6}$$

which is normalized to be one when the unitaries are fully commuting and is equal to the probability of applying the commutator in a single step. For our implementation, we select 50 pairs of U, V for which $N_c \leq 0.9$. Choosing a finite set of unitaries generated by fixed Hamiltonians, rather than sampling them randomly, allows us to compare our results to those of a classical experimenter given access to the same resources, but who cannot implement quantum superpositions of time evolutions. The optimal strategy for such a classical experimenter is discussed in Supplement 1. While polarization rotations implemented by wave plates alone are in general not invariant under counterpropagation, the specific set above is since it does not contain any σ_x terms. As was recently shown in [18], the restriction on the set of implementable unitaries in a Sagnac quantum switch can be lifted if one includes time-reversal symmetry breaking elements, making the implementation equivalent to previous Mach–Zehnder based quantum switches.

To benchmark the fidelity of the protocol, we ran it on the four input states $\{|H\rangle, |+\rangle, |-\rangle, |R\rangle\}$, corresponding to horizontally, diagonally, anti-diagonally, and right-handed circularly polarized light, respectively. This was independently repeated

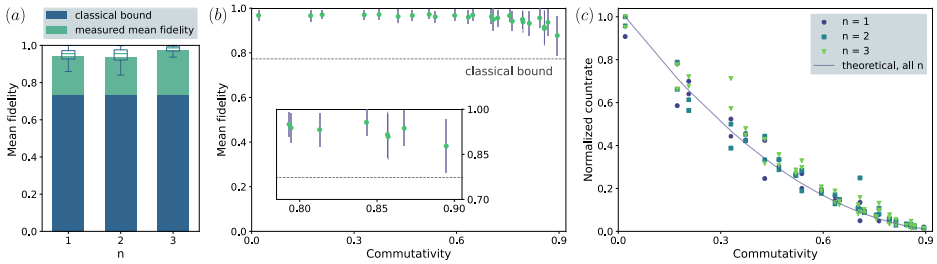


Fig. 5. Experimental results. (a) Experimental state fidelities. Each bar shows the measured state fidelity averaged over all 50 pairs of U, V, the four different input states, and three independent experimental runs, giving a total of 600 different reconstructed density matrices for each value of n. The combined measurement time for all n was approximately three weeks. The exact fidelities are $\mathcal{F}_1 = (0.94234 \pm 0.00023)$, $\mathcal{F}_2 = (0.93803 \pm 0.00041)$, $\mathcal{F}_3 = (0.97336 \pm 0.00043)$, with an average of $\mathcal{F}_m = (0.95129 \pm 0.00021)$. The superimposed box plot indicates the median and spread of the fidelities for each n. The higher fidelity for $n = 3$ can be attributed to higher polarization contrast in the setup (see Supplement 1). The dark blue bars show the highest theoretical fidelity for an experimenter unable to implement superpositions of time evolutions. (b) Fidelities to $U^{-n}|\Psi_i\rangle$ as a function of the commutativity (N_c). For a given N_c, the plotted fidelity is averaged over all runs and input states, with a total of 72 samples per point as several pairs of unitaries commute to the same degree. The error bars show the standard deviations of the fidelities, and not the uncertainty in the estimated mean fidelity, which is too small to be visible. At high commutativity, the experiment becomes more sensitive to several noise sources, such as detector dark counts, background photons, and the leakage of the interferometer due to finite visibility, whereas in the regime of high commutativity, the fidelity is limited by constant effects such as finite polarization contrast through the setup. (c) Commutativity (N_c) versus the normalized total event count rate for all implementations of V and U. Count rates are normalized to the maximal event rate separately for each n (to account for additional losses at higher n). The rates are averaged over all four input states and all three runs. $n = 1$ (circles), $n = 2$ (rectangles), $n = 3$ (triangles). Error bars (Poissonian standard deviation) are too small to be visible. The theoretically ideal behavior, depicted by the solid line, is given by N_c^2 and has a quadratic behavior due to the commutator being applied twice. The biggest deviation from the overall good agreement to the theory appears in the central region of the curve, where the interferometer has a higher sensitivity to noise.

three times for all 50 choices of time evolutions, and for three different sizes of time steps ($n = 1, 2, 3$), yielding a total of 1800 experimental runs with a combined measurement time of more than 500 h. In each experimental run, full quantum state tomography was performed on the output states ρ, and the fidelity $\langle\Psi_i|U^n\rho U^{-n}|\Psi_i\rangle$ was calculated. The density matrices of the output states were reconstructed using a maximum likelihood fit [19], and a background contribution originating from the detector dark counts was accounted for using a Monte Carlo simulation, which is how the uncertainties in the fidelities were calculated (see Supplement 1). The average fidelities for $n = 1, 2, 3$ were $\mathcal{F}_1 = (0.94234 \pm 0.00023)$, $\mathcal{F}_2 = (0.93803 \pm 0.00041)$, $\mathcal{F}_3 = (0.97336 \pm 0.00043)$. These fidelities, along with the classical bound, are shown in Fig. 5(a); it can be seen that the quantum protocol clearly outperforms the classical strategy, achieving a high fidelity independent of the length of the time evolution.

In our implementation, the fidelity of the final state is not fully independent of the choice of U, V. This is due to the fact that for pairs of unitaries that almost commute, photons are most likely to exit in the anticommutator port of the interferometer, which in turn makes the protocol more sensitive to experimental imperfections such as finite interferometric visibility and detector dark counts. In Fig. 5(b), the relationship between the degree of commutativity N_c and fidelity is illustrated. The mean fidelity stays at high levels over a broad range of N_c; only when the degree of commutativity approaches 0.9 can a small drop in the fidelity be seen.

Since it is expected that the event rate will drop with increasing values of N_c, we verify that out setup produces the correct scaling by comparing N_c to our normalized detected photon rate, separately for each n. The comparison is visualized in Fig. 5(c) where good agreement between relative rate and degree of commutativity can be seen. We attribute the undesired variance in rate to imperfect polarization compensation inside the Sagnac interferometer, as well as phase shifts originating from slight interferometer misalignment. The largest variance is seen in the neighbourhood around $N_c = 0.5$, where the sensitivity to phase noise is highest, due to the sinusoidal relationship between phase and output intensity in an interferometer.

4. DISCUSSION

In this work, we have demonstrated a universal time-rewinding protocol for two-level quantum systems. Unlike previously proposed protocols, ours can reach an arbitrarily high probability of success and is asymptotically optimal in the time required to perform the rewinding, answering the question of whether or not such processes are permitted by the laws of quantum mechanics. Remarkably, the experimenter performing the rewinding does not need any knowledge about the target quantum system, its internal dynamics, or even the specifics of the perturbed evolution. The optimality of the protocol is demonstrated in our implementation, where the total elapsed time (equivalent to the number of applications of U) grows linearly with the length of time to be rewound, with an optimal proportionality constant of one. We find that the experimental quantum protocol significantly outperforms the optimal classical strategy in terms of the resulting state fidelity.

We emphasize that our results are in principle not restricted to photonic quantum systems, since the concepts used do not make any assumptions about the physical system the protocol is applied to. We note that, while experiments using cold atom interferometers have demonstrated the necessary building blocks for the protocol [20], implementations utilizing massive particles would still likely prove challenging. In contrast, our photonic implementation offers a particularly simple and robust approach that utilizes a mature technological platform, in particular for implementing the commutator of the time evolutions through a quantum switch. Given the recent progress in integrated quantum photonics [21,22], we envision that fully monolithic architectures

capable of higher fidelity operations will facilitate demonstrations of the active error correction [Fig. 2(e)] in the near future. Additional follow-up investigations could include non-optical implementations of the protocol as well as extensions to higher dimensions, as described in [7].

Funding. European Commission (AppQInfo (no. 956071), EPIQUS (no. 899368)); Austrian Science Fund (BeyondC (F7113), Reseach Group 5 (FG5)); Air Force Office of Scientific Research (PhoQuGraph (FA8655-20-1-7030), QTRUST (FA9550-21-1-0355)); John Templeton Foundation (QISS (ID 61466)); Austrian Federal Ministry for Digital and Economic Affairs; National Foundation for Research, Technology and Development; Christian Doppler Forschungsgesellschaft.

Acknowledgment. P.S. and T.S. thank Robert Peterson and Lee Rozema for useful discussions. D.T. is a recipient of a DOC Fellowship of the Austrian Academy of Sciences at the Institute of Quantum Optics and Quantum Information (IQOQI), Vienna. T.S and P.S. designed and built the experimental setup, carried out the measurements, and analyzed the experimental data. D.T. and B.D. provided theoretical support. V.S. assisted with the experimental implementation. M.N. and P.W. supervised the project. All authors contributed to writing the manuscript.

Disclosures. The authors declare no conflicts of interest.

Data availability. All data used in the manuscript are made publicly available in Ref. [23].

Supplemental document. See Supplement 1 for supporting content.

[†]These authors contributed equally to this paper.

REFERENCES

1. A. Eddington, *The Nature of the Physical World* (Cambridge University, 1928).
2. H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics* (Wiley, 1998).
3. W. Kuperman, W. S. Hodgkiss, H. C. Song, T. Akal, C. Ferla, and D. R. Jackson, "Phase conjugation in the ocean: experimental demonstration of an acoustic time-reversal mirror," J. Acoust. Soc. Am. **103**, 25–40 (1998).
4. W. R. Tompkin, M. S. Malcuit, R. W. Boyd, and R. Y. Chiao, "Time reversal of Berry's phase by optical phase conjugation," J. Opt. Soc. Am. B **7**, 230–233 (1990).
5. A. L. Gaeta and R. W. Boyd, "Quantum noise in phase conjugation," Phys. Rev. Lett. **60**, 2618 (1988).
6. M. Navascues, "Resetting uncontrolled quantum systems," Phys. Rev. X **8**, 031008 (2018).
7. D. Trillo, B. Dive, and M. Navascues, "Translating uncontrolled systems in time," Quantum **4**, 374 (2020).
8. M. Gong, F. Xu, Z.-D. Li, *et al.*, "Verification of a resetting protocol for an uncontrolled superconducting qubit," NPJ Quantum Inf. **6**, 99 (2020).
9. Z.-D. Li, X.-F. Yin, Z. Wang, L.-Z. Liu, R. Zhang, Y.-Z. Zhang, X. Jiang, J. Zhang, L. Li, N.-L. Liu, X.-B. Zhu, F. Xu, Y.-A. Chen, and J.-W. Pan, "Photonic realization of quantum resetting," Optica **7**, 766–770 (2020).
10. M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, "Reversing unknown quantum transformations: universal quantum circuit for inverting general unitary operations," Phys. Rev. Lett. **123**, 210502 (2019).
11. Q. Feng, T. Feng, Y. Tian, M. Luo, and X. Zhou, "Experimentally undoing an unknown single-qubit unitary," Phy. Rev. A **102**, 012602 (2020).
12. I. S. B. Sardharwalla, T. S. Cubitt, A. W. Harrow, and N. Linden, "Universal refocusing of systematic quantum noise," arXiv, arXiv:1602.07963 (2016).
13. Y. Aharonov, J. Anandan, S. Popescu, and L. Vaidman, "Superpositions of time evolutions of a quantum system and a quantum time-translation machine," Phys. Rev. Lett. **64**, 2965–2968 (1990).
14. G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, "Quantum computations without definite causal structure," Phys. Rev. A **88**, 022318 (2013).
15. L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. A. Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, "Experimental superposition of orders of quantum gates," Nat. Commun. **6**, 7913 (2015).
16. D. Trillo, B. Dive, and M. Navascués, "A universal quantum rewinding protocol with an arbitrarily high probability of success," arXiv, arXiv:2205.01131 (2022).
17. R. Simon and N. Mukunda, "Minimal three-component SU(2) gadget for polarization optics," Phys. Lett. A **143**, 165–169 (1990).
18. T. Strömberg, P. Schiansky, R. W. Peterson, M. T. Quintino, and P. Walther, "Demonstration of a quantum switch in a Sagnac configuration," arXiv, arXiv:2211.12540 (2022).
19. J. B. Altepeter, E. R. Jeffrey, and P. G. Kwiat, "Photonic state tomography," Adv. At. Mol. Opt. Phys. **52**, 105–159 (2005).
20. A. D. Cronin, J. Schmiedmayer, and D. E. Pritchard, "Optics and interferometry with atoms and molecules," Rev. Mod. Phys. **81**, 1051 (2009).
21. J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, "Integrated photonic quantum technologies," Nat. Photonics **14**, 273–284 (2020).
22. F. Flamini, N. Spagnolo, and F. Sciarrino, "Photonic quantum information processing: a review," Rep. Prog. Phys. **82**, 016001 (2018).
23. P. Schiansky, T. Strömberg, D. Trillo, V. Saggio, B. Dive, M. Navascués, and P. Walther, "Demonstration of universal time-reversal for qubit processes," Zenodo (2022), https://zenodo.org/record/7354848#.Y9AUMnbMKM8.

Supplemental Document

Demonstration of universal time-reversal for qubit processes: supplement

P. Schiansky,[1,3,†] T. Strömberg,[1,†] D. Trillo,[2] V. Saggio,[1] B. Dive,[2] M. Navascués,[2] and P. Walther[1,4]

[1] *University of Vienna, Faculty of Physics & Research Network Quantum & Aspects of Space Time (TURIS), Boltzmanngasse 5, 1090 Vienna, Austria*
[2] *Institute for Quantum Optics and Quantum Information, Boltzmanngasse 3, 1090 Vienna, Austria*
[3] *e-mail: peter.schiansky@univie.ac.at*
[4] *e-mail: philip.walther@univie.ac.at*
[†] *These authors contributed equally to this paper.*

Supplement DOI: https://doi.org/10.6084/m9.figshare.21757478

Parent Article DOI: https://doi.org/10.1364/OPTICA.469109

Demonstration of universal time-reversal for qubit processes: supplemental document

1. CONNECTION WITH THE PROBE FRAMEWORK

Previous proposals for universal time translation [1], [2] are framed in a *scattering scenario*, where the target system, which sits still in a so-called scattering region, is made to sequentially interact with a number of quantum *probes*. These probes are prepared in a controlled lab and then released to the scattering region, where they interact with the target in an uncharacterized but repeatable way. As it turns out, the rewinding protocol introduced in the main text can also be realized as a scattering experiment. In the following we explain how this can be achieved.

Our goal is to implement a quantum SWITCH by means of scattered probes, which the experimenter has full control over. The probe system P, together with an ancillary qubit A, are initially prepared in the state:

$$|\omega\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle_A |\Phi\rangle_P + |1\rangle_A |\Psi\rangle_P \right). \tag{S1}$$

Here $|\Phi\rangle$ is a state that remains in the lab and does not interact with the target system, while $|\Psi\rangle$ represents a probe state that allows the probe to enter the scattering region and interact with the target.

We assume that the target system T, in the absence of probes, evolves via the (unknown) free Hamiltonian H_0. Similarly, when the probe remains in the lab, its evolution is governed by the (known) free Hamiltonian H_P. Finally, the interaction between the target and a probe in the scattering region is described by the (unknown) Hamiltonian H_I.

Taking the initial state of the target to be $|\psi\rangle$, we let the joint target-probe state $|\omega\rangle_{AP} |\psi\rangle_T$ evolve for time ΔT:

$$\frac{1}{\sqrt{2}} \left(|0\rangle_A U |\psi\rangle_T W |\Phi\rangle_P + |1\rangle_A e^{-iH_I \Delta T} |\psi\rangle_T |\Psi\rangle_P \right), \tag{S2}$$

where $U = e^{-iH_0 \Delta T}$ and $W = e^{-iH_P \Delta T}$. Subsequently, we apply the following (probabilistic) operation on the probe system:

$$|0\rangle \langle 0|_A \otimes \mathbb{1}_P + |1\rangle \langle 1|_A \otimes W |\Phi\rangle \langle \Xi|_P, \tag{S3}$$

where $|\Xi\rangle$ is any state of the probe with support in the lab. Defining:

$$V = (\langle \Xi|_P \otimes \mathbb{1}_T) e^{-iH_I \Delta T} (|\Psi\rangle_P \otimes \mathbb{1}_T), \tag{S4}$$

we can now write the joint state as:

$$\frac{1}{\sqrt{2}} \left(|0\rangle_A U |\psi\rangle_T + |1\rangle_A V |\psi\rangle_T \right) W |\Phi\rangle_P, \tag{S5}$$

Taking advantage of the fact that the probe is now in the state $|\Phi\rangle$, i.e. within the lab, we apply the following operation:

$$|0\rangle \langle 0|_A \otimes (|\Psi\rangle \langle \Phi| W^{-1})_P + |1\rangle \langle 1|_A (\otimes W^{-1})_P. \tag{S6}$$

yielding the state:

$$\frac{1}{\sqrt{2}} \left(|0\rangle_A U |\psi\rangle_T |\Psi\rangle_P + |1\rangle_A V |\psi\rangle_T |\Phi\rangle_P \right). \tag{S7}$$

Letting the probe and target systems evolve for another ΔT units of time and applying (S3) once more results in the state:

$$\frac{1}{\sqrt{2}}\left(|0\rangle_A VU|\psi\rangle_T + |1\rangle_A UV|\psi\rangle_T\right) W|\Phi\rangle_P . \tag{S8}$$

Since the state of the probe factors out, we ignore it from now on. Finally, we measure the ancillary qubit in the basis $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. Depending on the measurement result $\pm$, the final state of the target system will be

$$\frac{1}{2}\left(VU \pm UV\right)|\psi\rangle . \tag{S9}$$

This is an implementation of the quantum SWITCH gate.

2. CLASSICAL STRATEGIES

In this section we compute the fidelity of a classical rewinding protocol. We consider an experimenter given access to the same resources as a quantum one. Specifically they can choose to either let the system evolve freely for some length of time, or evolve the system using the perturbed time evolution V. The strategy should be understood as universal, meaning that the experimenter cannot tailor their interventions to the specific time evolution to be rewound. In contrast to the quantum experimenter, the classical one can only implement the free and perturbed evolutions sequentially, not in a coherent superposition. Thus, the most general classical strategy will have the following form:

$$C(H_0, V, n, t) := e^{-iH_0 t_n} V \cdots e^{-iH_0 t_1} V e^{-iH_0 t_0},$$

where $t_j \geq 0$ for all j, and

$$n + \sum_{i=0}^{n} t_i \leq (4+n)\Delta T.$$

This last condition ensures that the classical protocol does not last longer than the one we have used in the main text. Since we are using waveplates as gates, we have to consider a discretized version of time where each gate consumes ΔT units of time, and thus the most basic rewinding protocol, consisting of implementing $[U, V]U^n[U, V]$, lasts indeed $(4+n)\Delta T$ time units.

Our figure of merit for each n is the average over the fidelities between the final state and the result of rewinding the original state $|\psi\rangle \in \mathcal{S}$ by an amount $\Delta T n$. The classical expression is thus

$$F = \frac{1}{|\mathcal{P}||\mathcal{S}|} \sum_{(U,V)\in\mathcal{P}, \psi\in\mathcal{S}} |\langle\psi| U^n C(H_0, V, n, t)|\psi\rangle|^2, \tag{S10}$$

where $\mathcal{P}, \mathcal{S}$ are, respectively, the set of pairs of operators (U, V) and states ψ considered in the experiment.

We numerically maximize this expression in Mathematica, obtaining $F_c \approx 0.733713$ for $n = 1, 2, 3$. The optimal classical strategy for these choices of states and gates is to let the system evolve unperturbed for $5.91507 - n$ units of time. This particular result is a coincidence, since for other choices the optimal is a non-trivial strategy.

The numerical optimization over C implies that the classical experimenter posesses knowledge about the set of unitaries. An experimenter restricted to being ignorant about these sets - a constraint we impose on the quantum experimenter - does not have access to this optimal strategy.

3. EXPERIMENTAL DETAILS

Photon pairs centered at approximately 1546 nm are produced in a type-II spontaneous parametric down conversion (SPDC) source based on a periodically poled $KTiOPO_4$ crystal in a Sagnac configuration [3]. The source is pumped by a mode-locked Ti:sapphire laser (Coherent Mira HP) emitting 2 ps long pulses at 773.1 nm with a repetition rate of 76 MHz.

Two electro-optical (EO) switches (Agiltron NanoSpeed) are used throughout the experiment to route the photons in real time, enabling them to pass through the same part of the setup multiple times. Upon the detection of a signal photon, the electrical signal created by the detector is split off, with one copy being amplified to TTL levels using a fast comparator, whereupon it is fed into a waveform preserving electro-optical converter outputting an optical pulse at 1310 nm. The optical signal is sent back through the fiber link to the experimental setup, where it gets

re-converted to an electrical signal using an opto-electrical converter, and is then received by the FPGA controlling the EO-switches.

The roughly 400 ns rise time of the switches necessitates the use of long fiber delays, which lower the duty cycle of the experiment. These are 519 ns/106 m between the state preparation Λ_i and S_1, 533 ns/109 m between the output of the quantum SWITCH and S_2, and 760 ns/155 m inside of the U-loop (Fig. 3(b) in the main text). The 1.5 dB attenuation per pass through the EO-switches is the main contributor to the overall experimental loss, as the eight passes for $n = 3$ add up to 12 dB. The non-negligible leakage through the switches of around -13 dB also contributes to some experimental noise. Additional short fiber delays are used to offset the experimental signal in time to ensure that detection events originating from unused photon pairs are not separated from the real signal by an integer multiple of the pump pulse separation of 13.2 ns.

The quantum SWITCH is implemented using a bulk Sagnac interferometer to enable long term phase stability. Additionally, the common path geometry ensures that the polarization unitaries are sampled on the same physical spots on the waveplates for both values of the control qubit. The visibility of the interferometer is measured to be in excess of 0.99 for all the four input polarization states used in the experiment. Polarization-dependent phase shifts from the mirrors inside the interferometer are corrected using multi-order QWPs.

Superconducting nanowire single-photon detectors from Photon Spot, housed in a 1 K cryostat, are used for detection. The typical measured detection efficiencies are around 93 %. An approximately 100 m long optical fiber link separates the detectors from the experiment. Successful detection events are recorded by a time-tagger with 15.625 ps timing resolution.

While our SPDC source is able to generate single photons at a rate in excess of 1.5 MHz, many of which cannot be used since a single run of the experiment takes between 2.5 and 4.5 µs. Therefore, the FPGA discards all detection events from heralding (signal) photons when a run of the experiment is still in progress. We therefore attenuate the laser pump power until the point where the rate of successful trigger events by the FPGA begins to fall. This also lets us bias the heralding detector at a greater voltage, leading to a higher heralding efficiency.

The Sagnac interferometer constituting the quantum SWITCH is housed in and isolated by three different layers of thinsulate, acrylic and neoprene. This is done to decrease airflow and temperature fluctuations. Gold coated mirrors are used throughout the setup as they exhibit low polarization-dependent loss at our working wavelength (0.034, 0.035 for S, P respectively). Their relatively poor reflectivity of 0.96 adds around 3.5 dB to the total loss. Similarly, a beam splitter with low polarisation dependent loss and splitting ratio is used for the Sagnac. Multi-order quarter-wave plates are used inside the interferometer to compensate the unwanted polarization-dependent phase shifts caused by the mirrors, with typical polarization contrasts in excess of 40 dB. After acquiring data for the cases $n = 1, 2$ we were able to exchange the CW-laser used for this compensation to a model with broader wavelength tuning-range to more closely match our single-photon central wavelength. This lead to superior polarization compensation performance, which in turn explains the increased fidelity $\mathcal{F}_3$ compared to $\mathcal{F}_{1,2}$. While the polarization rotations induced by most components in the setup are not strongly wavelength dependent, the EO-switches are an exception to this, and therefore benefit from a CW-wavelength that more closely matches that of the idler photons. The polarization unitaries U_p, V_q, Λ_i are implemented with three sets of three waveplates in a QWP-HWP-QWP-configuration, mounted in motorized piezo-electric rotation mounts. While polarization transformations implemented purely with linear retarders will not in general be the same for both propagation directions, our restricted gate set consisting only of linear combinations of σ_y and σ_z is invariant under change of propagation direction. Since the applications of U outside the quantum SWITCH involve the photons hitting a different spot on the wave-plates, we verify their uniformity by performing quantum process tomography on six randomly generated unitaries. The wave-plates are sampled by four different beams offset horizontally by 2 mm each, and we find that the resulting gate fidelities do not differ by more than 0.1 %. To verify that the free time-evolution unitaries U_p are faithfully implemented we also perform quantum tomography on them inside the setup and obtain an average fidelity of $\mathcal{F}_U = (0.9928 \pm 0.00035)$, averaged over all values of p and the states $|H\rangle$ and $|+\rangle$.

The applicability of the protocol to mixed states is verified by reconstructing density matrices from convex combinations of the acquired data for the pure states $|+\rangle$ and $|-\rangle$. They are compared to expected outcomes for input states of the form $\rho = \alpha\,|+\rangle\langle+| + (1-\alpha)\,|-\rangle\langle-|$, for 23 values of $\alpha \in (0, 1)$, all 50 combinations of (p, q) and all 3 experimental runs. In Fig. S1, the fidelities to

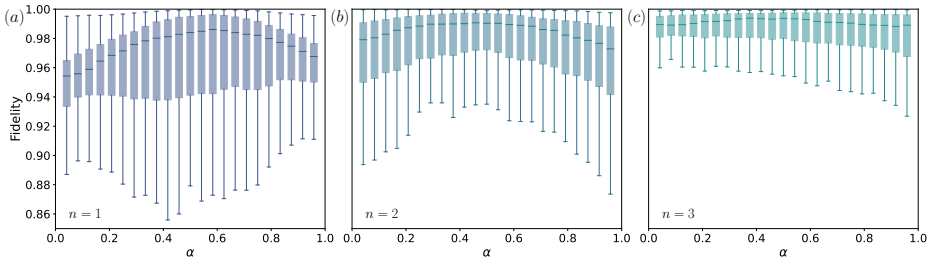


Fig. S1. Fidelities for mixed states. Density matrices are reconstructed out of convex combinations of data measured for input states $|+\rangle$ and $|-\rangle$. Fidelities to the expected states $U^{-n}\rho U^{-n^\dagger}$, with $\rho = \alpha\,|+\rangle\langle+| + (1-\alpha)\,|-\rangle\langle-|$, are calculated for varying mixing parameter α, all combinations of (U, V), all 3 experimental runs, and all n. Each box plot shows the median and spread of the state fidelities. **(a)** $n = 1$, **(b)** $n = 2$, **(c)** $n = 3$.

these expected states are plotted for each value of n.

4. SIGNAL PROCESSING

To analyse the data we generate coincidence histograms between the heralding detector and the two detectors connected to the tomography stage. As previously stated, due to the long fiber delays the idler photon takes several microseconds to traverse the entire setup. Heralding photons detected within this time window will be ignored by the FPGA, but will still be recorded by the time-tagger. In order to filter out these unused heralding photons in the coincidence analysis, the FPGA outputs a trigger signal whenever it initiates a new pulse sequence. This signal is transmitted back to the time-tagger by the same electro-optic conversion procedure as in the Methods. Only trigger events for which a corresponding signal was received from the FPGA are used in the analysis. Conditioning the photon detection events on the FPGA trigger signal significantly reduces the background noise, as illustrated in Fig. S2.

Due to the presence of active switches in the setup, in any given run of the experiment there exists multiple possible paths that a photon could have taken from the source to the tomography stage. For example, during the state depicted in Fig. 3(c) in the main text photons can travel directly from the source to the tomography stage. While most such events can be filtered out by virtue of the fact that the difference in arrival time between the signal and idler photons will not match that of the real signal, there are also higher order contributions consisting of signal and idler photons emitted at different times. Since the photons propagating straight to the tomography stage avoid most of the experimental loss, and are not attenuated by the success probability, the rate of these events becomes comparable to the signal even though the intrinsic rate of double-pair emission from the source is significantly lower than the single-pair emission rate. To offset this large noise contribution we add a small fiber delay between S1 and S2. The result is that the signal sits between the major noise peaks, as shown in Fig. S3.

While the SNSPDs have a very low dark count rate, ranging from about 30 to 300 Hz, the high rate of heralding photons nevertheless leads to a small number of accidental coincidence events that form a uniform background in the coincidence landscape. For pairs of implemented unitaries that nearly commute, the resulting low rate of detected signal photons makes the background of accidental coincidences non-negligible. To estimate the impact of this noise in our signal we sample the background in several regions of the coincidence histograms that don't contain any signal. These regions, separated by 13.2 ns, are indicated in Fig. S3. The mean value of the noise is used as input to a Monte Carlo simulation, from which the mean and standard deviation of the fidelity is obtained. At every step of the simulation a density matrix is constructed using a maximum likelihood fit. The Monte Carlo simulation is allowed to run until the 0.95 confidence interval on the mean fidelity reaches a value below $2 \cdot 10^{-3}$. In order to further increase the signal to noise, a narrow coincidence window of $0.3 - 0.7$ ns is used.

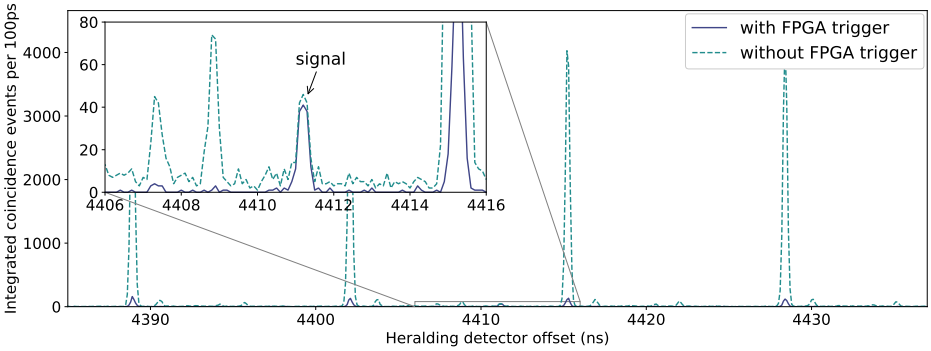


Fig. S2. Influence of FPGA trigger events on noise background. A comparison between two coincidence histograms with 100 ps wide bins taken from the first run of $(p, q, i, n) = (6, 2, H, 1)$. Background noise originating from higher-order emission events in the SPDC-source as well as detector dark counts can be greatly reduced by filtering out trigger photons that were ignored by FPGA while it was already executing a measurement sequence (blue, solid line). An unprocessed histogram is shown as dashed line. The inset shows a magnified region centered around the signal peak. The suppression of the noise peaks originating from unrelated photon pairs is greater than the ratio of unused to used trigger events. This is because the dead time of the detectors (on the order of 100 ns) acts as an additional filter on the heralding photons in the region around our signal, and as a consequence the majority of the events contributing to these peaks in the unfiltered signal comes from heralding photons that were not triggered on.

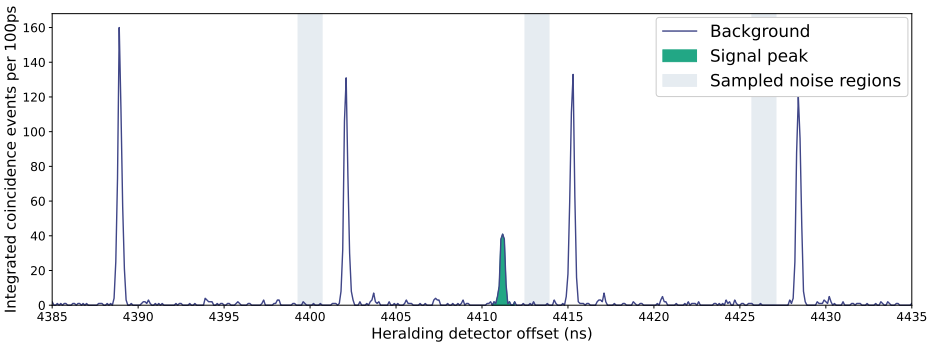


Fig. S3. Background noise sampling. An example of a coincidence histogram with 100 ps wide bins taken from the first run of $(p, q, i, n) = (6, 2, H, 1)$. It shows the number of coincidence events between the H-port of the tomography stage and the heralding detector, integrated over the measurement time, as a function of a time-offset in the heralding detector. The signal in our experiment is the dark green peak in the center of the graph. Slightly offset from the signal peak one would in an ideal experiment expect zero coincidence events due to the strong time correlation between the signal and idler photons, however due to a small but non-negligible detector dark count rate (on the order of 100 Hz) some coincidence events nevertheless occur. These form a uniform background, which we sample and include as an input to the Monte Carlo simulation that estimates the reconstructed state fidelities. The small side peaks offset from the signal by 13.2 ns (which is the reciprocal of the 76 MHz laser repetition rate) are caused by signal photons emitted before or after the pair the FPGAs triggered on. The remaining peaks are caused by coincidences from uncorrelated photon pairs, and this signal is strongly suppressed by conditioning the coincidence counting on the FPGA output signal.

5. INDIVIDUAL FIDELITIES FOR ALL INPUT STATES, TIMESTEPS AND PAIRS OF U AND V

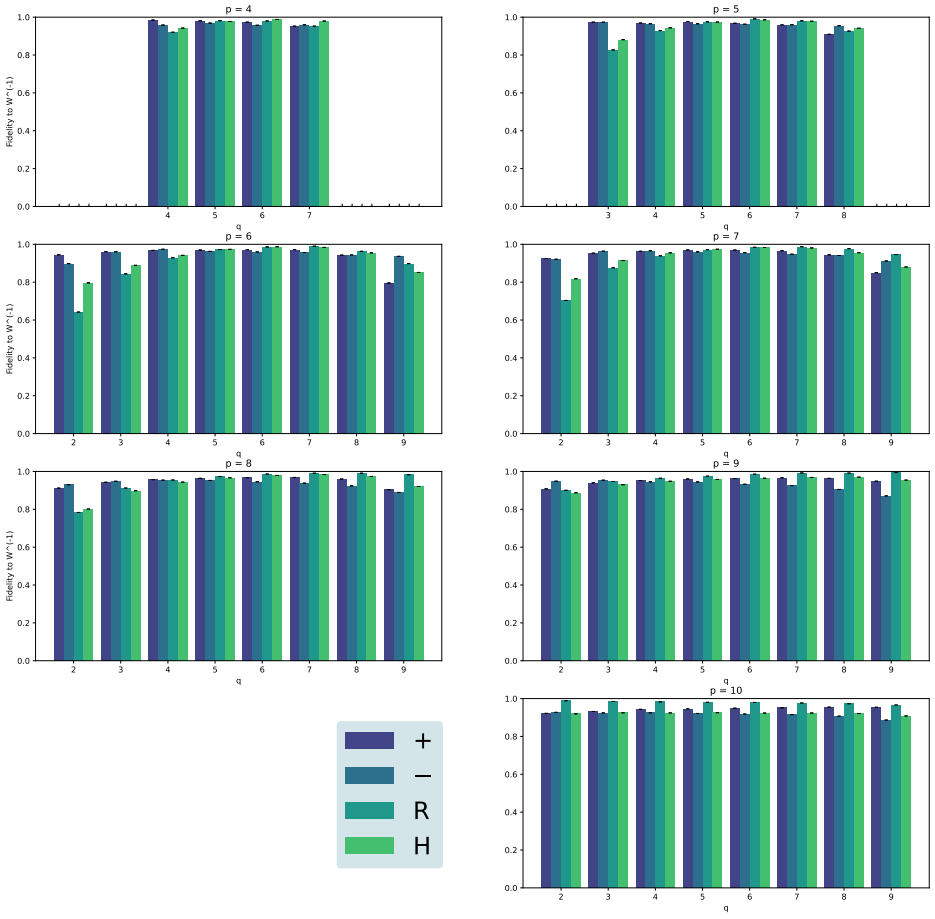


Fig. S4. Fidelities to $U^{-1}|\Psi_i\rangle$. For all 50 pairs of (U,V) with $N_c \leq 0.9$ and 4 input states each: $+$(violet), $-$(prussian), R(teal), H(green). Averaged over three measurement sets. For each measurement set, the standard deviation of the fidelities calculated by the Monte Carlo simulation is computed. The shown error bars are the root square-sum of these standard deviations.

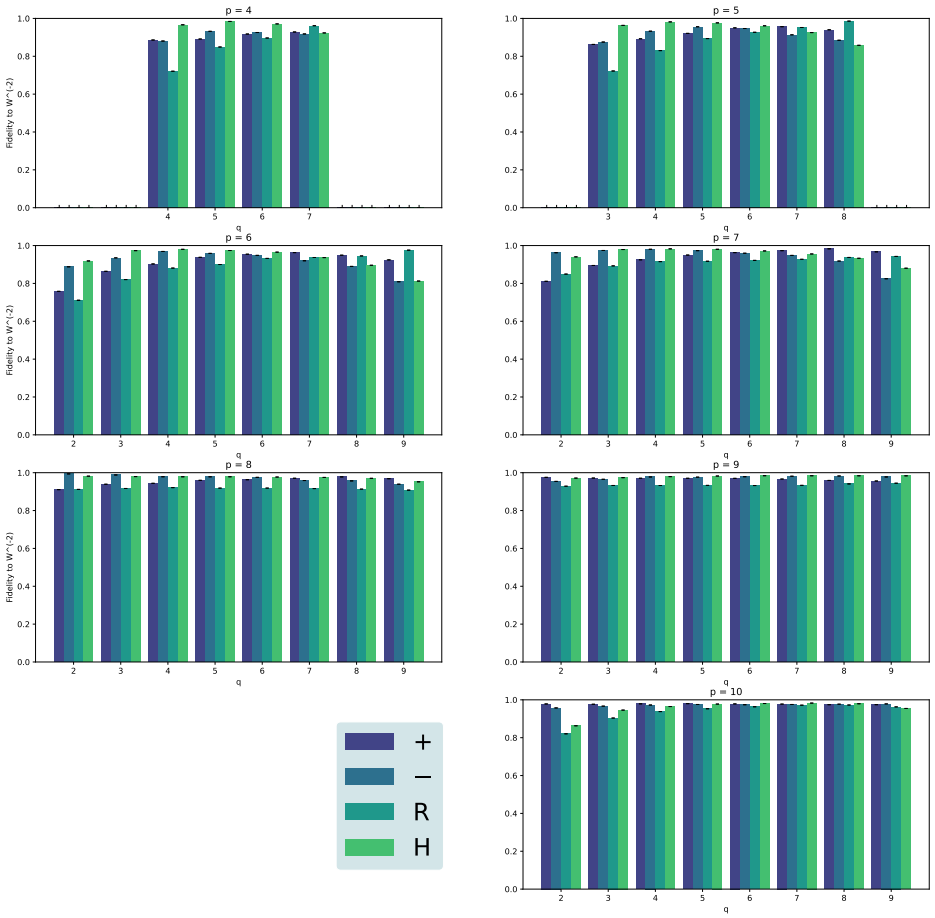


Fig. S5. Fidelities to $U^{-2}\,|\Psi_i\rangle$. For all 50 pairs of (U,V) with $N_c \leq 0.9$ and 4 input states each: $+$(violet), $-$(prussian), R(teal), H(green). Averaged over three measurement sets. For each measurement set, the standard deviation of the fidelities calculated by the Monte Carlo simulation is computed. The shown error bars are the root square-sum of these standard deviations.

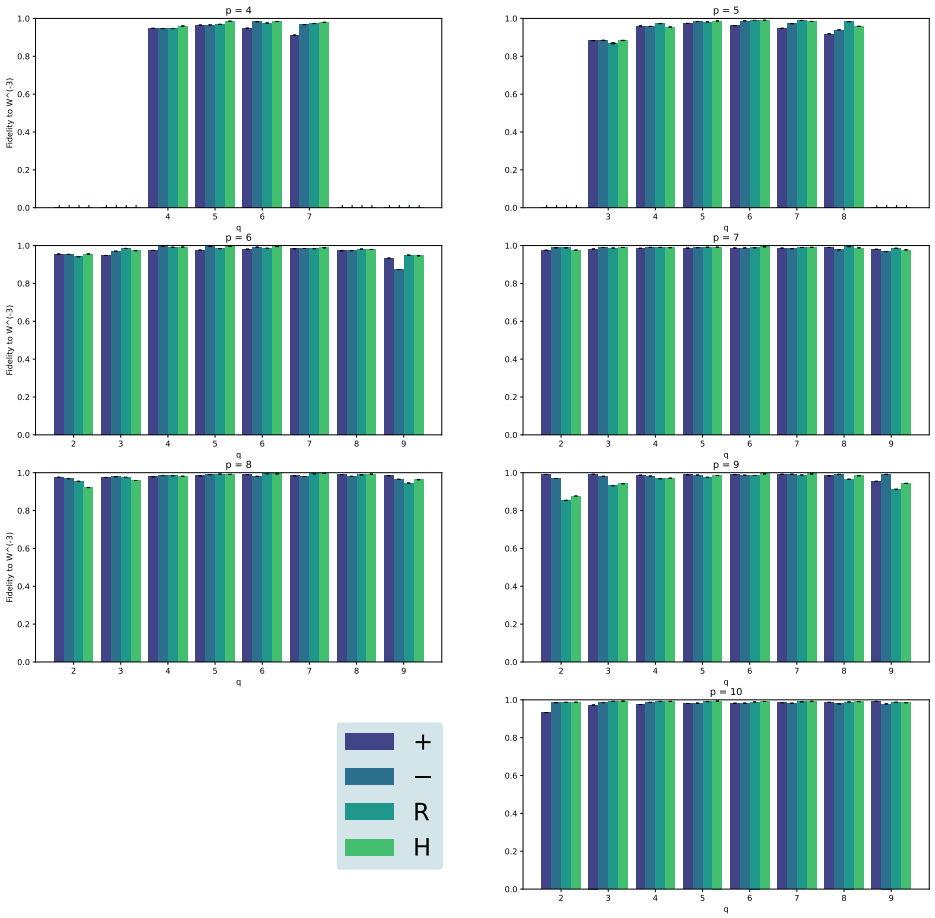


Fig. S6. Fidelities to $U^{-3}\,|\Psi_i\rangle$. For all 50 pairs of (U,V) with $N_c \leq 0.9$ and 4 input states each: $+$(violet), $-$(prussian), R(teal), H(green). Averaged over three measurement sets. For each measurement set, the standard deviation of the fidelities calculated by the Monte Carlo simulation is computed. The shown error bars are the root square-sum of these standard deviations.

REFERENCES

1. M. Navascues, "Resetting uncontrolled quantum systems," Phys. Rev. X 8, 031008 (2018) (2017).
2. D. Trillo, B. Dive, and M. Navascues, "Translating uncontrolled systems in time," Quantum 4, 374 (2020) (2019).
3. C. Greganti, P. Schiansky, I. A. Calafell, L. M. Procopio, L. A. Rozema, and P. Walther, "Tuning single-photon sources for telecom multi-photon experiments," Opt. express **26**, 3286–3302 (2018).

10 Reprint of "Experimental superposition of a quantum evolution with its time reverse"

10.1 Contribution statement

The experiment was designed by myself and Peter Schiansky, and was both assembled and carried out by myself, with input and assistance from Peter Schiansky and Iris Agresti. The collected data was analysed by myself and Marco Túlio Quintino with support from Peter Schiansky. The manuscript was primarily written by myself, Marco Túlio Quintino and Iris Agresti, in close collaboration with Peter Schiansky and under the supervision of Philip Walther, with all authors providing input. The figures were made by myself, Peter Schiansky and Iris Agresti. Michael Antesberger and Lee Rozema proposed the experiment. Marco Túlio Quintino derived the theoretical bounds on the success probabilities.

Experimental superposition of a quantum evolution with its time reverse

Teodor Strömberg,[1,*] Peter Schiansky,[1] Marco Túlio Quintino,[2,3,4] Michael Antesberger,[1] Lee Rozema,[2] Iris Agresti,[2] Časlav Brukner,[3,2] and Philip Walther[2,†]

[1]*University of Vienna, Faculty of Physics & Vienna Doctoral School in Physics, Boltzmanngasse 5, A-1090 Vienna, Austria*
[2]*University of Vienna, Faculty of Physics & Research Network Quantum Aspects of Space Time (TURIS), Boltzmanngasse 5, 1090 Vienna, Austria*
[3]*Institute for Quantum Optics and Quantum Information, Boltzmanngasse 3, 1090 Vienna, Austria*
[4]*Sorbonne Université, CNRS, LIP6, F-75005 Paris, France*

In the macroscopic world, time is intrinsically asymmetric, flowing in a specific direction, from past to future. However, the same is not necessarily true for quantum systems, as some quantum processes produce valid quantum evolutions under time reversal. Supposing that such processes can be probed in both time directions, we can also consider quantum processes probed in a coherent superposition of forwards and backwards time directions. This yields a broader class of quantum processes than the ones considered so far in the literature, including those with indefinite causal order. In this work, we demonstrate for the first time an operation belonging to this new class: the quantum time flip. Using a photonic realisation of this operation, we apply it to a game formulated as a discrimination task between two sets of operators. This game not only serves as a witness of an indefinite time direction, but also allows for a computational advantage over strategies using a fixed time direction, and even those with an indefinite causal order.

INTRODUCTION

In recent years, the framework of quantum theory has been generalised to describe agents interacting through quantum processes with indefinite causal orders [1–3]. These processes have been realised experimentally using photonic platforms [4–8], thereby witnessing the implementation of causally non-separable series of events. Remarkably, these are not the most general processes allowed by quantum mechanics. Take, for example, the quantum SWITCH [2]: even though the causal order of the constituent events is indefinite, each operation is accessed only in a single time direction. By considering processes where the time direction of the underlying operations is indefinite, one can go beyond the framework of indefinite causality. Indeed, a quantum superposition of evolutions with opposite thermodynamic arrows of time was first proposed in [9]. By associating the "forwards" temporal direction with a positive change in the entropy generated in a thermodynamic process, and its "time-reversing" counterpart with a negative change, the corresponding superposition of processes exhibits a quantum-mechanically undefined thermodynamic arrow of time.

More generally, processes with an indefinite time direction can be studied by considering operations that exhibit a time symmetry; these operations admit a change of reference frame that yields a valid quantum evolution in which the time coordinate is inverted. Unitary channels are an example of such operations, and in particular they admit the following time-reversal symmetries: for every evolution U, both the inverse $U \mapsto U^\dagger$ and the transpose $U \mapsto U^T$ are valid time-reversal operations.

Given quantum operations that can in principle be accessed in both time directions, we can consider coherent superpositions of transformations made in the forwards and backwards time-directions. This amounts to a new kind of process, which we will refer to as being *inseparable in its time direction*, an example of which - called the *quantum time flip* - was recently introduced in Ref. [10]. This process cannot be realised within the quantum circuit model. In this work we nevertheless present a photonic implementation of the quantum time flip by exploiting *device dependent symmetries* of our experimental apparatus. A quantum state undergoing a time evolution is encoded in the polarization degree of freedom of a single photon, while a control qubit determining the time direction is encoded in its path degree of freedom. We show that polarization operations with waveplates naturally implement different time directions for forwards and backwards propagation directions through the waveplates, given the correct Stokes-parameter convention. This results in a deterministic time-reversal, in contrast to more general approaches which may involve multiple uses of the input operation in combination with probabilistic or non-exact methods [11–20]. We can furthermore realise the quantum time flip deterministically by passing the photon through the waveplates in a superpositon of the two propagation directions.

We certify the indefinite time direction by demon-

* Corresponding author: teodor.stroemberg@univie.ac.at
† Corresponding author: philip.walther@univie.ac.at

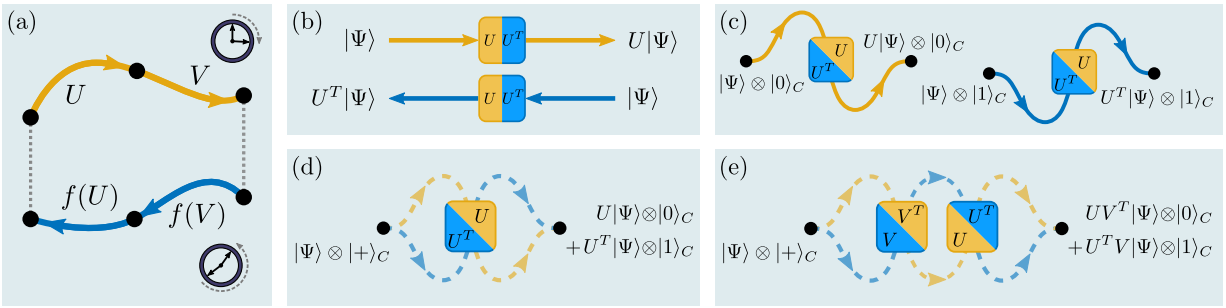


FIG. 1. **Time-reversal and the quantum time flip. (a)** The forwards (top) and backwards (bottom) directions of the same time-evolution are shown in yellow and blue, respectively. The backwards time-evolution is given by some function f of the forwards evolution, and decomposing the total time evolution into steps shows that f must be order reversing. The inverse and transpose are examples of such order reversing functions. **(b)** Quantum gates are often modelled as black boxes with an input and an output. In this work, we consider black boxes that can be accessed in two different directions, producing either the forwards or backwards time-evolution depending on which direction the box is accessed in. Here, the backwards time-evolution is taken to be the transpose. **(c)** A control degree of freedom can be introduced to control in which direction the black box is accessed. **(d)** By putting the control qubit in a coherent superposition of the two states in (c) the box is accessed in a superposition of both directions, and the input state is propagated in a superposition of time directions. This is a realisation of the quantum time flip. **(e)** The quantum time flip can be applied to more than a single gate. This figure illustrates a scenario where two gates are accessed in a superposition of orders, in which they always have the opposite time directions. As described in the main text, this use of two quantum time flips can yield a computational advantage.

strating an information-theoretic advantage of the quantum time flip in the context of a computational game. In this setting, the quantum time flip not only outperforms strategies that utilise operations with a fixed time direction, but even strategies that exploit operations with an indefinite causal order [4, 21].

QUANTUM CIRCUITS, UNITARY TRANSPOSITION, AND PROCESSES WITH INDEFINITE TIME DIRECTION

The standard quantum circuit formalism provides solid grounds for quantum computing and forms the basis for quantum complexity theory [22, 23]. However, it also imposes limitations on how we apply quantum theory. In a circuit, operations necessarily respect a definite causal order and the strict notion of input and output. The existence of time reversal processes such as unitary transposition is forbidden by the standard circuit formalism when given access to one [15, 16] or even two [10] uses of an unknown unitary. However, for practical and foundational reasons, researchers have been designing and pursuing non-exact and probabilistic schemes aimed towards this goal [11–19]. Remarkably, a very recent work shows that in the qubit case, when four uses of the input operation are available, there exists a quantum circuit to invert arbitrary unitary operations [20].

In quantum theory, reversible operations are described by unitary operators. Processes which reverse a composition of such operations may be expressed by a function f satisfying:

$$f(UV) = f(V)f(U), \quad \forall U, V, \tag{1}$$

for all unitary operators U and V (see Fig. 1). Under natural assumptions, it can be proven that, up to a unitary transformation, there are only two time reversal functions f, unitary transposition $f(U) = U^T$ and unitary inversion $f(U) = U^{-1}$ [10]. For two-dimensional systems, unitary transposition and unitary inversion are unitarily equivalent via a Pauli σ_Y operation. This follows from the identity, $U^{-1} = \sigma_Y U^T \sigma_Y$ which holds for all operators $U \in \mathcal{SU}(2)$. Hence, for qubits, universal unitary transposition is possible if and only if unitary inversion is possible. When focusing on a particular physical implementation, the general aspects of the standard quantum circuit formalism may limit our view and lead to an apparent mismatch between theory and practice. A known illustrative example is the universal coherent control of unitary operations, where an arbitrary unitary U is applied to the target system conditional on the state of a control qubit: $U \mapsto \mathbb{1} \otimes |0\rangle\langle 0|_C + U \otimes |1\rangle\langle 1|_C$. While it is not possible to design a quantum circuit to perform universal control, a simple Mach-Zehnder optical interferometer can be used for this task [24–26]. Indeed, experimental control of black box quantum gates has been demonstrated [27, 28]. Such experimental implementations exploit the knowledge of the position of the physical device per-

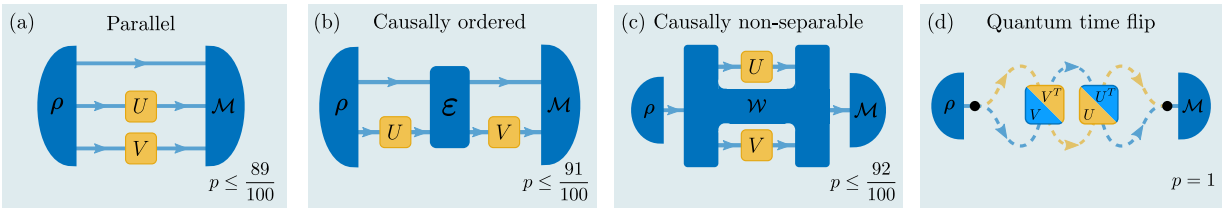


FIG. 2. **Classes of game strategies.** The figure depicts the different strategies for the game described in the main text and their corresponding maximum winning probabilities p. These maximum winning probabilities are obtained through an optimization over all possible choices of the resources shown in dark blue. The state ρ, for example, is allowed to contain any number of auxiliary degrees of freedom, and analogous statements hold for the measurement $\mathcal{M}$, channel $\mathcal{E}$ and process $\mathcal{W}$. The three strategies differ in how they are able to access the gates picked by the referee. The strategies (a)-(c) are shown here in the forwards time direction, but are also valid in the backwards time direction in which both gates are transposed. Each subsequent strategy is strictly better than the previous one, and only players who have access to a quantum time flip process can win the game with unity probability. **(a)** Parallel gate order. **(b)** Causally ordered gate sequence. **(c)** Process without a definite causal order. **(d)** Quantum time flip.

forming the gate, circumventing this apparent limitation imposed by the quantum circuit formalism.

Although time reversal processes such as unitary transposition are not possible within the standard circuit formalism when given access to one [15, 16] or even two [10] uses of an unknown unitary, in this work we implement general qubit unitary transposition, as well as the quantum time flip process, using a particular optical construction. Similar to the case of universal coherent control, we make use of knowledge about our specific experimental apparatus to realise a black box unitary that may be used in two different directions. As shown in Fig. 1.b, this box implements U in the 'forwards' direction, while in the 'backwards' direction it has the effect of the transposed operation U^T.

Moreover, in addition to "simply" reversing a quantum evolution, we also coherently superpose the forwards and backwards time evolutions, and in so doing perform an optical implementation of a process with an indefinite time direction [10], i.e. one which cannot be described as a convex mixture of processes in which each gate is accessed only in one time direction. The process that we implement optically is the quantum time flip for unitary transposition, a process which acts on unitary operations as:

$$U \mapsto U \otimes |0\rangle\langle 0|_C + U^T \otimes |1\rangle\langle 1|_C \,. \tag{2}$$

We then compose the time flip process of Eq. (2) with its flipped version, $V \mapsto V^T \otimes |0\rangle\langle 0|_C + V \otimes |1\rangle\langle 1|_C$, to obtain a process which acts on a pair of unitary operators as:

$$(U, V) \mapsto UV^T \otimes |0\rangle\langle 0|_C + U^T V \otimes |1\rangle\langle 1|_C \,. \tag{3}$$

In addition to having an indefinite time direction, the process described in Eq. (3) cannot be described by general process matrices with indefinite causality such as the quantum switch [2] or the Oreshkov-Costa-Brukner (OCB) process [3]. In the next section, we will explain how to witness this property.

GAME DESCRIPTION

We now describe a discrimination task, first introduced in Ref. [10], where the quantum time flip process will be used as a resource to increase our performance. In this game, a referee provides the player with two black box unitaries, U and V, belonging to either the set $\mathcal{M}_+$ or $\mathcal{M}_-$, which are known to respect the property:

$$\mathcal{M}_+ := \left\{ (U, V) : UV^T = +U^T V \right\} \tag{4}$$

$$\mathcal{M}_- := \left\{ (U, V) : UV^T = -U^T V \right\}. \tag{5}$$

The player is then challenged to determine which of the two sets the gates were picked from, while only being allowed to access each of the black boxes once.

As discussed in the previous section, a player able to perform the quantum time flip may implement the process:

$$(U, V) \mapsto UV^T \otimes |0\rangle\langle 0|_C + U^T V \otimes |1\rangle\langle 1|_C \,. \tag{6}$$

Consider as a strategy an initial state of the form $|\psi\rangle_T \otimes |+\rangle_C$, where $|\pm\rangle_C = \frac{|0\rangle_C \pm |1\rangle_C}{\sqrt{2}}$, $|\psi\rangle$ is an arbitrary state, and the subscripts C and T refer to the control and target qubits. Sending this state through the gate in Eq. (6) gives the state:

$$\left[\frac{UV^T + U^T V}{2}\right] |\psi\rangle_T |+\rangle_C + \left[\frac{UV^T - U^T V}{2}\right] |\psi\rangle_T |-\rangle_C. \tag{7}$$

Since the states $|\pm\rangle$ are orthogonal, a player using this strategy can always correctly determine which set was chosen by the referee. In contrast, players who do not have access to indefinite time strategies may not be able to ascertain with certainty to which set a given pair of unitaries (U, V) belongs. In order to make this claim concrete, Ref. [10] considers a particular game where the set $\mathcal{M}_+$ has 13 pairs of unitary operators respecting $UV^T = +U^TV$, and $\mathcal{M}_-$ has 8 pairs of unitary operators respecting $UV^T = -U^TV$; these two sets of unitary operators are presented in the Methods. Here, we consider an average case variation of the aforementioned game, which goes as follows: with uniform probability $p = \frac{1}{13+8}$, the referee picks a pair of unitary operators (U, V) from $\mathcal{M}_+$ or $\mathcal{M}_-$ and lets the player make a single use of each. We then consider the optimal success probability of players who have access to different kinds of resources. As indicated by Eq. (7), players who have access to the quantum time flip can always win with unity probability. The three other classes of strategies, shown in Fig. 2, only have access to a single time direction, forwards or backwards, and convex combinations of these strategies will be called separable in their time direction; a detailed mathematical characterisation of these strategies is presented in the methods. Employing the computer-assisted proof methods of Ref. [29] we obtain upper bounds on the maximal success probabilities for players restricted to particular classes of strategies. The code for this is openly available in our online repository, see Methods for details.

The first alternative strategy we consider is one in which the player is restricted to using U and V in parallel, and this results in a maximal success probability that is bounded by $\frac{88}{100} \leq p_{\text{par}} \leq \frac{89}{100}$. Next, we consider players restricted to causally ordered strategies, whose maximal success probability is found to be bounded by $\frac{90}{100} \leq p_{\text{causal}} \leq \frac{91}{100}$. Finally, players given access to process matrices with indefinite causality (also called indefinite testers [30]), but with definite time direction, have their maximal success probability bounded by $\frac{91}{100} < p_{\text{i.c.}} \leq \frac{92}{100}$. Unlike the task in [31], in which causally ordered and general non-quantum-circuit-model strategies perform equally well, this game is hence an example of a channel discrimination task with strict hierarchy between four different classes of strategies. Additionally, while the operations selected by the referee are treated as being fully characterised in the above analysis, there are no assumptions made about the measurements performed by the player, and these can remain unknown. This is therefore an example of a semi-device-independent certification of an indefinite time direction [32, 33]. This stands in contrast to witness based approaches, previously used to certify advantages in channel discrimination tasks [5], in which one needs well characterised measurement devices in order to evaluate the witness operator.

EXPERIMENT

Our photonic implementation of the game described in the previous section makes use of the quantum time-flip strategy from Eq. (7) to achieve a success probability exceeding that of any strategy only using the gates in one time direction. To coherently apply the quantum time flip, we employ polarization optics in a partially common-path interferometer, depicted in Fig. 3, with the control and target qubits being encoded in the path and polarization degrees of freedom of a single-photon, respectively. Our experiment makes use of two quantum time flips, sequentially applied to the two unitaries V and U. The resulting controlled channel is the one of Eq. (7) where the gates UV^T and U^TV act on the target (polarization) qubit and are implemented using two Simon-Mukunda polarization gadgets consisting of three waveplates each [34], for which the transpose operation is obtained by reversing the propagation direction.

Such polarization gadgets generally do not realise the transpose operation in the backwards propagation direction, but rather a related operation:

$$U_{\text{fw}} \to U_{\text{bw}} = PU_{\text{fw}}^T P^\dagger, \tag{8}$$

where P is a matrix describing the change of reference frame to the backwards direction, and the subscripts indicate the propagation direction. While it is possible to construct a gadget that implements the transpose by introducing time-reversal symmetry breaking elements (T.S, R. Peterson, and P.W., SU(2) gadgets for counterpropagating polarization optics, Manuscript under preparation), here we instead exploit the fact that the transpose is a basis-dependent operation. More concretely, by adopting the convention $(S_1, S_2, S_3) \leftrightarrow (-Z, -Y, -X)$ for our Stokes parameters [35] we find that $P = \mathbb{1}$, and the polarization gadgets transform as the transpose under counterpropagation (see Methods). Superimposing two propagation directions through a gadget therefore allows us to implement the quantum time flip, with the photon path acting as a control degree of freedom. The specific coherent superposition of time flips in Eq. (3) is achieved through the use of fiber optic circulators.

The optical circuit in Fig. 3 begins with a bulk beamsplitter that initializes the control qubit into the state $|+\rangle_C = \frac{1}{\sqrt{2}}(|0\rangle_C + |1\rangle_C)$, after which two fiber circulators guide the photons through the V gadget in two

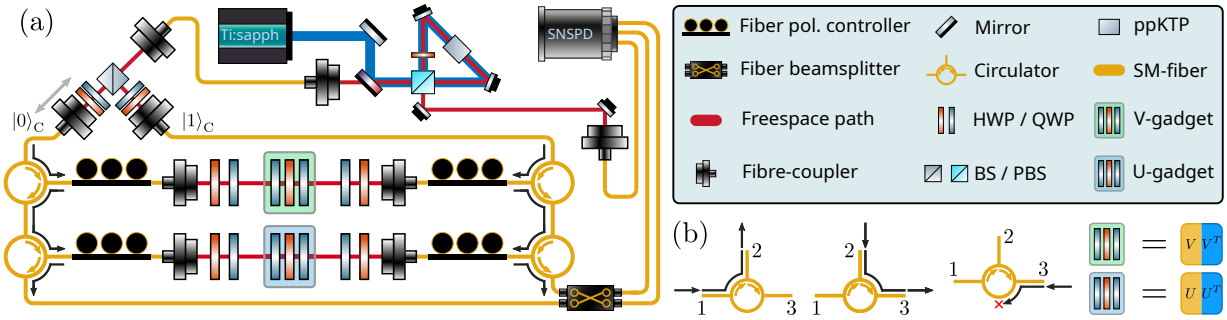


FIG. 3. **Experimental apparatus. (a)** A type-II spontaneous parametric down-conversion source generates frequency degenerate single-photon pairs at 1546 nm in a ppKTP crystal (top). The signal photon is sent to a heralding detector, while the idler photon is routed to a balanced bulk beam-splitter and coupled into single-mode fiber. A piezo-electric actuator attached to one of the fiber couplers allows for control over the interferometric phase, pairs of HWPs / QWPs are used for polarization compensation through the fibers. After the initial beam-splitter two fiber circulators guide the photon through the V-gadget. Propagating through the gadget in the 'forwards' direction implements the unitary V, while propagating 'backwards' has the effect of applying the transposed operation relative to the 'forward' direction. One of the two paths through first the gadget therefore results in the V^T being applied instead of V. Two additional circulators then route the photon through a gadget implementing U (U^T) in 'forwards' ('backwards') direction. Finally, the signal photon is sent to a fiber beam-splitter, which applies a Hadamard gate on the path degree of freedom, and correlates the two spatial output modes with the sets $\mathcal{M}_+$, $\mathcal{M}_-$. Detection is performed by superconducting nanowire single-photon detectors (SNSPDs) housed in a 1 K cryostat. Additional QWP/HWP pairs are used to compensate fiber-induced polarization rotations. **(b)** The fiber circulators route the light from port 1 $\to$ 2, from 2 $\to$ 3 and block light entering in port 3. The bidirectional boxes in Fig. 1 are realised using sets of three waveplates. Depending on the propagation direction, they implement either the unitary operation U/V or U^T/V^T.

different directions, giving the joint control-target state:

$$\frac{1}{\sqrt{2}}(V|\psi\rangle_T \otimes |0\rangle_C + V^T|\psi\rangle_T \otimes |1\rangle_C). \tag{9}$$

Entering the circulators from a different port, the photons are then directed to the U gadget, which they once again propagate through in opposite directions, transforming the joint state to:

$$\frac{1}{\sqrt{2}}(U^TV|\psi\rangle_T \otimes |0\rangle_C + UV^T|\psi\rangle_T \otimes |1\rangle_C). \tag{10}$$

At the end of the optical circuit, a fiber beam-splitter applies a Hadamard gate on the control qubit, giving the state:

$$\left[\frac{UV^T + U^TV}{2}\right]|\psi\rangle_T|+\rangle_C + \left[\frac{UV^T - U^TV}{2}\right]|\psi\rangle_T|-\rangle_C. \tag{11}$$

A projective measurement on the control (path) qubit in the computational basis then reveals whether (U, V) belong to $\mathcal{M}_+$ or $\mathcal{M}_-$.

The partially common-path structure of the interferometer has two distinct advantages: (1) photons in the two different propagation directions of the interferometer hit exactly the same spots on the waveplates and the physical symmetries of the gadget therefore ensures the faithful implementation of the time flip independently of any imperfections in the waveplates, (2) the paths traversed in both directions do not contribute any phase noise to the interferometer, thereby simplifying the phase stabilization. More specifically, only the paths connecting the two beam-splitters with the fiber circulators, as well as the fibers directly between the circulators, add phase noise to the interferometer. These fiber components, as well as the bulk beam-splitter at the interferometer input, are housed in a thermally and acoustically insulated box. The passive stabilization of these elements is sufficient to bring the phase drift down to a value of approximately 10 mrad min^{-1}. The use of a bulk-beamsplitter at the input was chosen in order to balance the losses induced by the fiber circulators through the free-space to fiber coupling, and to give control over the interferometer phase , through a piezo-electric actuator. This piezo was used to reset the phase of the interferometer prior to beginning the measurements, and was not employed for active feedback. The fiber beam-splitter at the output ensures perfect spatial mode overlap for high interferometric visibility.

RESULTS

Before demonstrating the quantum time flip in the context of the game, we first verified the ability of a polarization gadget to implement both a unitary and its transpose simultaneously, in the two different propagation directions of the light. To this end, we per-

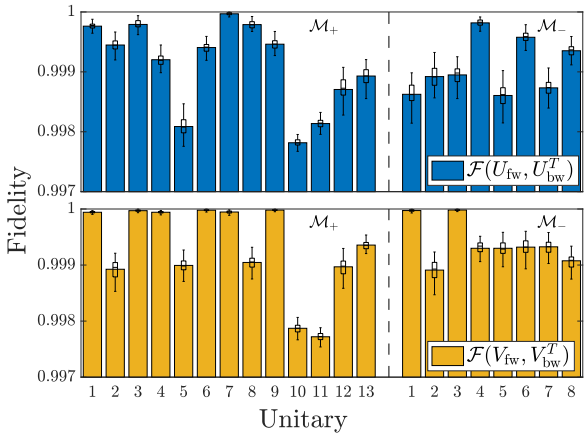


FIG. 4. **Unitary transposition fidelity.** The yellow and blue bars indicate the fidelity, $\mathcal{F}$, of the unitaries U (top) and V (bottom) from the sets $\mathcal{M}_+$ (left) and $\mathcal{M}_-$ (right), measured in the forward propagation direction, with respect to the transpose of the reconstructed unitary measured in the backwards propagation direction. Taller bars indicate a higher fidelity between the unitaries in the two propagation directions. The average fidelity is $0.9992 \pm 6.5 \times 10^{-4}$ indicating that the gadgets faithfully implement the transpose. The uncertainties were estimated using a Monte-Carlo simulation of the tomography accounting for errors in the waveplate angles, and the superimposed box plot indicates the spread of the reconstructed fidelities. We attribute the residual errors to imperfect waveplate retardance in the tomography, and angle differences between the setting of the forwards and backwards unitaries, since in principle the gadgets perfectly implement the transpose of the unitary in the forward direction.

formed quantum process tomography on the implemented unitaries from the sets $\mathcal{M}_+$ and $\mathcal{M}_-$, in both propagation directions. We then compared the fidelity $\mathcal{F} = \left\langle (U_{\rm fw}|\Psi\rangle)^\dagger U_{\rm bw}^T |\Psi\rangle \right\rangle_{|\Psi\rangle}$ between the reconstructed unitaries in the forward direction, $U_{\rm fw}$ and $V_{\rm fw}$, with the transposed reconstructed unitaries in the backwards direction, $U_{\rm bw}^T$ and $V_{\rm bw}^T$ (see Methods). The results of this are shown in Fig. 4. The average fidelity is greater than 0.999, indicating that the gadgets correctly implement the transpose. Note that the fidelity of the transpose is independent of any errors in the retardance of the waveplates in the gadget itself. Such imperfections would cause the fidelity in the implementation of a desired unitary to drop, but would affect the forward and backward directions symmetrically. The same is true for undesired offsets in the waveplate angles, however in the measurements shown in Fig. 4 the unitaries in the two directions were measured in separate runs, causing them to indeed be sensitive to waveplate angle errors, in addition to errors in the tomography itself.

Having verified the ability to implement a given unitary and its transpose with a single black box simultaneously, we then realised the game discussed in the previous sections. First, two-photon coincidence events for the different elements of $\mathcal{M}_+$ and $\mathcal{M}_-$ were collected sequentially to reduce the time spent rotating the waveplates. Second, the game itself was played using the collected data. In each round the referee uniformly randomly selects a pair of channels, and the player outputs an answer, '+' or '−', given by a unique two-photon event from the corresponding measurement set. Fig. 5 shows the relative frequencies $f^{\rm rel}_{\pm,k} = N^{\pm}_k / N_k$, where $N^{\pm}_k$ is the number of times the player output the answer '±' when the channels (U_k, V_k) were picked, and N_k is the total number of times these channels were selected by the referee. It can be seen that the player outputs the correct answer with a relative frequency higher than the indefinite tester bound of 0.92 for every setting, and by extension any strategy that is separable in its time direction. More specifically, the average winning frequency is found to be 0.9945, with the best and worst case frequencies being 0.9993 and 0.9860, respectively.

The formulation of the indefinite-time-direction witness as a game with only two outcomes, win or lose, allows for a straightforward statistical interpretation of the results. Since we have an upper bound $p_{\rm i.c.} \leq \frac{92}{100}$ on the probability of success for an indefinite tester , we can calculate the probability P of such a player having obtained v or more victories in N rounds:

$$P = \sum_{k=v}^{N} \binom{N}{k} p_{\rm i.c.}^k (1 - p_{\rm i.c.})^{N-k}. \tag{12}$$

This probability is exactly the P-value for the experimentally implemented process not being indefinite in its time direction. Out of the $N = 10^6$ rounds played in the experiment, $v = 994,512$ were won by successfully identifying the correct set, while 5,488 rounds were lost. Using a Chernoff bound tailored for the binomial distribution, we can provide an upper bound on the P-value, given by:

$$\sum_{k=v}^{N} \binom{N}{k} p_{\rm i.c.}^k (1 - p_{\rm i.c.})^{N-k} \leq \exp\left(-N\, D\left(\frac{v}{N} || p_{\rm i.c.}\right)\right), \tag{13}$$

where exp is the exponential function and:

$$D\left(\frac{v}{N} || p\right) := \frac{v}{N} \ln\left(\frac{v}{Np}\right) + \left(1 - \frac{v}{N}\right) \ln\left(\frac{1 - v/N}{1 - p}\right) \tag{14}$$

is the relative entropy. Direct calculation using $p_{\rm i.c.} = 0.92$ shows that $D\left(\frac{v}{N} || p_{\rm i.c.}\right) \approx 0.0627$, hence the P-value is upper bounded by $P \leq e^{-10^4}$, which is an extremely small number. This rules out any explanation of the data in terms of convex mixtures of quantum processes

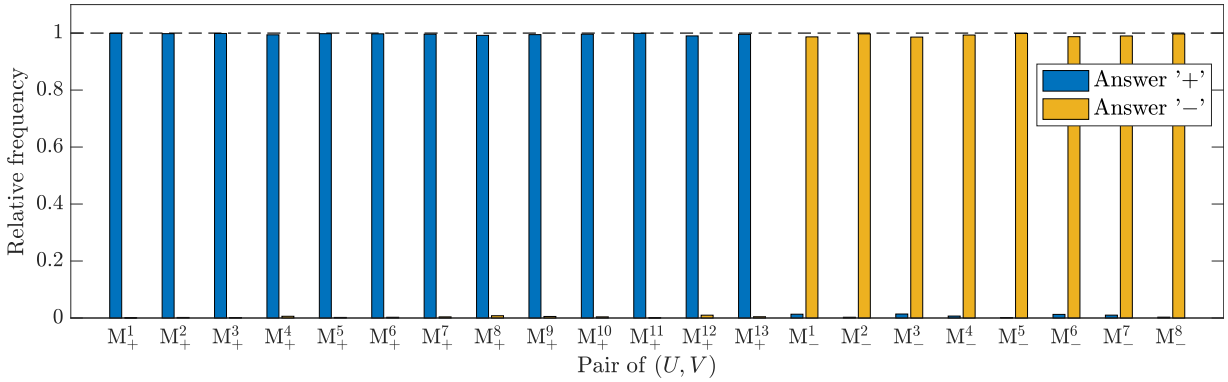


FIG. 5. **Observed relative outcome frequencies.** The figure shows the observed relative frequency of answers $f_{\pm}^{\text{rel}}$ in the quantum flip game for all the pairs of unitaries in the sets $\mathcal{M}_+$ and $\mathcal{M}_-$. For the gates in the set $\mathcal{M}_+$ ($\mathcal{M}_-$) the game is won when the player outputs the answer '+' ('−'). The observed average winning frequency is 0.9945. Since the bars corresponds to the actual number of times the different outcomes were recorded there is no associated uncertainty (see Methods).

that access the gates in a definite time direction. Since this is the defining characteristic for the class of processes with an indefinite time direction, we therefore conclude that the implemented process belongs to this class.

DISCUSSION

In this work we have demonstrated, for the first time, a process that is inseparable in its time direction. Using an optical interferometer, we implemented a coherent superposition of arbitrary unitary transformations and their time-reversal. Such a process can only be probabilistically simulated by a quantum circuit with a definite time direction. Even agents equipped with two copies of the gates and able to combine them in an indefinite order cannot realise the process deterministically, unless they are given the ability of pre- and post-selecting quantum systems [2, 36–41]. It is worth noting that our implementation of controlled unitary transposition is not in contradiction with the no-go theorem, stating that there is no quantum circuit that can transform an unknown quantum unitary gate to its transpose [10, 15, 42]. Our implementation adopts a device that implements a single-qubit gate U, and while this gate can remain unknown, the physical device itself is neither arbitrary nor unknown. Indeed, it is the particular symmetries of the physical device that necessarily and deterministically generate the transposed gate U^T. While time itself does not flow backwards in any part of the experimental apparatus, our demonstration highlights the limitations of the quantum circuit model for describing the full range of quantum information processing protocols. This is analogous to the impossibility of perfect unitary coherent control within the quantum circuit model [10, 25, 28, 43]. Through a channel discrimination game, in which we outperform any strategy with a definite time direction, we furthermore certify that the coherent superposition of time directions yields a process that is inseparable in its time direction.

The study of indefinite causality led to the discovery and realisation of quantum information protocols with practical advantages [44, 45], as well as a lively debate about the interpretation of these realisations [46–49]. We envision that future studies of processes with an indefinite time direction will similarly expand both the theoretical and experimental toolkit and open up new avenues for quantum information processing. In this context we note that universal transposition of single-qubit gates is a sufficient building block for the transposition of multi-qubit gates, for instance using a Reck decomposition [50], or through the inclusion of a reciprocal symmetric two-qubit gate [51]. Finally, the investigation of time reversed quantum processes also holds applications in quantum thermodynamics. Indeed, in [10], it was shown that the processes for which the quantum time flip produces another valid process are exactly those which do not increase the entropy in either time direction, and the application of superpositions of two time directions in the context of thermodynamic work was recently studied in [9, 52].

ACKNOWLEDGEMENTS

We thank Robert Peterson for fruitful discussions. This research was funded in whole, or in part, by

the Austrian Science Fund (FWF) through [F7113] (BeyondC), [FG5] (Research Group 5) and [TAI 483]. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. M.T.Q. and C.B. acknowledge the Austrian Science Fund (FWF) through the SFB project BeyondC (sub-project F7103), a grant from the Foundational Questions Institute (FQXi) as part of the Quantum Information Structure of Spacetime (QISS) Project (qiss.fr). The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 801110. It reflects only the authors' view, the EU Agency is not responsible for any use that may be made of the information it contains. ESQ has received funding from the Austrian Federal Ministry of Education, Science and Research (BMBWF).

AUTHOR CONTRIBUTIONS

T.S. and P.S. carried out the experiment and collected the data. T.S. and I.A. build the experimental setup. T.S., P.S. and M.T.Q. analysed the data. T.S. and P.S. designed the experiment. M.T.Q. carried out the computer assisted proofs. M.A. and L.R. conceived the experiment. P.W., C.B. and I.A. supervised the project. All authors contributed to writing the manuscript.

DATA AVAILABILITY

All data used in this work is openly available in Zenodo under: 10.5281/zenodo.7352614

CODE AVAILABILITY

The code used to perform the computer assisted proofs is openly available at the following online repository: https://github.com/mtcq/UnitaryTransposition

[1] L. Hardy, Probability Theories with Dynamic Causal Structure: A New Framework for Quantum Gravity, arXiv e-prints , gr-qc/0509120 (2005), arXiv:gr-qc/0509120 [gr-qc].

[2] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Phys. Rev. A **88**, 022318 (2013), arXiv:0912.0195 [quant-ph].

[3] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nature Communications **3**, 1092 (2012), arXiv:1105.4464 [quant-ph].

[4] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, Experimental superposition of orders of quantum gates, Nature Communications **6**, 7913 (2015), arXiv:1412.4006 [quant-ph].

[5] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, Experimental verification of an indefinite causal order, Science Advances **3**, e1602589 (2017), arXiv:1608.01683 [quant-ph].

[6] G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M. Zych, Č. Brukner, and P. Walther, Experimental entanglement of temporal order, Quantum **6**, 621 (2022), arXiv:1712.06884 [quant-ph].

[7] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, Indefinite Causal Order in a Quantum Switch, Phys. Rev. Lett. **121**, 090503 (2018), arXiv:1803.04302 [quant-ph].

[8] G. Rubino, L. A. Rozema, D. Ebler, H. Kristjánsson, S. Salek, P. Allard Guérin, A. A. Abbott, C. Branciard, Č. Brukner, G. Chiribella, and P. Walther, Experimental quantum communication enhancement by superposing trajectories, Physical Review Research **3**, 013093 (2021), arXiv:2007.05005 [quant-ph].

[9] G. Rubino, G. Manzano, and Č. Brukner, Quantum superposition of thermodynamic evolutions with opposing time's arrows, Communications Physics **4**, 1–10 (2021), arXiv:2008.02818 [quant-ph].

[10] G. Chiribella and Z. Liu, Quantum operations with indefinite time direction, Communications Physics **5**, 190 (2022), arXiv:2012.03859 [quant-ph].

[11] G. Chiribella and D. Ebler, Optimal quantum networks and one-shot entropies, New Journal of Physics **18**, 093053 (2016), arXiv:1606.02394 [quant-ph].

[12] I. S. B. Sardharwalla, T. S. Cubitt, A. W. Harrow, and N. Linden, Universal Refocusing of Systematic Quantum Noise, ArXiv e-prints (2016), arXiv:1602.07963 [quant-ph].

[13] M. Navascués, Resetting Uncontrolled Quantum Systems, Phys. Rev. X **8**, 031008 (2018), arXiv:1710.02470 [quant-ph].

[14] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Reversing Unknown Quantum Transformations: Universal Quantum Circuit for Inverting General Unitary Operations, Phys. Rev. Lett., **123**, 210502 (2019), arXiv:1810.06944 [quant-ph].

[15] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Probabilistic exact universal quantum circuits for transforming unitary operations, Phys. Rev. A **100**, 062339 (2019), arXiv:1909.01366 [quant-ph].

[16] M. T. Quintino and D. Ebler, Deterministic transformations between unitary operations: Exponential advantage with adaptive quantum circuits and the power of indefinite causality, Quantum **6**, 679 (2022), arXiv:2109.08202 [quant-ph].

[17] D. Trillo, B. Dive, and M. Navascués, Translating Uncontrolled Systems in Time, Quantum **4**, 374 (2020), arXiv:1903.10568 [quant-ph].

[18] D. Trillo, B. Dive, and M. Navascués, A universal quantum rewinding protocol with an arbitrarily high probability of success, arXiv e-prints (2022), arXiv:2205.01131 [quant-ph].

[19] P. Schiansky, T. Strömberg, D. Trillo, V. Saggio, B. Dive, M. Navascués, and P. Walther, Demonstration of universal time-reversal for quantum processes, arXiv e-prints (2022), arXiv:2205.01122 [quant-ph].

[20] S. Yoshida, A. Soeda, and M. Murao, Reversing unknown qubit-unitary operation, deterministically and exactly, arXiv e-prints (2022), arXiv:2209.02907 [quant-ph].

[21] M. M. Taddei, J. Cariñe, D. Martínez, T. García, N. Guerrero, A. A. Abbott, M. Araújo, C. Branciard, E. S. Gómez, S. P. Walborn, L. Aolita, and G. Lima, Computational advantage from the quantum superposition of multiple temporal orders of photonic gates, PRX Quantum **2**, 010320 (2021), arXiv:2002.07817 [quant-ph].

[22] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, 2000).

[23] J. Watrous, Quantum Computational Complexity, arXiv e-prints (2008), arXiv:0804.3401 [quant-ph].

[24] A. Soeda, Limitations on quantum subroutine designing due to the linear structure of quantum operators (2013).

[25] M. Araújo, A. Feix, F. Costa, and Č. Brukner, Quantum circuits cannot control unknown operations, New Journal of Physics **16**, 093026 (2014), arXiv:1309.7976 [quant-ph].

[26] A. Bisio, M. Dall'Arno, and P. Perinotti, Quantum conditional operations, Phys. Rev. A **94**, 022340 (2016), arXiv:1509.01062 [quant-ph].

[27] X.-Q. Zhou, T. C. Ralph, P. Kalasuwan, M. Zhang, A. Peruzzo, B. P. Lanyon, and J. L. O'Brien, Adding control to arbitrary unknown quantum operations, Nature Communications **2**, 413 (2011), arXiv:1006.2670 [quant-ph].

[28] J. Thompson, K. Modi, V. Vedral, and M. Gu, Quantum plug n' play: modular computation in the quantum regime, New Journal of Physics **20**, 013004 (2018).

[29] J. Bavaresco, M. Murao, and M. T. Quintino, Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination, Phys. Rev. Lett. **127**, 200504 (2021), 2011.08300 [quant-ph].

[30] G. Chiribella and D. Ebler, Quantum speedup in the identification of cause-effect relations, Nature Communications **10**, 1472 (2019), arXiv:1806.06459 [quant-ph].

[31] A. A. Abbott, J. Wechs, D. Horsman, M. Mhalla, and C. Branciard, Communication through coherent control of quantum channels, Quantum **4**, 333 (2020).

[32] J. Bavaresco, M. Araújo, Č. Brukner, and M. T. Quintino, Semi-device-independent certification of indefinite causal order, Quantum **3**, 176 (2019), arXiv:1903.10526 [quant-ph].

[33] H. Cao, J. Bavaresco, N.-N. Wang, L. A. Rozema, C. Zhang, Y.-F. Huang, B.-H. Liu, C.-F. Li, G.-C. Guo, and P. Walther, Experimental semi-device-independent certification of indefinite causal order, arXiv e-prints (2022), arXiv:2202.05346 [quant-ph].

[34] R. Simon and N. Mukunda, Minimal three-component su(2) gadget for polarization optics, Physics Letters A **143**, 165–169 (1990).

[35] N. J. Frigo, F. Bucholtz, G. A. Cranch, and J. M. Singley, On the choice of conventions in polarization evolution calculations and representations, J. Lightwave Technol. **40**, 179–190 (2022).

[36] R. Oeckl, General boundary quantum field theory: Foundations and probability interpretation, arXiv e-prints , hep-th/0509122 (2005), arXiv:hep-th/0509122 [hep-th].

[37] O. Oreshkov and N. J. Cerf, Operational quantum theory without predefined time, New Journal of Physics **18**, 073037 (2016), arXiv:1406.3829 [quant-ph].

[38] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Phys. Rev. A **88**, 022318 (2013), arXiv:0912.0195 [quant-ph].

[39] R. Silva, Y. Guryanova, A. J. Short, P. Skrzypczyk, N. Brunner, and S. Popescu, Connecting processes with indefinite causal order and multi-time quantum states, New Journal of Physics **19**, 103022 (2017), arXiv:1701.08638 [quant-ph].

[40] G. Svetlichny, Time Travel: Deutsch vs. Teleportation, International Journal of Theoretical Physics **50**, 3903–3914 (2011), arXiv:0902.4898 [quant-ph].

[41] S. Lloyd, L. Maccone, R. Garcia-Patron, V. Giovannetti, Y. Shikano, S. Pirandola, L. A. Rozema, A. Darabi, Y. Soudagar, L. K. Shalm, and A. M. Steinberg, Closed Timelike Curves via Postselection: Theory and Experimental Test of Consistency, Phys. Rev. Lett. **106**, 040403 (2011), arXiv:1005.2219 [quant-ph].

[42] G. Chiribella, E. Aurell, and K. Życzkowski, Symmetries of quantum evolutions, Physical Review Research **3**, 033028 (2021), arXiv:2101.04962 [quant-ph].

[43] S. Nakayama, A. Soeda, and M. Murao, Universal construction of controlled-unitary gates using dynamical decoupling and the quantum zeno effect, AIP Conference Proceedings **1633**, 183–185 (2014), https://aip.scitation.org/doi/pdf/10.1063/1.4903131.

[44] P. Allard Guérin, A. Feix, M. Araújo, and Č. Brukner, Exponential Communication Complexity Advantage from Quantum Superposition of the Direction of Communication, Physical review letters **117**, 100502 (2016), arXiv:1605.07372 [quant-ph].

[45] K. Wei, N. Tischler, S.-R. Zhao, Y.-H. Li, J. M. Arrazola, Y. Liu, W. Zhang, H. Li, L. You, Z. Wang, Y.-A. Chen, B. C. Sanders, Q. Zhang, G. J. Pryde, F. Xu, and J.-W. Pan, Experimental Quantum Switching for Exponentially Superior Quantum Communication Complexity, Phys. Rev. Lett. **122**, 120504 (2019), arXiv:1810.10238 [quant-ph].

[46] O. Oreshkov, Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics, Quantum **3**, 206 (2019), arXiv:1801.07594 [quant-ph].

[47] N. Paunković and M. Vojinović, Causal orders, quantum circuits and spacetime: distinguishing between definite and superposed causal orders, Quantum **4**, 275 (2020), arXiv:1905.09682 [quant-ph].

[48] M. Fellous-Asiani, R. Mothe, L. Bresque, H. Dourdent, P. A. Camati, A. A. Abbott, A. Auffèves, and C. Branciard, Comparing the quantum switch and its

simulations with energetically-constrained operations, arXiv preprint arXiv:2208.01952 , arXiv:2208.01952 (2022), arXiv:2208.01952 [quant-ph].

[49] V. Vilasini and R. Renner, Embedding cyclic causal structures in acyclic spacetimes: no-go results for process matrices, arXiv preprint arXiv:2203.11245 , arXiv:2203.11245 (2022), arXiv:2203.11245 [quant-ph].

[50] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Experimental realization of any discrete unitary operator, Physical review letters **73**, 58 (1994).

[51] I. Alonso Calafell, J. Cox, M. Radonjić, J. Saavedra, F. García de Abajo, L. Rozema, and P. Walther, Quantum computing with graphene plasmons, npj Quantum Information **5**, 37 (2019).

[52] G. Rubino, G. Manzano, L. A. Rozema, P. Walther, J. M. R. Parrondo, and Č. Brukner, Inferring work by quantum superposing forward and time-reversal evolutions, Physical Review Research **4**, 013208 (2022), arXiv:2107.02201 [quant-ph].

[53] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009), arXiv:0904.4483 [quant-ph].

[54] M. Ziman, Process positive-operator-valued measure: A mathematical framework for the description of process tomography experiments, Phys. Rev. A **77**, 062112 (2008), arXiv:0802.3862 [quant-ph].

METHODS

Arbitrary unitary transposition

The description of linear retarders depends on the convention used for the polarization states, i.e. which Pauli matrices are associated with which Stokes parameters. The most commonly used convention in quantum optics is:

$$(S_1, S_2, S_3) \leftrightarrow (Z, X, Y), \tag{15}$$

corresponding to the $\{H, V\}$, $\{+, -\}$ and $\{L, R\}$ polarizations being the eigenstates of Z, X and Y respectively. Under this convention, a linear retarder, such as a waveplate, at an angle θ to the vertical axis, is described by the following matrix:

$$\begin{aligned} U(\theta) &= e^{-\frac{i}{2}\theta Y} e^{-\frac{i}{2} r Z} e^{\frac{i}{2}\theta Y} \\ &= \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} e^{i\frac{r}{2}} & 0 \\ 0 & e^{-i\frac{r}{2}} \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \end{aligned} \tag{16}$$

where r is the retardance of the element. Note that the matrix $U(\theta)$ is symmetric since:

$$\begin{aligned} U(\theta)^T &= (e^{\frac{i}{2}\theta Y})^T (e^{-\frac{i}{2} r Z})^T (e^{-\frac{i}{2}\theta Y})^T \\ &= e^{-\frac{i}{2}\theta Y} e^{-\frac{i}{2} r Z} e^{\frac{i}{2}\theta Y}. \end{aligned} \tag{17}$$

Propagating through such an element backwards has the effect of taking $\theta \to -\theta$. This transformation can be written as:

$$Z U(\theta) Z = U(-\theta) \tag{18}$$

since

$$Z e^{-\frac{i}{2}\theta Y} Z = e^{\frac{i}{2}\theta Y}. \tag{19}$$

For a general polarization gadget consisting of several linear retarders described by the unitary $U_{G,\mathrm{fw}}$ in the forwards direction we find the unitary for the backwards propagation direction, $U_{G,\mathrm{bw}}$, by transposing the order of the individual linear retarders and changing the sign of their respective angles:

$$U_{G,\mathrm{fw}} = U_1(\theta_1) \dots U_n(\theta_n) \to U_{G,\mathrm{bw}} = U_n(-\theta_n) \dots U_1(-\theta_1) \tag{20}$$

which can be written:

$$U_{G,\mathrm{fw}} \to Z U_{G,\mathrm{fw}}^T Z \tag{21}$$

since:

$$\begin{aligned} Z(U_1(\theta_1) \dots U_n(\theta_n))^T Z &= Z U_n(\theta_n) \dots U_1(\theta_1) Z \\ &= Z U_n(\theta_n) Z \dots Z U_1(\theta_1) Z \\ &= U_n(-\theta_n) \dots U_1(-\theta_1). \end{aligned} \tag{22}$$

The transformation in Eq. (21) is not useful for realising the transpose, since the Z gates around the unitary $U_{G,\mathrm{fw}}^T$ have to be undone to recover the transpose.

However, this problem can be overcome by picking a different convention for the polarization basis states, such as $(S_1, S_2, S_3) \leftrightarrow (X, Y, Z)$ which is a cyclic permutation of the aforementioned one (corresponding to a rotation of the basis vectors by $\pi/3$ around the vector $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$), and which is commonly used in polarimetry. In this work, we chose the convention:

$$(S_1, S_2, S_3) \leftrightarrow (-Z, -Y, -X). \tag{23}$$

The minus signs are necessary to preserve the handedness of the coordinate system when exchanging X and Y. That this convention yields the desired transformation under counterpropagation can be realised by noting that the Stokes parameters of a unitary always transform as $(S_1, S_2, S_3) \to (S_1, -S_2, S_3)$, however for completeness we will perform the calculation explicitly. In the convention of Eq. (23) a linear retarder at an angle θ is written as:

$$U(\theta) = e^{\frac{i}{2}\theta X} e^{\frac{i}{2} r Z} e^{-\frac{i}{2}\theta X} \tag{24}$$

and the corresponding unitary in the backwards direction is

$$\begin{aligned} U(-\theta) &= e^{-\frac{i}{2}\theta X} e^{\frac{i}{2} r Z} e^{\frac{i}{2}\theta X} \\ &= (e^{\frac{i}{2}\theta X} e^{\frac{i}{2} r Z} e^{-\frac{i}{2}\theta X})^T \\ &= U(\theta)^T. \end{aligned} \tag{25}$$

It then follows that a general waveplate gadget also transforms as the transpose:

$$\begin{aligned} U_{G,\text{fw}} = U_1(\theta_1) \dots U_n(\theta_n) \to U_{G,\text{bw}} &= U_n(-\theta_n) \dots U_1(-\theta_1) \\ &= (U_1(\theta_1) \dots U_n(\theta_n))^T \\ &= U_{G,\text{fw}}^T. \end{aligned} \tag{26}$$

One could alternatively get around the problem with Eq. (21) by introducing two more polarization gadgets implementing Z operators on either side of the gadget in Eq. (21), and making sure that these additional gadgets only act on one propagation direction. For example, by physically displacing the beam paths of the two propagation directions, so that the gadgets act on different spatial modes in the different propagation directions. This would, however, change the interpretation of the experiment with respect to the implementation in the main text, since the transformations in the two propagation directions would no longer be related by a physical symmetry. Instead they would depend on the transformations realised by the additional gadgets.

Sets of unitary operators used in the game

In this section, we explicitly list the two sets of unitary operators used in the discrimination task considered in this work. These sets of operators were first presented at Ref. [10].

$$\begin{aligned} \mathcal{M}_+ := \Bigg\{ & (I,I), (I,X), (I,Z), (X,I), (X,X), (X,Z), (Z,I), (Z,X), (Z,Z), \\ & \left(\frac{X-Y}{\sqrt{2}}, \frac{X+Y}{\sqrt{2}}\right), \left(\frac{X+Y}{\sqrt{2}}, \frac{X-Y}{\sqrt{2}}\right), \left(\frac{Z-Y}{\sqrt{2}}, \frac{Z+Y}{\sqrt{2}}\right), \left(\frac{Z+Y}{\sqrt{2}}, \frac{Z-Y}{\sqrt{2}}\right) \Bigg\} \\ \mathcal{M}_- := \Bigg\{ & (Y,I), (Y,X), (Y,Z), (I,Y), (X,Y), (Z,Y), \\ & \left(\frac{I+iY}{\sqrt{2}}, \frac{I-iY}{\sqrt{2}}\right), \left(\frac{I-iY}{\sqrt{2}}, \frac{I+iY}{\sqrt{2}}\right) \Bigg\}. \end{aligned} \tag{27}$$

Obtaining upper bounds for different classes of strategies

We now detail how to obtain an upper bound on the winning probability of the game described in the main manuscript. Let N be the total number of pairs of unitary operators contained in the set $\mathcal{M}_+$ and $\mathcal{M}_-$. Following a uniform distribution, i.e., with probably $1/N$, the referee picks a pair of unitary operators (U_i, V_i). The player should then employ a quantum strategy to guess whether (U_i, V_i) belongs to $\mathcal{M}_+$ or $\mathcal{M}_-$. Let $p(\pm|(U_i, V_i))$ the probability that the player guesses $(U_i, V_i) \in \mathcal{M}_\pm$. The probability of such player to win the game is then given by

$$p = \frac{1}{N} \left(\sum_{(U_i,V_i)\in\mathcal{M}_+} p(+|(U_i,V_i)) + \sum_{(U_i,V_i)\in\mathcal{M}_-} p(-|(U_i,V_i)) \right). \tag{28}$$

For the qubit scenario considered here, we can analyse the case where unitary gates act backwards by simply considering the case where all involved unitary operators are transposed. This is true because, as discussed earlier, there are only two anti-homomorphisms from $\mathcal{SU}(d)$ to $\mathcal{SU}(d)$, and for any $U \in \mathcal{SU}(2)$, we have that $U^{-1} = \sigma_Y U^T \sigma_Y$. More explicitly, the winning probability for players using the unitary gates backwards is given by

$$p = \frac{1}{N}\left(\sum_{(U_i,V_i)\in\mathcal{M}_+} p(+|(U_i^T, V_i^T)) + \sum_{(U_i^T,V_i^T)\in\mathcal{M}_-} p(-|(U_i^T, V_i^T)) \right). \tag{29}$$

Also, as we show more explicitly later, since the success probability is linear function of the strategies, convex combinations of forward and backwards strategies cannot increase the maximal success probability. Hence it is enough to analyse the forward and backwards case.

When the player is restricted to parallel strategies, the most general approach consists of preparing a quantum state ρ, sending part of this state to the operators U_i and V_i, and then performing a quantum measurement with outcomes labelled as $+$ or $-$, that is,

$$p_{\text{par}}(\pm|(U_i, V_i)) = \text{tr}\Big[M_\pm \Big(U_i \otimes V_i \otimes \mathbb{1})\rho(U_i^\dagger \otimes V_i^\dagger \otimes \mathbb{1})\Big)\Big], \tag{30}$$

where $M_+, M_- \geq 0$ are the POVM operators associated to the outcomes $+$ and $-$, see Fig. 2 for a pictorial illustration.

Parallel strategies may be analysed in the (parallel) tester formalism [29, 53], also known as process POVM [54]. Let us label the linear spaces corresponding to the input and output spaces as $\mathcal{H}_I$ and $\mathcal{H}_O$ respectively. We can then write $U_i \otimes V_i : \mathcal{H}_I \to \mathcal{H}_O$ with $\mathcal{H}_I \cong \mathcal{H}_O \cong \mathbb{C}_2 \otimes \mathbb{C}_2$. In the tester formalism, operations are viewed as states and Eq. (30) may be written as the generalized Born's rule. More formally, we have that:

$$p_{\text{par}}(\pm|(U_i, V_i)) = \text{tr}\Big[T_\pm |U_i \otimes V_i\rangle\!\rangle\langle\!\langle U_i \otimes V_i|\Big], \tag{31}$$

where[1] $T_+, T_- \in \mathcal{L}(\mathcal{H}_I \otimes \mathcal{H}_O)$ are tester elements and $|U_i \otimes V_i\rangle\!\rangle \in (\mathcal{H}_I \otimes \mathcal{H}_O)$ is the Choi vector of $U_i \otimes V_i$ defined as:

$$|U_i \otimes V_i\rangle\!\rangle := \sum_l |l\rangle \otimes \Big(U_i \otimes V_i\,|l\rangle\Big), \tag{32}$$

where $\{|l\rangle\}$ is the computational basis for $\mathcal{H}_I$. The operators T_+ and T_- are parallel testers when $T_+, T_- \geq 0$ and their sum respects:

$$T_+ + T_- = \sigma_I \otimes \mathbb{1}_O, \tag{33}$$

where $\sigma \in \mathcal{L}(\mathcal{H}_I)$ is a quantum state. As shown in Refs. [29, 53, 54], all parallel strategies as in Eq. (30) can be represented by testers such as those in Eq. (31), and *vice versa*. Hence, when optimizing over all possible strategies, instead of considering all possible states ρ and measurements $M_\pm$ as in Eq. (30), we may optimize over all valid testers $T_\pm$ as in Eq. (31).

One advantage of using the tester formalism, is that the maximal probability of winning the discrimination game can be written in terms of a semidefinite program (SDP) via the following optimisation problem:

$$\max \frac{1}{N}\left[\sum_{(U_i,V_i)\in\mathcal{M}_+} \text{tr}\Big(T_+ \,|U_i \otimes V_i\rangle\!\rangle\langle\!\langle U_i \otimes V_i|\,\Big) + \sum_{(U_i,V_i)\in\mathcal{M}_-} \text{tr}\Big(T_- \,|U_i \otimes V_i\rangle\!\rangle\langle\!\langle U_i \otimes V_i|\Big)\right] \tag{34}$$

$$\text{s.t.: } T_+, T_- \geq 0 \tag{35}$$

$$T_+ + T_- = \sigma_I \otimes \mathbb{1}_O \tag{36}$$

$$\text{tr}(\sigma) = 1. \tag{37}$$

[1] Here $\mathcal{L}(\mathcal{H}_I \otimes \mathcal{H}_O)$ denotes the set of linear operators from $\mathcal{H}_I \otimes \mathcal{H}_O$ (linear endomorphisms).

Following the steps of Ref. [29], the dual problem is given by:

$$\min \operatorname{tr}(C)/d_I \tag{38}$$

$$\text{s.t.: } \frac{1}{N}\sum_{(U_i,V_i)\in\mathcal{M}_+} |U_i\otimes V_i\rangle\!\rangle\langle\!\langle U_i\otimes V_i| \leq C \tag{39}$$

$$\frac{1}{N}\sum_{(U_i,V_i)\in\mathcal{M}_-} |U_i\otimes V_i\rangle\!\rangle\langle\!\langle U_i\otimes V_i| \leq C \tag{40}$$

$$\operatorname{tr}_O(C) = \operatorname{tr}_{IO}(C)\,\frac{\mathbb{1}_I}{d_I}, \tag{41}$$

where d_I is the dimension of $\mathcal{H}_I$ (for our particular problem, $d_I = 4$). By the definition of dual problem, if we find a linear operator C satisfying the feasibility constraints of inequality (39), inequality (40), and Eq. (41), the quantity $\operatorname{tr}(C)/d_I$ is an upper bound on the maximal success probability. In order to obtain a computer-assisted-proof upper bound with fraction of integers, we use standard and efficient floating-point arithmetic algorithms to solve the SDP, obtain an operator C which satisfies the constraints of the dual problem and truncate it in such a way that the feasibility constraints are still satisfied. We refer to our online repository (see Code Availability) for an implementation of this procedure and to Ref. [29] for a detailed explanation on how to perform the truncation step.

When the player is restricted to causal strategies (also referred to as sequential strategies), the most general approach consists of preparing a quantum state ρ, sending part of this state to the operators U_i (or to V_i), applying a quantum channel $\mathcal{E}$, then performing the operation V_i (or U_i), and finally performing a quantum measurement with outcomes labelled as $+$ or $-$, that is:

$$p_{\text{seq}}(\pm|(U_i,V_i)) = \operatorname{tr}\Big[M_\pm\,(V_i\otimes\mathbb{1})\mathcal{E}\Big(U_i\otimes\mathbb{1}\,\rho\,U_i^\dagger\otimes\mathbb{1}\Big)(V_i^\dagger\otimes\mathbb{1})\Big].$$

Using the concept of sequential testers [29, 53], we can also write the problem of finding the optimal causal strategy as an SDP. Since there is a notion of causal order, we label the input and output space of the first operation as $\mathcal{H}_{I_1}$ and $\mathcal{H}_{O_1}$ respectively. Analogously, we use $\mathcal{H}_{I_2}$ and $\mathcal{H}_{O_2}$ for the second operations. If the player uses the operation U_i first and V_i second, we have that $U_i : \mathcal{H}_{I_1} \to \mathcal{H}_{O_1}$ and $V_i : \mathcal{H}_{I_2} \to \mathcal{H}_{O_2}$. Following Ref. [29], the primal and dual problem for causal strategies are respectively given by

$$\max \frac{1}{N}\left[\sum_{(U_i,V_i)\in\mathcal{M}_+} \operatorname{tr}\Big(T_+\,|U_i\otimes V_i\rangle\!\rangle\langle\!\langle U_i\otimes V_i|\,\Big) + \sum_{(U_i,V_i)\in\mathcal{M}_-} \operatorname{tr}\Big(T_-\,|U_i\otimes V_i\rangle\!\rangle\langle\!\langle U_i\otimes V_i|\Big)\right] \tag{42}$$

$$\text{s.t.: } T_+, T_- \geq 0 \tag{43}$$

$$T_+ + T_- = W_{I_1O_1I_2}\otimes\mathbb{1}_{O_2} \tag{44}$$

$$\operatorname{tr}_{I_2}(W_{I_1O_1I_2}) = \sigma_{I_1}\otimes\mathbb{1}_{O_1} \tag{45}$$

$$\operatorname{tr}(\sigma) = 1. \tag{46}$$

and

$$\min \operatorname{tr}(C)/d_I \tag{47}$$

$$\text{s.t.: } \frac{1}{N}\sum_{(U_i,V_i)\in\mathcal{M}_+} |U_i\otimes V_i\rangle\!\rangle\langle\!\langle U_i\otimes V_i| \leq C \tag{48}$$

$$\frac{1}{N}\sum_{(U_i,V_i)\in\mathcal{M}_-} |U_i\otimes V_i\rangle\!\rangle\langle\!\langle U_i\otimes V_i| \leq C \tag{49}$$

$$\operatorname{tr}_{O_2}(C) = \operatorname{tr}_{I_2O_2}(C)\otimes\frac{\mathbb{1}_{I_2}}{d_{I_2}} \tag{50}$$

$$\operatorname{tr}_{O_1I_2O_2}(C) = \operatorname{tr}_{I_1O_1I_2O_2}(C)\,\frac{\mathbb{1}_{I_1}}{d_{I_1}}. \tag{51}$$

Another sequential strategy would be to use V_i before U_i. For this case, the semidefinite program is then exactly the same as the one before, but we exchange the roles of V_i and U_i. Our methods show that, when U_i precedes V_i, the success probability is bounded by $\frac{90}{100} \leq p_{UV} \leq \frac{91}{100}$, and when V_i precedes U_i, the success probability is bounded by $\frac{90}{100} \leq p_{VU} \leq \frac{91}{100}$. Since the two bounds coincide, we have $\frac{90}{100} \leq p_{\text{causal}} \leq \frac{91}{100}$.

When the player is restricted to general quantum strategies without a definite causal order, the strategies are described by means of an indefinite tester [30], which are positive semidefinite operators that add up to a process matrix [3], that is $T_+ + T_- = W$, where W is a bipartite process matrix. Following Ref. [29], and defining the trace-and-replace maps as ${}_iX := \text{tr}_i(X) \otimes \mathbb{1}_i$, the primal and the dual problem are respectively given by

$$\max \frac{1}{N}\Bigg[\sum_{(U_i,V_i)\in\mathcal{M}_+} \text{tr}\Big(T_+ \, |U_i \otimes V_i\rangle\!\rangle\langle\!\langle U_i \otimes V_i| \, \Big) + \sum_{(U_i,V_i)\in\mathcal{M}_-} \text{tr}\Big(T_- \, |U_i \otimes V_i\rangle\!\rangle\langle\!\langle U_i \otimes V_i|\Big)\Bigg] \tag{52}$$

$$\text{s.t.: } T_+, T_- \geq 0 \tag{53}$$

$$T_+ + T_- = W \tag{54}$$

$${}_{I_2O_2}W = {}_{O_1I_2O_2}\, W \tag{55}$$

$${}_{I_1O_1}W = {}_{O_2I_1O_1}\, W \tag{56}$$

$$W = {}_{O_1}\, W + {}_{O_2}\, W - {}_{O_1O_2}\, W \tag{57}$$

$$\text{tr}(W) = \text{tr}(\mathbb{1}_{O_1O_2}). \tag{58}$$

and

$$\min \text{tr}(C)/d_I \tag{59}$$

$$\text{s.t.: } \frac{1}{N} \sum_{(U_i,V_i)\in\mathcal{M}_+} |U_i \otimes V_i\rangle\!\rangle\langle\!\langle U_i \otimes V_i| \leq C \tag{60}$$

$$\frac{1}{N} \sum_{(U_i,V_i)\in\mathcal{M}_-} |U_i \otimes V_i\rangle\!\rangle\langle\!\langle U_i \otimes V_i| \leq C \tag{61}$$

$${}_{O_1}C = {}_{I_1O_1}\, C \tag{62}$$

$${}_{O_2}C = {}_{I_2O_2}\, C. \tag{63}$$

Data analysis

As described in the main text, the game was played by having the referee pick pairs of unitaries from the sets $\mathcal{M}_\pm$ in a uniformly random way in every round. The player's outcome was determined by the first unused photon detection event in the event list corresponding to that choice of unitary by the referee. More concretely, let $O_\pm^{j,k}$ be the k-th element in the time ordered list of detection events $\mathcal{O}_\pm^j$ for the implemented pair of unitaries $M_\pm^j$. Then the outcome of the n-th round of the game, in which the referee picked the pair of unitaries $M_\pm^j$ for the k-th time, is $O_\pm^{j,k}$.

During the course of this game the player outputs the answer '+' ('−') a total of N_j^+ (N_j^-) times in the N_j rounds that the referee selects the pair of channels (U_j, V_j). The relative frequencies with which the player outputs these answers can be written as:

$$f_{+,j}^{\text{rel}} = \frac{N_j^+}{N^j}, \quad f_{-,j}^{\text{rel}} = \frac{N_j^-}{N^j}. \tag{64}$$

These observed relative frequencies are shown in Fig. 5 in the main text. The values of these observed relative frequencies do not by themselves have an associated uncertainty, and are purely observed quantities. In many single photon experiments, quantities such as these are interpreted as empirical estimates of underlying probabilities, and such estimates do carry uncertainties. Even in perfect experimental realisations, finite counting statistics would

introduce Poissonian noise in this type of estimation. However, the statistical method we use to determine the confidence in our conclusion - the calculation of the P-value - allows us to make statements about the underlying probability distribution without directly estimating it. Specifically, that its expectation value exceeds the bound imposed on the winning probability of any strategy with a definite time direction.

In order to filter out background events resulting from various back-reflections in the experimental setup, as well as detector dark counts, two-fold coincidence events between the signal and idler photons were used to time filter the detection events.

The superconducting nanowire detectors used in the experiment have a slight polarization dependence in their detection efficiency, and due to the different pairs of unitaries generating different target qubit states the event rates for different implemented unitaries varied. This difference in efficiency was not necessary to account for, because the number of events for each pair of unitaries was truncated, in reverse chronological order, to match the setting with the fewest events. To find the numbers of rounds won and lost, the data was sampled from once, drawing 10^6 different samples from unique, chronologically ordered (for each setting) detection events. The exact number of won and lost rounds in this sampling were $994,512$ won and $5,488$ lost.

A detection efficiency imbalance is also present in the two output ports of the interferometer, corresponding to the two different measurement outcomes of the control qubit. This efficiency difference could quite easily be characterised and corrected for, however such actions are equivalent to classical post-processing and is captured by the indefinite tester. Imbalanced detection efficiency could therefore not lead to a violation of the bound, and is not necessary to correct for since the data already violates the bound. This is a different way of stating the semi-device independence of our methods.

The measurement of the fidelity between the unitary implemented in one direction and the transpose of the unitary in the other direction was performed with coherent light. To estimate the fidelity, the two unitaries were first fitted to the data using a maximum likelihood estimation and then the fidelity was calculated by evaluating the following average:

$$\mathcal{F} = \left\langle (U_{\text{fw}}|\Psi\rangle)^\dagger U_{\text{bw}}^T |\Psi\rangle \right\rangle_{|\Psi\rangle}, \tag{65}$$

taken over 1000 Haar-random states $|\Psi\rangle$. This was done in every step of a Monte-Carlo simulation to estimate the measurement uncertainties induced by the waveplate errors.

Semi-device independence of demonstration

In this section we will elaborate on what is meant by our certification methods being semi-device independent. Our usage of this term is consistent with the notion of semi-device independence introduced in [32]. That our demonstration is semi-device independent means that the measurement that the player performs does not have to be characterised. Equivalently, the player does not have to trust that their measurement device implements a specific measurement. It is a statement about the required assumptions on the measurement.

The basis for the claim that our demonstration is semi-device independent lies in the fact that the derivation of the bounds for the strategies depicted in Fig. 2.a-c in the main text included an optimization over all possible binary measurements the player could perform. This means that there is no measurement that a player using these strategies could perform that would allow them to violate the bounds we derived. Hence, a violation of these bounds has the same interpretation regardless of what measurements the player performed.

It is worth noting that semi-device independence does not imply that the ability of the player to violate the bounds is independent of the measurement they perform. Indeed, measurement imperfections can reduce the winning rate of the player. This can cause them to fail to certify that they employ a certain strategy, even if they do in fact employ that strategy.

A concrete consequence of the semi-device independence is that imperfections in the measurement do not need to be accounted for, and the measurement itself does not need to be modelled in the data analysis. This is in contrast to device-dependent methods, which rely on well characterised measurements to draw conclusions about the observed results. A device-dependent verification method that frequently appears in experimental quantum information science is the witness operator, for example entanglement witnesses or causal witnesses. Such witness operators can also be constructed for the task described in the main text. A witness operator $\hat{S}$ can be used to

certify a certain statement about a quantum system or process by experimentally evaluating its expectation value, and confirming that it satisfies some bound:

$$\langle \hat{S} \rangle < B. \tag{66}$$

Empirically evaluating $\langle \hat{S} \rangle$ requires the witness operator to be decomposed in terms of experimentally measurable observables, and the expectation values of these observables to be estimated. Imperfections in the measurement devices induce uncertainties in these estimates, which in turn propagate as uncertainties into the expectation value of the witness operator. A statistically significant violation of the inequality (66) therefore requires well characterised measurement devices.

11 Reprint of "Demonstration of a Quantum Switch in a Sagnac Configuration"

11.1 Contribution statement

The idea for this work grew out of the study of polarization gadgets with novel properties under counterpropagation, which was carried out by myself and Robert Peterson, with input from Peter Schiansky. I conceived and designed the experiment. After preliminary experimental work by myself, Peter Schiansky assembled the optical setup, carried out the experiment and collected the data. Marco Túlio Quintino and I contributed to data analysis carried out by Peter Schiansky. The manuscript was written by myself together with Marco Túlio Quintino, with input from Peter Schiansky and Robert Peterson, and under the supervision of Philip Walther. All the figures were made by Peter Schiansky and myself. Marco Túlio Quintino carried out the computer assisted proofs. The proof of the polarization gadget's universality was performed by myself.

Demonstration of a Quantum Switch in a Sagnac Configuration

Teodor Strömberg,[1,2,*] Peter Schiansky,[1,2] Robert W. Peterson,[2] Marco Túlio Quintino,[3,4,5] and Philip Walther[2,†]
[1]University of Vienna, Faculty of Physics & Vienna Doctoral School in Physics, Boltzmanngasse 5, A-1090 Vienna, Austria
[2]University of Vienna, Faculty of Physics & Research Network Quantum Aspects of Space Time (TURIS), Boltzmanngasse 5, 1090 Vienna, Austria
[3]Sorbonne Université, CNRS, LIP6, F-75005 Paris, France
[4]University of Vienna, Faculty of Physics, Boltzmanngasse 5, 1090 Vienna, Austria
[5]Institute for Quantum Optics and Quantum Information, Boltzmanngasse 3, 1090 Vienna, Austria

(Received 24 November 2022; accepted 12 July 2023; published 11 August 2023)

The quantum switch is an example of a process with an indefinite causal structure, and has attracted attention for its ability to outperform causally ordered computations within the quantum circuit model. To date, realizations of the quantum switch have made a trade-off between relying on optical interferometers susceptible to minute path length fluctuations and limitations on the range and fidelity of the implementable channels, thereby complicating their design, limiting their performance, and posing an obstacle to extending the quantum switch to multiple parties. In this Letter, we overcome these limitations by demonstrating an intrinsically stable quantum switch utilizing a common-path geometry facilitated by a novel reciprocal and universal SU(2) polarization gadget. We certify our design by successfully performing a channel discrimination task with near unity success probability.

DOI: 10.1103/PhysRevLett.131.060803

Introduction.—Quantum information processing tasks are most commonly described within the framework of the quantum circuit model. In this framework an initial state gradually evolves by passing through a fixed sequence of gates. This, however, is not the most general model of computation that quantum mechanics admits, and in [1] a processes that effects a superposition of quantum circuits was proposed. This process, known as the quantum switch, has attracted significant theoretical [2–5] and experimental [6–13] interest. Together with the so-called Oreshkov-Costa-Brukner process [14] it was the first example of a quantum process without a definite causal structure, and motivated the study of more general causal structures within quantum mechanics that could help bridge the gap between general relativity and quantum mechanics. The quantum switch is also of practical interest, since it has been shown to allow for a computational advantage over standard quantum circuits [15,16], an advantage which has been demonstrated experimentally [17].

In its simplest form, the quantum switch is a map that acts on two gates, U and V, and transforms them into a controlled superposition of the gates being applied in two different orders:

$$(U, V) \mapsto UV \otimes |0\rangle\langle 0|_C + VU \otimes |1\rangle\langle 1|_C. \quad (1)$$

To date, all experimental realizations of the quantum switch have been done using single photons as the physical system encoding the input and output state of the process. These implementations typically rely on folded Mach-Zehnder interferometers (MZIs) and polarization optics to couple different internal degrees of freedom of the single photons. A technical challenge associated with such implementations is that the phase of the interferometer needs to be kept constant even as different choices of U and V in (1) change the interference condition. In practice, most experimental quantum switches have relied on passive phase stability during operation, limiting not only their fidelity, but also their duty cycle due to the need to periodically reset the phase. Furthermore, the geometry of the MZI means that single photons in the two different arms of the interferometer interact with different parts of the polarization optics [see Fig. 1(a)]. The reliance on optical geometries that suffer from phase instability is necessitated by the nonreciprocity of the optical components that effect the unitary transformations U and V on the photon polarization. The works [8,10] remedied this by using the polarization as a control degree of freedom (d.o.f.), instead of a target, thereby enabling a common-path geometry; however, this came at

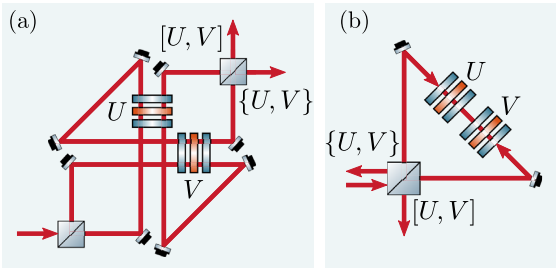


FIG. 1. Path-polarization quantum switch. (a) The most common implementation of the photonic quantum switch utilizes the path degree of freedom of a single photon inside a Mach-Zehnder interferometer to coherently control the order in which two polarization operations U and V are applied. In this geometry, photons in different arms of the interferometer propagate through different parts of the polarization optics. (b) An implementation based on a Sagnac interferometer is fundamentally simpler and more robust, but necessitates two different propagation directions through the polarization gadgets effecting the transformations U and V. In general, the operations in the two different propagation directions are not the same, limiting the use of this geometry to special cases.

the expense of more cumbersome and low fidelity target qubit operations.

In [17] a common-path geometry was realized in a different way, by encoding the target system in the temporal degree of freedom. This implementation, however, was limited to generalized versions of the Pauli X and Z operators, and could therefore not prepare superposition states. These operations furthermore required both ultrafast phase modulators and the manual replacement of optical components, increasing the experimental requirements while reducing the programmability of the setup.

In this Letter, we overcome these limitations by designing a fully reciprocal polarization gadget capable of realizing any $U \in \mathrm{SU}(2)$, thereby enabling the use of a passively stable Sagnac geometry with perfect spatial mode overlap, while still using the polarization d.o.f. for the target qubit, without imposing any restrictions on the unitaries applied on this system inside the quantum switch.

Photonic quantum switch.—In the commonly used path-polarization quantum switch, shown in Fig. 1(a), the path d.o.f. is used to coherently superpose two different orders through the wave plate gadgets that act on the target (polarization) qubit. These wave plate gadgets, first introduced by Simon and Mukunda, consist of two quarter-wave plates and one half-wave plate, and are universal for SU(2) [18]. If one tries to use a common path geometry, as depicted in Fig. 1(b), the photon now travels through the polarization gadgets in two different directions. The action of a Simon-Mukunda gadget in the backwards direction is $U^{\mathrm{bw}} = PU^T P^\dagger \neq U$, where U is the unitary operation in the forwards direction, and P is a unitary operator describing the basis change to the backwards propagation direction. By picking a convention for the polarization states in which the diagonal polarizations are associated with the eigenstates of the Pauli Y matrix one finds that $P = \mathbb{1}$ [19], however, the residual transpose, a consequence of the fact that the order of the wave plates is transposed in the backwards direction, cannot be undone this way. The Simon-Mukunda gadget is therefore only reciprocal for a two parameter subset of SU(2), and common-path quantum switches such as [20], were thus far not able to implement arbitrary polarization unitaries.

In this Letter, we adopt the convention $(S_1, S_2, S_3) \leftrightarrow (X, Y, Z)$ for the Stokes parameters and Pauli matrices. Written in this convention, the Simon-Mukunda gadget transforms as $U \mapsto U^{\mathrm{bw}} = ZU^T Z$ under counterpropagation. Given a unitary parametrized as $U = \exp[-i(\theta/2)\vec{\sigma}\cdot\vec{n}]$, where $\vec{\sigma}$ is the Pauli vector and $\vec{n}$ the rotation axis of the unitary operation on the Bloch sphere, this transformation corresponds to $[\theta, n_x, n_y, n_z] \mapsto [\theta, -n_x, n_y, n_z]$. Since this transformation applies to any unitary operation implemented by a sequence of linear retarders under counterpropagation, we also consider circular retarders, more specifically Faraday rotators. It is a well known fact that Faraday rotators are nonreciprocal, due to the magneto-optic effect breaking Lorentz reciprocity [21]. This property has enabled a multitude of widely adopted optical devices such as Faraday mirrors [22], optical circulators, and optical isolators [23]. Quantitatively, the nonreciprocity manifests itself as the following transformation under counterpropagation $[\theta, n_y] \mapsto [-\theta, n_y]$.

Faraday rotators are usually sold with a fixed circular retardance of $\theta = (\pi/2)$, corresponding to a rotation of linear polarization by 45°. We will therefore restrict our discussion to only these devices, and show how they can be used to construct a fully reciprocal polarization device.

A reciprocal polarization gadget.—Since only the X component of the Simon-Mukunda gadget exhibits nonreciprocity, we begin by constructing a gadget capable of realizing a reciprocal X rotation:

$$G_x(\theta) = H\left(\frac{\pi}{8}\right) F_- Q\left(\frac{\pi}{2}\right) H\left(\frac{\theta + 2\pi}{4}\right) Q\left(\frac{\pi}{2}\right) F_+ H\left(\frac{\pi}{8}\right). \tag{2}$$

Here H and Q refer to half- and quarter-wave plates at a given angle from the vertical axis, and $F_\pm$ are Faraday rotators with circular retardance of $\pm(\pi/2)$. The three middle wave plates constitute a Simon-Mukunda gadget implementing a nonreciprocal X rotation:

$$R_x(\theta + 2\pi) = -R_x(\theta) \mapsto -R_x(-\theta). \tag{3}$$

The action of the G_x gadget in the two different propagation directions can therefore be expressed as

$$G_x^{\mathrm{fw}}(\theta) = -H\left(\frac{\pi}{8}\right) F_- R_x(\theta) F_+ H\left(\frac{\pi}{8}\right) \tag{4}$$

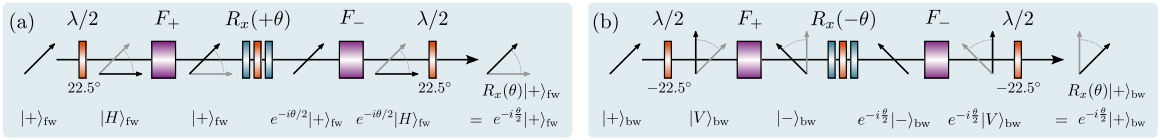


FIG. 2. Reciprocal polarization gadget. The evolution of a diagonally polarized state in the two different propagation directions through a reciprocal R_x gadget. The arrows indicating the polarization states are drawn in the comoving frame. Note that only differential phases between the two polarization components are indicated. (a) The Faraday rotators $F_\pm$ rotate linear polarization by 45°, and in the forwards direction these rotations cancel the half-wave plate (HWP) rotations, indicated by $\lambda/2$. The state before the central wave plate gadget is therefore $|+\rangle$, and the correct phase gets applied. (b) In the backwards propagation direction the rotations of the HWP and Faraday rotators add, rotating the linear polarization by 90°. The $R_x(-\theta)$ rotation applies the desired phase to what is now the $|-\rangle$ state, which subsequently gets rotated back to $|+\rangle$, yielding the correct transformation.

$$G_x^{\mathrm{bw}}(\theta) = -H\left(-\frac{\pi}{8}\right)F_- R_x(-\theta)F_+ H\left(-\frac{\pi}{8}\right), \tag{5}$$

since for a single linear retarder at an angle φ to the vertical axis the effect of reversing the propagation direction is $\varphi \mapsto -\varphi$. The superscripts fw and bw refer to the forwards and backwards propagation directions, respectively. To see that the full gadget is reciprocal, note that

$$H\left(\frac{\pi}{8}\right)F_- = F_+ H\left(\frac{\pi}{8}\right) = -iX, \tag{6}$$

$$H\left(-\frac{\pi}{8}\right)F_- = F_+ H\left(-\frac{\pi}{8}\right) = -iZ, \tag{7}$$

as shown in the Supplemental Material [24]. The action of the gadget in the two propagation directions can therefore be simplified to

$$G_x^{\mathrm{fw}}(\theta) = XR_x(\theta)X = R_x(\theta), \tag{8}$$

$$G_x^{\mathrm{bw}}(\theta) = ZR_x(-\theta)Z = R_x(\theta). \tag{9}$$

A graphical examination of the reciprocity of the gadget is shown in Fig. 2. Using this gadget as a building block, it becomes possible to construct a fully reciprocal gadget capable of implementing arbitrary unitaries:

$$G_R = Q(\theta)H(\phi)G_x(\gamma)H(-\phi)Q(-\theta). \tag{10}$$

This gadget is reciprocal due to its palindromic order, and a proof of its universality, as well as a method to find the angles θ, ϕ, and γ for a given U, is provided in the Supplemental Material [24]. An implementation of this algorithm is available in an open repository [25].

Advantage in a channel discrimination task.—To certify that our experimental platform is capable of realizing an indefinite causal order, we now present a channel discrimination task for which the quantum switch strictly outperforms any causally ordered strategy. This channel discrimination problem was originally presented as a causal witness in Ref. [15] and was inspired by the task introduced in Ref. [3]. Let U_i, V_j be two qubit unitary operators belonging to the set

$$\mathcal{G} \coloneqq \left\{\mathbb{1}, X, Y, Z, \frac{X+Y}{\sqrt{2}}, \frac{X+Z}{\sqrt{2}}, \frac{Y+Z}{\sqrt{2}}\right\}. \tag{11}$$

Using this set, we define two sets of pairs of operators (U_i, V_j) that either commute or anticommute:

$$\mathcal{G}_\pm \coloneqq \{(U_i, V_j) | U_i, V_j \in \mathcal{G}, U_i V_j = \pm V_j U_i\}. \tag{12}$$

Let (U_i, V_j) be a pair of channels belonging to either $\mathcal{G}_+$ or $\mathcal{G}_-$, and consider the task of deciding to which set they belong, given only a single use of the channels. It is well known that this task can be performed deterministically when given access to the quantum switch. This can be seen by setting the state of the control qubit in (1) to $|+\rangle_C$ and considering the action on any target state $|\Psi\rangle_T$:

$$\begin{aligned}&(UV \otimes |0\rangle\langle 0| + VU \otimes |1\rangle\langle 1|)|\Psi\rangle_T \otimes |+\rangle_C \\ &= \frac{1}{2}(UV + VU)|\Psi\rangle_T \otimes |+\rangle_C + \frac{1}{2}(UV - VU)|\Psi\rangle_T \otimes |-\rangle_C.\end{aligned} \tag{13}$$

A measurement of the control qubit then reveals to which set (U_i, V_j) belongs.

For this discrimination task, the probability of successfully guessing the set can be expressed as

$$p_{\mathrm{s}}(i,j) \coloneqq p[\pm|(U_i, V_j)], \quad \text{if } (U_i, V_j) \in \mathcal{G}_\pm. \tag{14}$$

By making use of the semidefinite programming methods presented in Ref. [26], we find that any causally ordered strategy necessarily obeys $\min[p_{\mathrm{s}}(i,j)] \leq 0.841$. Moreover, if the pairs of channels (U_i, U_j) are uniformly picked from $\mathcal{G}_+$ and $\mathcal{G}_-$, the average probability of correctly guessing the set with a causally ordered strategy is bounded by $(1/N)\sum_{i,j} p_{\mathrm{s}}(i,j) \leq 0.904$ where the indices i, j run over all the pairs of gates that commute or anticommute. As discussed in Ref. [15], this average success probability

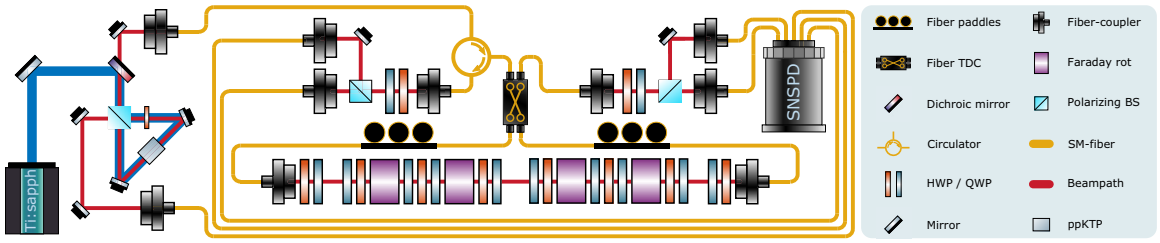


FIG. 3. Experimental setup. Single photons are generated by a type-II spontaneous parametric down-conversion source using a ppKTP crystal. Detection of the signal photon using superconducting nanowire single-photon detectors (SNSPDs) heralds the presence of the idler photon. A tunable directional coupler (TDC) configured for a balanced splitting ratio sends the idler photon through a free-space path in a superposition of two propagation directions. This path contains two reciprocal polarization gadgets consisting of Faraday rotators, quarter-wave plates (QWPs) and half-wave plates (HWPs). These gadgets implement the operators U and V. The photon finally exits in one of the two TDC ports depending on whether U and V commute or anticommute, and is then detected using a polarization resolving measurement. A fiber circulator is used to pick off photons exiting the Sagnac in the input port.

approach can be phrased in terms of a causal nonseparability witness, and in the Supplemental Material [24] we explicitly present such a witness.

Experiment.—Before experimentally performing the channel discrimination task, we first show that the gadget in (10) is indeed reciprocal and universal. To this end we perform quantum process tomography on both gadgets used in the experiment for 100 random unitaries. The resulting gate fidelities, defined as the average state fidelity under the reconstructed unitaries, are presented in the Supplemental Material [24]. Achieving an average gate fidelity of 0.9970 ± 0.0018, and an average fidelity between the two propagation directions of 0.9972 ± 0.0018, we conclude that the gadget is reciprocal and universal.

We now turn to the experimental realization of the quantum switch, pictured in Fig. 3. For the certification of the indefinite causal structure of the implemented process, we employ single photons generated using type-II spontaneous parametric down-conversion [27]. The signal photon is used as a herald for the idler photon, which is made to propagate through the quantum switch. Initially, a beam splitter in the form of a tunable directional coupler (TDC) applies a Hadamard operation on the path (control) d.o.f., thereby preparing the state: $|\Psi\rangle_T \otimes (|0\rangle_C + |1\rangle_C)/\sqrt{2}$, where the subscripts C and T refer to the control and target degrees of freedom, respectively. The photon then passes through the polarization gadgets in a superposition of the two propagation directions, correlating the applied gate order with the control d.o.f.: $(1/\sqrt{2})UV|\Psi\rangle_T \otimes |0\rangle_C + (1/\sqrt{2})VU|\Psi\rangle_T \otimes |1\rangle_C$. Finally, the photon propagates back to the TDC which once again applies a Hadamard gate on the control qubit:

$$\frac{1}{2}\{U,V\}|\Psi\rangle_T \otimes |0\rangle_C + \frac{1}{2}[U,V]|\Psi\rangle_T \otimes |1\rangle_C. \tag{15}$$

Measuring the photon's location then reveals, with unity probability, whether the gates (U,V) commute or anticommute.

Our implementation makes use of a combination of free-space and fiber optics, which is facilitated by the intrinsic phase stability of the common-path geometry. The use of a fiber TDC allows for precise control over the splitting ratio as well as providing perfect spatial mode overlap. These two factors combine to yield a high interferometric visibility in excess of 0.9995. The two inner ports of the TDC are connected to fiber collimators that launch the photons into free space where they propagate through the two polarization gadgets in opposite directions.

A fiber circulator is placed at the input port of the TDC to separate the backwards propagating photons from the input light. Finally, two measurement stations are used to measure the polarization of the photons in either output arm of the Sagnac. The polarization resolving measurements allow for the polarization dependent detection efficiencies in the superconducting nanowire single-photon detectors used in the experiment to be corrected for. The fiber circulator induces a small amount of differential loss in the two interferometer outputs which is also corrected for (see Supplemental Material [24]).

Results.—Each of the 52 pairs of (anti)commuting unitary operators in the sets (12) were implemented six independent times, and for each pair of operations single-photon events were recorded for 60 sec, giving a total measurement time of approximately 5 h, with the only downtime being the time spent rotating the wave plates. The success probabilities were then calculated separately for each run. The results of this are shown in Fig. 4. We find a minimum success probability of $\min[p_s(i,j)] = 0.9895$, and an average success probability of $\langle p_s \rangle = 0.99639 \pm 0.00007$, far exceeding the causally separable bounds of 0.841 and 0.904, respectively. Our observed average success probability can be directly compared with a non-common-path implementation of an analogous channel discrimination task presented in [6], where a success probability of 0.973 was achieved. The observed success probabilities $p_s(i,j)$ for the individual pairs of gates

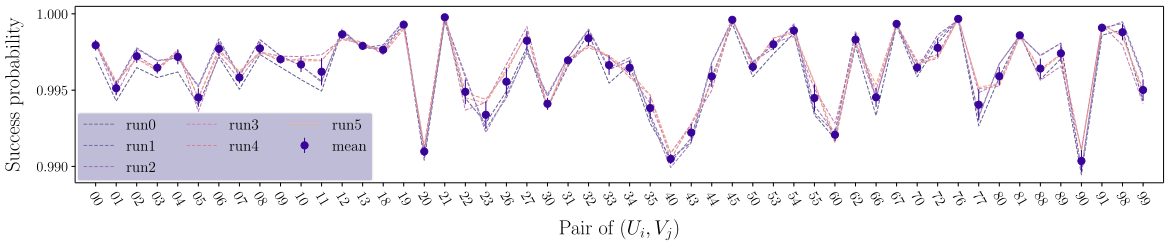


FIG. 4. Success probabilities $p_s(i,j)$ that correspond to the probability for the photon to exit in the correct port of the interferometer given a pair of commuting or anticommuting unitaries (U_i, V_j). The six different runs of the experiment, plotted separately, exhibit high repeatability. Dashed lines show these six different experimental runs, while the solid dots indicate the mean and standard deviation of the runs. The average success probability is $\langle p_s \rangle = 0.9964$, while the highest and lowest success probabilities are $\max[p_s(i,j)] = 0.99997$ and $\min[p_s(i,j)] = 0.9895$, respectively.

(U_i, V_j) additionally display a remarkably low variance, with a recorded standard deviation of $\sigma^2 = 0.0024$, demonstrating the robustness of our design. The uncertainty in the experimentally evaluated success probability is the error-propagated observed standard deviation for the constituent success probabilities $p_s(i,j)$ in the six runs.

Discussion.—We have demonstrated for the first time a path-polarization quantum switch that utilizes a passively stable common-path geometry. Our novel design greatly simplifies the construction and operation of the device, while simultaneously increasing its fidelity, robustness and duty cycle compared to previous demonstrations. The implementation is facilitated by a new polarization gadget that combines different forms of nonreciprocity to unlock fully reciprocal and universal polarization transformations. The methods used to engineer the reciprocity of the polarization gadgets can be applied more broadly to map balanced interferometers onto common-path geometries. This will enable simple and robust bulk-optics realizations of important primitives such as reconfigurable beam splitters and variable partially polarizing beam splitters [28]. We anticipate that this will lead to straightforward realizations of generalized measurements directly on polarization qubits [29], as well as the demonstration of multiparty quantum switches [16].

All data used in this work is openly available at [30].

R. W. P. acknowledges support from the ESQ Discovery program (Erwin Schrödinger Center for Quantum Science and Technology), hosted by the Austrian Academy of Sciences (ÖAW). P. W. acknowledges support from the research platform TURIS, the European Commission through EPIQUS (No. 899368) and AppQInfo (No. 956071), from the Austrian Science Fund (FWF) through BeyondC (F7113) and Research Group 5 (FG5), from the AFOSR via PhoQuGraph (FA8655-20-1-7030) and QTRUST (FA9550-21- 1-0355), from the John Templeton Foundation via the Quantum Information Structure of Spacetime (QISS) project (ID 61466), and from the Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association. T. S. thanks Francesco Massa for helpful discussions.

*Corresponding author.
teodor.stroemberg@univie.ac.at
†Corresponding author.
philip.walther@univie.ac.at

[1] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Phys. Rev. A **88**, 022318 (2013).
[2] T. Colnaghi, G. M. D'Ariano, S. Facchini, and P. Perinotti, Quantum computation with programmable connections between gates, Phys. Lett. A **376**, 2940 (2012).
[3] G. Chiribella, Perfect discrimination of no-signalling channels via quantum superposition of causal structures, Phys. Rev. A **86**, 040301(R) (2012).
[4] A. Feix, M. Araújo, and Č. Brukner, Quantum superposition of the order of parties as a communication resource, Phys. Rev. A **92**, 052326 (2015).
[5] X. Zhao, Y. Yang, and G. Chiribella, Quantum Metrology with Indefinite Causal Order, Phys. Rev. Lett. **124**, 190503 (2020).
[6] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, Experimental superposition of orders of quantum gates, Nat. Commun. **6**, 7913 (2015).
[7] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, Experimental verification of an indefinite causal order, Sci. Adv. **3**, e1602589 (2017).
[8] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, Indefinite Causal Order in a Quantum Switch, Phys. Rev. Lett. **121**, 090503 (2018).
[9] G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M. Zych, Č. Brukner, and P. Walther, Experimental entanglement of temporal order, Quantum **6**, 621 (2022).

[10] K. Goswami, Y. Cao, G. A. Paz-Silva, J. Romero, and A. G. White, Increasing communication capacity via superposition of order, Phys. Rev. Res. **2**, 033292 (2020).
[11] G. Rubino, L. A. Rozema, D. Ebler, H. Kristjánsson, S. Salek, P. Allard Guérin, Č. Brukner, A. A. Abbott, C. Branciard, G. Chiribella, and P. Walther, Experimental quantum communication enhancement by superposing trajectories, Phys. Rev. Res. **3**, 013093 (2021).
[12] Y. Guo, X.-M. Hu, Z.-B. Hou, H. Cao, J.-M. Cui, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, and G. Chiribella, Experimental Transmission of Quantum Information Using a Superposition of Causal Orders, Phys. Rev. Lett. **124**, 030502 (2020).
[13] H. Cao, J. Bavaresco, N.-N. Wang, L. A. Rozema, C. Zhang, Y.-F. Huang, B.-H. Liu, C.-F. Li, G.-C. Guo, and P. Walther, Semi-device-independent certification of indefinite causal order in a photonic quantum switch, Optica **10**, 561 (2023).
[14] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nat. Commun. **3**, 1092 (2012).
[15] M. Araújo, F. Costa, and Č. Brukner, Computational Advantage from Quantum-Controlled Ordering of Gates, Phys. Rev. Lett. **113**, 250402 (2014).
[16] M. J. Renner and Č. Brukner, Computational Advantage from a Quantum Superposition of Qubit Gate Orders, Phys. Rev. Lett. **128**, 230503 (2022).
[17] K. Wei, N. Tischler, S.-R. Zhao, Y.-H. Li, J. M. Arrazola, Y. Liu, W. Zhang, H. Li, L. You, Z. Wang, Y.-A. Chen, B. C. Sanders, Q. Zhang, G. J. Pryde, F. Xu, and J.-W. Pan, Experimental Quantum Switching for Exponentially Superior Quantum Communication Complexity, Phys. Rev. Lett. **122**, 120504 (2019).
[18] R. Simon and N. Mukunda, Minimal three-component SU(2) gadget for polarization optics, Phys. Lett. A **143**, 165 (1990).
[19] T. Strömberg, P. Schiansky, M. T. Quintino, M. Antesberger, L. Rozema, I. Agresti, Č. Brukner, and P. Walther, Experimental superposition of time directions, arXiv:2211.01283.
[20] P. Schiansky, T. Strömberg, D. Trillo, V. Saggio, B. Dive, M. Navascués, and P. Walther, Demonstration of universal time-reversal for qubit processes, Optica **10**, 200 (2023).
[21] V. S. Asadchy, M. S. Mirmoosa, A. Diaz-Rubio, S. Fan, and S. A. Tretyakov, Tutorial on electromagnetic nonreciprocity and its origins, Proc. IEEE **108**, 1684 (2020).
[22] M. Martinelli, A universal compensator for polarization changes induced by birefringence on a retracing beam, Opt. Commun. **72**, 341 (1989).
[23] B. E. Saleh and M. C. Teich, *Fundamentals of Photonics* (John Wiley & Sons, New York, 2019).
[24] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.131.060803 for technical details, universality proof and characterisation data.
[25] M. T. Quintino, https://github.com/mtcq (2022).
[26] J. Bavaresco, M. Murao, and M. T. Quintino, Strict Hierarchy between Parallel, Sequential, and Indefinite-Causal-Order Strategies for Channel Discrimination, Phys. Rev. Lett. **127**, 200504 (2021).
[27] C. Greganti, P. Schiansky, I. A. Calafell, L. M. Procopio, L. A. Rozema, and P. Walther, Tuning single-photon sources for telecom multi-photon experiments, Opt. Express **26**, 3286 (2018).
[28] J. Flórez, N. J. Carlson, C. H. Nacke, L. Giner, and J. S. Lundeen, A variable partially polarizing beam splitter, Rev. Sci. Instrum. **89**, 023108 (2018).
[29] P. Kurzyński and A. Wójcik, Quantum Walk as a Generalized Measuring Device, Phys. Rev. Lett. **110**, 200404 (2013).
[30] T. Strömberg *et al.*, Data for Demonstration of a Quantum Switch in a Sagnac Configuration (2023), 10.5281/zenodo.8171117.

Supplementary Material for 'Demonstration of a quantum SWITCH in a Sagnac configuration'

Teodor Strömberg,[1,2,*] Peter Schiansky,[1,2] Robert W. Peterson,[2] Marco Túlio Quintino,[3,4,5] and Philip Walther[2,†]

[1] *University of Vienna, Faculty of Physics & Vienna Doctoral School in Physics, Boltzmanngasse 5, A-1090 Vienna, Austria*

[2] *University of Vienna, Faculty of Physics & Research Network Quantum Aspects of Space Time (TURIS), Boltzmanngasse 5, 1090 Vienna, Austria*

[3] *Sorbonne Université, CNRS, LIP6, F-75005 Paris, France*

[4] *University of Vienna, Faculty of Physics, Boltzmanngasse 5, 1090 Vienna, Austria*

[5] *Institute for Quantum Optics and Quantum Information, Boltzmanngasse 3, 1090 Vienna, Austria*

I. EXPERIMENTAL DETAILS

The single photons in the experiment were generated using spontaneous parametric down-conversion in a periodically poled $KTiOPO_4$-crystal phase matched for a type-II collinear process. This crystal was pumped by a pulsed Ti:Sapphire laser (Coherent Mira 900HP), with a pulse repetition rate of 76 MHz and tuned to a wavelength of $\lambda_p = 773$ nm, thereby generating degenerate photon pairs at $\lambda_s = \lambda_i = 1546$ nm. The two photons were separated using a polarizing beam-splitter, and the signal photon was sent directly to a superconducting nanowire single-photon detector (SNSPD) from PhotonSpot, housed in a 1 K cryostat. These single-photon detectors had a detection efficiency of around 95 %, and were separated from the experimental setup by 100 m of optical fiber. The idler photon was sent to the experimental setup through approximately 10 m of single-mode fiber, and was then injected into the tunable directional coupler (TDC) using a fiber-optic circulator. This circulator contributed approximately 1 dB of optical loss per pass (2 dB total).

Two 5 m fibers spooled inside fiber polarization controllers were used to connect the output of the TDC to the fiber couplers in the centre of the Sagnac. The free-space optical path loop was approximately 80 cm long, in order to have sufficient room for all the polarization optics and leaving enough space to fit polarizers between the elements for characterisation measurements. Due to the small 5 mm aperture of the Faraday rotators in the polarization gadgets, fiber collimators producing a small beam-diamter (Thorlabs PAF2A-7C) were used to ensure the spatial profile of the photons was not clipped by the polarization elements. The single-mode coupling efficiency in the free-space part was in excess of 85 %. In order to reduce backreflections in the interferometer, the fibers connected to the fiber collimators used anti-reflection coated APC connectors. The output ports of the TDC in the backwards direction were connected to two polarization measurement stations, and the photons in the TDC output port overlapping with the input port were separated by the fiber circulator. The two different polarization components of the light were separated using polarizing beam-splitters and coupled into different single-mode fibers, connected to a total of four SNSPDs. This was done in order to account for the polarization dependent detection efficiencies in these detectors.

Before performing the measurements described in the main text, polarization compensation was first performed on the fibers inside the Sagnac. This was done by injecting H (+)-polarized CW light in the input port of the interferometer, and using the fiber paddles and quarter- / half-wave plates next to the fiber collimators to minimize the transmission through a V (−)-polarizer. Polarization contrasts in excess of 40 dB were achieved for both input polarization states. As a final step, the splitting ratio of the TDC was finetuned. This was done by configuring both polarization gadgets to implement the identity operation, such that destructive interference is observed in one output port of the interferometer. The splitting ratio of the TDC was optimized by minimizing the optical power in this dark port.

* Corresponding author: teodor.stroemberg@univie.ac.at
† Corresponding author: philip.walther@univie.ac.at

II. DEFINITIONS AND CONVENTIONS

In this work we use the following convention for our polarization states:

$$\begin{aligned} |H\rangle &= \begin{bmatrix}1\\0\end{bmatrix}, & |V\rangle &= \begin{bmatrix}0\\1\end{bmatrix}, \\ |+\rangle &= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, & |-\rangle &= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}, \\ |L\rangle &= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\i\end{bmatrix}, & |R\rangle &= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix}. \end{aligned} \tag{1}$$

Under this convention, quarter-wave and half-wave plates are defined as:

$$Q(\theta) = R_y(2\theta)R_z(\pi/2)R_y(-2\theta) \tag{2}$$

$$H(\theta) = R_y(2\theta)R_z(\pi)R_y(-2\theta), \tag{3}$$

where

$$R_k(\theta) = \exp\left[-i\frac{\theta}{2}\sigma_k\right] = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}\sigma_k. \tag{4}$$

Similarly, the fixed Faraday rotators are expressed as:

$$F_\pm = R_y(\pm\pi/2). \tag{5}$$

III. GADGET DERIVATION

Having established these definitions, we explicitly show the simplification used in the derivation of the reciprocal gadget in the main text:

$$\begin{aligned} H\Big(\frac{\pi}{8}\Big)F_- &= R_y\Big(\frac{\pi}{4}\Big)R_z(\pi)R_y\Big(-\frac{\pi}{4}\Big)R_y\Big(-\frac{\pi}{2}\Big) \\ &= R_z(\pi)R_y\Big(-\frac{\pi}{4}\Big)R_y\Big(-\frac{\pi}{4}\Big)R_y\Big(-\frac{\pi}{2}\Big) \\ &= R_z(\pi)R_y(-\pi) \\ &= (-iZ)(iY) \\ &= -iX, \end{aligned} \tag{6}$$

where the first step used:

$$R_y(\theta)R_z(\pi) = R_z(\pi)R_y(-\theta). \tag{7}$$

The other three simplifications follow using the same steps.

IV. UNIVERSALITY OF THE RECIPROCAL GADGET

In this section we will give a proof that the reciprocal gadget presented in the main text is capable of implementing any $U \in \mathrm{SU}(2)$. We first recall the construction of this gadget:

$$\begin{aligned} G_R &= Q(\theta)H(\phi)G_x(\psi)H(-\phi)Q(-\theta) \\ &= Q(\theta)H(\phi)R_x(\psi)H(-\phi)Q(-\theta). \end{aligned} \tag{8}$$

In [1] it was shown that a combination of one half-wave and quarter-wave plate implements a two-parameter subset of SU(2) parameterized as:

$$Q(\theta)H(\phi) = R_y(\alpha)R_z(\pi/2)R_y(\beta). \tag{9}$$

This subset can be equivalently expressed as:

$$R_y(\gamma)R_z(\delta)R_x(\pi/2) \iff R_y(\alpha)R_z(\pi/2)R_y(\beta). \tag{10}$$

Given a two-waveplate gadget with the above parameterisation, the description in the backwards propagation direction is:

$$H(-\phi)Q(-\theta) = R_x(-\pi/2)R_z(\delta)R_y(\gamma). \tag{11}$$

Substituting in these parameterisations in (8) we find:

$$\begin{aligned} G_R &= R_y(\gamma)R_z(\delta)R_x(\pi/2)R_x(\psi)R_x(-\pi/2)R_z(\delta)R_y(\gamma) \\ &= R_y(\gamma)R_z(\delta)R_x(\psi)R_z(\delta)R_y(\gamma). \end{aligned} \tag{12}$$

We now multiply this expression from the left by $R_x(\pi)$ and use the trivial relation $R_x(\pi)R_x(-\pi) = \mathbb{1}$:

$$\begin{aligned} R_x(\pi)G_R &= R_x(\pi)R_y(\gamma)R_x(-\pi)R_x(\pi)R_z(\delta)R_x(\psi)R_z(\delta)R_y(\gamma) \\ &= R_x(\pi)R_y(\gamma)R_x(-\pi)R_x(\pi)R_z(\delta)R_x(-\pi)R_x(\psi+\pi)R_z(\delta)R_y(\gamma) \\ &= R_y(-\gamma)R_z(-\delta)R_x(\psi+\pi)R_z(\delta)R_y(\gamma) \\ &= R_y(-\gamma)R_z(-\delta)R_x(\psi')R_z(\delta)R_y(\gamma), \end{aligned} \tag{13}$$

where $\psi' = \psi + \pi$ and we made use of the identities:

$$R_x(\pi)R_y(\gamma)R_x(-\pi) = R_y(-\gamma) \tag{14}$$

$$R_x(\pi)R_z(\delta)R_x(-\pi) = R_z(-\delta), \tag{15}$$

in the last step. To show that (13) is universal it suffices to show that it can apply a phase $\lambda/2$ to an arbitrary state $|u\rangle$:

$$R_x(\pi)G_R\,|u\rangle = U\,|u\rangle = e^{i\lambda/2}\,|u\rangle\,. \tag{16}$$

To this end, we choose ψ', δ and γ such that $R_x(\psi')R_z(\delta)R_y(\gamma)$ maps $|u\rangle$ to $|-\rangle$ times some phase ϕ:

$$R_x(\psi')R_z(\delta)R_y(\gamma)\,|u\rangle = e^{i\phi}\,|-\rangle = e^{i(\psi'+\mu)/2}\,|-\rangle\,, \tag{17}$$

where $\mu = \phi - \psi'$. This is always possible since $R_x(\psi')R_z(\delta)R_y(\gamma)$ is a Tait-Bryan rotation. It's evident that the mapping $|v\rangle \to e^{i\mu/2}\,|-\rangle$ has to be done by $R_z(\delta)R_y(\gamma)$, since $|-\rangle$ is an eigenstate of $R_x(\psi')$:

$$\begin{aligned} R_z(\delta)R_y(\gamma)\,|u\rangle &= e^{i\mu/2}\,|-\rangle \\ R_x(\psi')\,|-\rangle &= e^{i\psi'/2}\,|-\rangle\,. \end{aligned} \tag{18}$$

We therefore have:

$$\begin{aligned} R_x(\pi)G_R\,|u\rangle &= R_y(-\gamma)R_z(-\delta)R_x(\psi')R_z(\delta)R_y(\gamma)\,|u\rangle \\ &= R_y(-\gamma)R_z(-\delta)e^{i(\psi'+\mu)/2}\,|-\rangle \\ &= e^{i\psi'/2}\,|u\rangle\,. \end{aligned} \tag{19}$$

There is hence always a G_R such that:

$$R_x(\pi)G_R = U, \tag{20}$$

for any $U \in \mathrm{SU}(2)$, and choosing $U = R_x(\pi)V$ for some $V \in \mathrm{SU}(2)$ shows that that G_R is universal.

V. WAVEPLATE ANGLE CALCULATION

In this section we give an explicit method for determining the waveplate angles α, θ and ϕ in the reciprocal gadget given an arbitrary unitary $U \in \mathrm{SU}(2)$:

$$G_R = Q(\theta)H(\phi)H\left(\frac{\pi}{8}\right)F_-Q\left(\frac{\pi}{2}\right)H(\alpha)Q\left(\frac{\pi}{2}\right)F_+H\left(\frac{\pi}{8}\right)H(-\phi)Q(-\theta) = U. \tag{21}$$

A Python implementation of the algorithm can be found in the online repository [2]. The method essentially consists of going through the proof presented in the previous section backwards. First, define a new unitary V:

$$V = R_x(\pi)U, \tag{22}$$

then find the eigenphase $-\lambda/2$ and corresponding eigenvector $|v_+\rangle$:

$$V|v_+\rangle = e^{-i\lambda/2}|v_+\rangle. \tag{23}$$

The angles γ and δ should then be chosen to rotate $|v_+\rangle$ to $|+\rangle$. This can be done by taking:

$$\gamma = \text{arctan2}\Big(\text{tr}\big[Z|v_+\rangle\langle v_+|\big], \text{tr}\big[X|v_+\rangle\langle v_+|\big]\Big) \tag{24}$$

$$\delta = -\text{arctan2}\Big(\text{tr}\big[YR_y(\gamma)|v_+\rangle\langle v_+|R_y^\dagger(\gamma)\big], \text{tr}\big[XR_y(\gamma)|v_+\rangle\langle v_+|R_y^\dagger(\gamma)\big]\Big). \tag{25}$$

The angle ψ is simply given by:

$$\psi = \lambda - \pi. \tag{26}$$

Next, the rotation angles γ and δ need to be mapped to the corresponding waveplate angles θ and ϕ:

$$Q(\theta)H(\phi) = R_y(\gamma)R_z(\delta)R_x(\pi/2). \tag{27}$$

To find these waveplate angles, first construct the state:

$$|L'\rangle = R_x(-\pi/2)R_z(-\delta)R_y(-\gamma)|L\rangle, \tag{28}$$

where:

$$R_x(-\pi/2)R_z(-\delta)R_y(-\gamma) = \big(Q(\theta)H(\phi)\big)^{-1}. \tag{29}$$

The quarter-wave plate angle for the order $H(\phi')Q(\theta') = \pm R_y(\gamma)R_z(\delta)R_x(\pi/2)$ can then be found as:

$$\theta' = \frac{1}{2}\text{arctan2}\Big(\text{tr}\big[X|L'\rangle\langle L'|\big], \text{tr}\big[Z|L'\rangle\langle L'|\big]\Big) + \frac{\pi}{4}. \tag{30}$$

To find the half-wave plate angle, construct the state:

$$|H'\rangle = Q(\theta')R_x(-\pi/2)R_z(-\delta)R_y(-\gamma)|H\rangle. \tag{31}$$

The angle ϕ' is then given as:

$$\phi' = \frac{1}{4}\text{arctan2}\Big(\text{tr}\big[X|H'\rangle\langle H'|\big], \text{tr}\big[Z|H'\rangle\langle H'|\big]\Big). \tag{32}$$

Note that the ambiguity in the overall sign of the unitary doesn't matter due to palindromic order of the total gadget, since $\pm H(\phi')Q(\theta') = \pm ZQ(-\theta')H(-\phi')Z$, and any minus signs cancel. The angles for the order $Q(\theta)H(\phi)$ are found using the waveplate permutation rule:

$$H(\alpha)Q(\beta) = Q(2\alpha - \beta)H(\alpha), \tag{33}$$

and hence:

$$\phi = \phi' \tag{34}$$

$$\theta = 2\phi - \theta'. \tag{35}$$

The last angle, the one of the middle half-wave plate, can be calculated directly from ψ:

$$\alpha = \psi/4 + \pi/2. \tag{36}$$

Using the waveplate reduction rules [1]:

$$Q(a)H(b)H(c) = Q(a + \pi/2)H(a - b + c - \pi/2) \tag{37}$$

$$H(a)H(b)Q(c) = H(a - b + c - \pi/2)Q(c + \pi/2), \tag{38}$$

the gadget can be simplified, removing two waveplates:

$$G_R = Q(\theta_1)H(\phi_1)F_-Q\Big(\frac{\pi}{2}\Big)H(\alpha)Q\Big(\frac{\pi}{2}\Big)F_+H(\phi_2)Q(\theta_2) = U, \tag{39}$$

with:

$$\theta_1 = \theta + \pi/2 \tag{40}$$
$$\phi_1 = \theta - \phi + \pi/8 - \pi/2 \tag{41}$$
$$\theta_2 = -\theta + \pi/2 \tag{42}$$
$$\phi_2 = \pi/8 + \phi - \theta - \pi/2. \tag{43}$$

VI. DEFINITIONS OF $G^{[,]}$, $G^{\{,\}}$

The commuting and anti-commuting subsets $\mathcal{G}_+$, $\mathcal{G}_-$ of

$$\mathcal{G} = \left\{\mathbb{1}, X, Y, Z, \frac{X+Y}{\sqrt{2}}, \frac{X-Y}{\sqrt{2}}\frac{X+Z}{\sqrt{2}}, \frac{X-Z}{\sqrt{2}}\frac{Y+Z}{\sqrt{2}}, \frac{Y-Z}{\sqrt{2}}\right\} \tag{44}$$

are

$$\begin{aligned}\mathcal{G}_+ &= \{U_i, V_j \in \mathcal{G} | [U_i, V_j] = 0\} \\ &= \{U_i, V_j \in \mathcal{G} | i = 0 \vee i = j \vee j = 0\}\end{aligned} \tag{45}$$

and

$$\begin{aligned}\mathcal{G}_- &= \Big\{U_i, V_j \in \mathcal{G} | \{U_i, V_j\} = 0\Big\} \\ &= \Big\{U_i, V_j \in \mathcal{G} | i = 1, j \in \{2,3,8,9\} \vee \\ &\qquad\qquad i = 2, j \in \{1,3,6,7\} \vee \\ &\qquad\qquad i = 3, j \in \{1,2,4,5\} \vee \\ &\qquad\qquad i = 4, j \in \{3,5\} \vee \\ &\qquad\qquad i = 5, j \in \{3,4\} \vee \\ &\qquad\qquad i = 6, j \in \{2,7\} \vee \\ &\qquad\qquad i = 7, j \in \{2,6\} \vee \\ &\qquad\qquad i = 8, j \in \{1,9\} \vee \\ &\qquad\qquad i = 9, j \in \{1,8\}\Big\}.\end{aligned} \tag{46}$$

VII. SUPPLEMENTARY DATA

In this section we present alternative visualisations of the data presented in the main text, as well as some supplementary data used to generate the main result.

Fig. 1 shows the expectation values for each pair of unitaries in the witness averaged over all runs.

Fig. 2 shows the heralding efficiency in the $|\pm\rangle$ ports of the control qubit. Since the total photon number should be conserved, a linear fit to this data gives the relative detection efficiencies in the two ports. This value was in turn used in the evaluation of the success probabilities.. Fig. 3 shows a histogram of the winning probabilities with a bin size of 0.001. This data includes all the settings for the six different runs. It can be seen that in the majority of rounds the winning probability exceeds 0.996.

To verify that the gadgets can faithfully implement any unitary, we implement 100 random unitaries W_i, use quantum state tomography to determine the states $W_i^\delta|\Psi\rangle$ for $|\Psi\rangle \in \{|\mathrm{H}\rangle, |+\rangle\}$ and $\delta \in \{\mathrm{fw}, \mathrm{bw}\}$, and calculate the quantum state fidelities to the expected states. We define the gate fidelity as the average over Ψ, and show the resulting fidelities as histograms in Fig. 4. The mean gate fidelity over all gadgets and directions of 0.9970 ± 0.0018 indicates that the gadgets are indeed capable of implementing arbitrary unitaries.

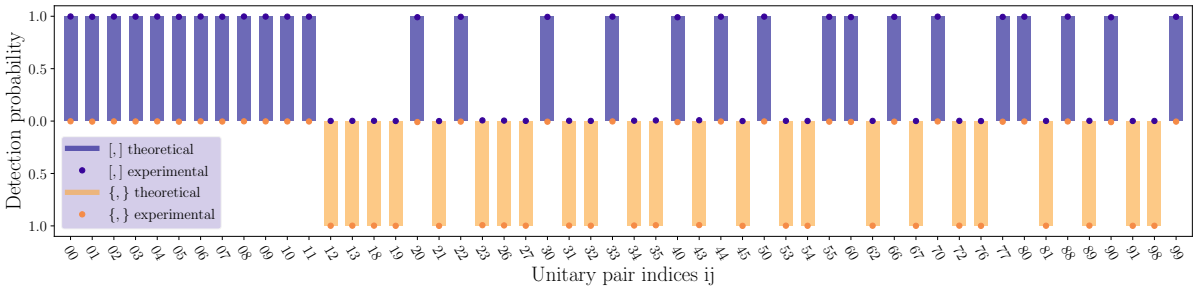


FIG. 1. **Average winning probabilities.** The figure shows the relative probability of a photon to be detected in the commutator and anti-commutator ports of the quantum SWITCH, for every pair of (U, V) in the sets $G^{[,]}$ and $G^{\{,\}}$. The theoretical probabilities $p \in \{0, 1\}$ are shown as solid bars, and the experimentally recorded ones, averaged over all six runs, are indicated by the colored dots. The indices (i, j) on the x-axis specify the pair of unitary operators $(U, V) = (\mathcal{G}_i, \mathcal{G}_j)$.

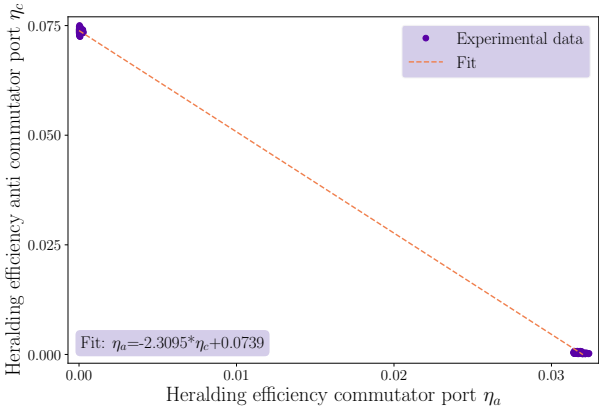


FIG. 2. **Relative heralding efficiencies**. Imbalanced detection efficiencies, as well as loss, between the two output modes of the quantum SWITCH may influence the calculated winning probabilities. As photons must exit the SWITCH in one of either ports, the observed relative heralding efficiencies of unitaries that commute to various degrees follow a linear slope.

Finally, to check that the gadgets are reciprocal, we use the very same measurements, but now calculate the fidelities between $W_i^{\text{fw}} |\Psi\rangle$ and $W_i^{\text{bw}} |\Psi\rangle$. Similarly to before, we define the gadget reciprocity as an average over Ψ, and show the results in Fig. 5. With a mean reciprocity of 0.9972 ± 0.0018, it can be concluded that unitaries implemented by the gadgets are reciprocal. We would like to point out that all unitaries were implemented independently for each direction, hence the reciprocity is affected by imperfect repeatability of the rotation motors moving the gadgets' waveplates.

In order to remove the influence of unwanted fiber polarization rotations on the tomography, it was not performed using the same polarization measurement stations as in the actual experiment. Instead, the tomography was carried out inside the Sagnac interferometer itself. In one propagation direction this was facilitated by using part of one polarization gadget to set the measurement basis for the tomography on the other gadget, and in the opposite propagation direction two additional motorized waveplates were introduced into the setup. A sketch of the tomography setup is shown in Fig. 6.

VIII. CAUSAL WITNESS

To characterise the causal structure of quantum processes one can make use of the process matrix formalism [3], in which a quantum process is represented by a positive semidefinite matrix W. The set of all process matrices that represent a definite causal structure, also called causally separable processes matrices, form a convex subset of all process matrices. Consequently, it is always possible to find a hyperplane separating any causally indefinite process matrix from the set of causally separable ones [4]. This in turn implies the existence of a witness operator $\mathcal{S}$, that can

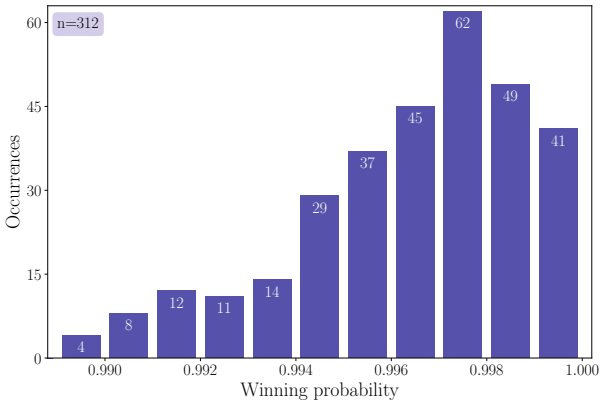


FIG. 3. Winning probability histogram (3 runs: n=312)

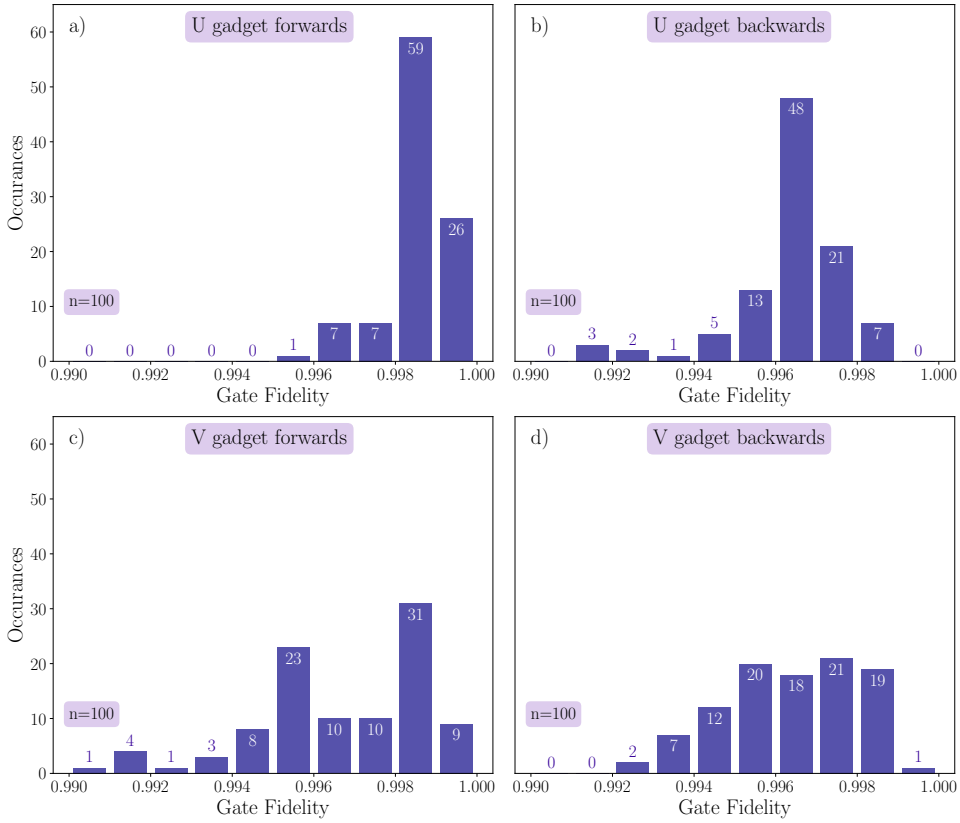


FIG. 4. Gate fidelity histograms for both gadgets in the forwards and backwards directions. 100 random unitaries W_i were implemented with each gadget, the states W_i |H⟩ and W_i $|+\rangle$ were measured in both directions independently and the quantum state fidelity between the expected and measured state determined. Depicted are the average fidelities for $|H\rangle$ and $|+\rangle$ for **a)** U_{fw} (Mean fidelity 0.99852 ± 0.00079). **b)** U_{bw} (0.9964 ± 0.0014). **c)** V_{fw} (0.9967 ± 0.0021). **d)** V_{bw} (0.9964 ± 0.0016). The mean fidelity over all gadgets and directions is 0.9970 ± 0.0018. All uncertainties are standard deviations.

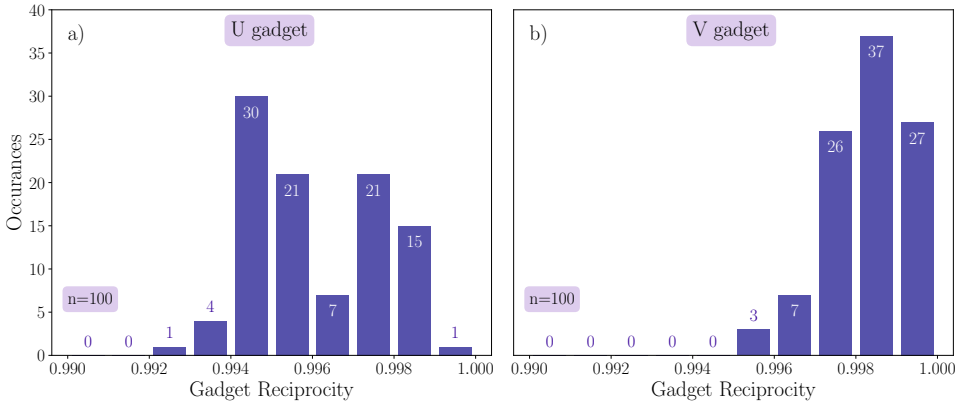


FIG. 5. Reciprocity histograms for the two gadgets used in the experiment. 100 random unitaries W_i were implemented on both gadgets, the states $W_i|\mathrm{H}\rangle$ and $W_i|+\rangle$ were measured in both directions independently and the quantum state fidelity between the expected and measured state determined. Depicted are the average fidelities for **a)** the U gadget, with mean fidelity of 0.9960 ± 0.0016 and **b)** the V gadget, with a mean fidelity of 0.99834 ± 0.00099. The mean fidelity for both gadgets is 0.9972 ± 0.0018. All uncertainties are standard deviations.

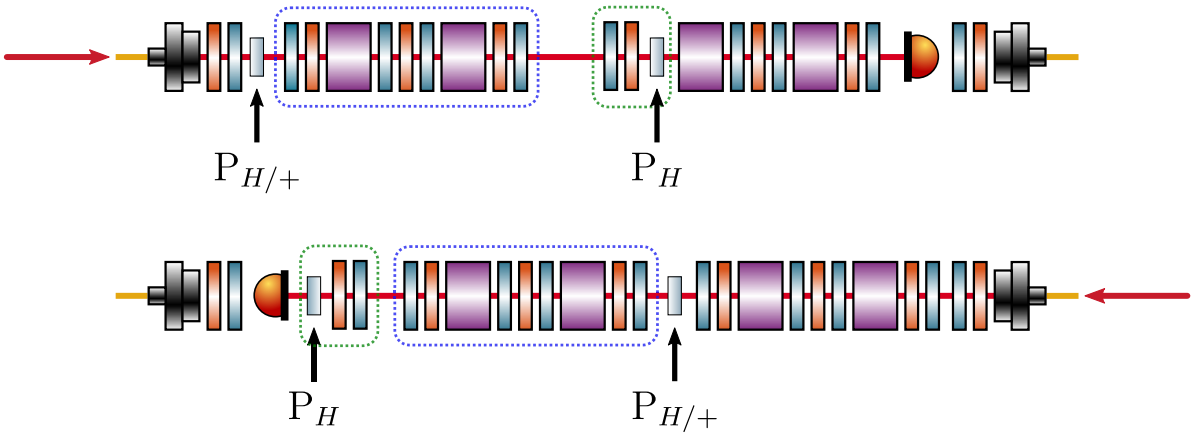


FIG. 6. **Tomography setup.** The tomography on the polarization unitaries was performed inside the Sagnac interferometer in order to avoid introducing measurement errors caused by imperfect polarization compensation in optical fibers. In the forward direction (top) the tomography on the first gadget was performed using the first two waveplates of the second gadget. A polarizer ($\mathrm{P}_{H/+}$) preparing the input polarization H / $+$ was placed before the gadget, and a horizontally aligned measurement polarizer (P_{H}) was placed inside the second gadget. In the backwards propagation direction (bottom) the state preparation polarizer was placed between the two gadgets, and two additional motorized waveplates were introduced into the setup to perform the tomography. The gadget being measured is circled in blue, while the measurement apparatus is circled in green. The tomography on the second gadget in forward (backward) direction was carried out analogously to the one on the first gadget in backward (forward) direction.

be used to certify the indefinite causal structure of a process. Such a witness has previously been used experimentally to validate the casual non-separability of the quantum SWITCH [5]. Here, we adapt a version of the witness from [4], which is inspired by the task presented in Ref. [6]. Making use of the Choi-Jamiołkowski ismorphism, which allows us to represent linear maps and quantum channels as matrices, the witness can be defined in terms of the operators:

$$G_{\pm}^{i,j} = |U_i\rangle\!\rangle\langle\!\langle U_i| \otimes |V_j\rangle\!\rangle\langle\!\langle V_j| \otimes |\pm\rangle\langle\pm|_C \,, \tag{47}$$

with:

$$U_i, V_j \in \mathcal{G} = \left\{ \mathbb{1}, X, Y, Z, \frac{X \pm Y}{\sqrt{2}}, \frac{X \pm Z}{\sqrt{2}}, \frac{Y \pm Z}{\sqrt{2}} \right\}. \tag{48}$$

We then define

$$\mathcal{G}_+ := \left\{ (U_i, V_j) \,\middle|\, U_i, V_j \in \mathcal{G},\ U_i U_j = V_j U_i \right\} \tag{49}$$

$$\mathcal{G}_- := \left\{ (U_i, V_j) \,\middle|\, U_i, V_j \in \mathcal{G},\ U_i V_j = -V_j U_i \right\} \tag{50}$$

The witness itself is given by:

$$\mathcal{S} = \frac{1}{N} \sum_{i,j}^{10} q_{ij}^{[,]} G_+^{i,j} + q_{ij}^{\{,\}} G_-^{i,j}, \tag{51}$$

where $q_{ij}^{[,]}$ and $q_{ij}^{\{,\}}$ are weights chosen such that:

$$\begin{aligned} \{(i,j) : [U_i, V_j] = 0\} : g_{ij}^{[,]} = 1 \\ \{(i,j) : \{U_i, V_j\} = 0\} : g_{ij}^{\{,\}} = 1, \end{aligned} \tag{52}$$

and are zero otherwise. Here, similarly to the main text, $N = 52$ is the total number of commuting or anti-commuting pairs of unitaries in $\mathcal{G}$, and the coefficients above select exactly these subsets. The expectation value of the witness is evaluated as $\langle \mathcal{S} \rangle = \mathrm{tr}[\mathcal{S}W]$, where W is a process matrix [3]. The causal separability bound for the witness described above can be evaluated numerically using the semidefinite programming methods of Ref. [4]. Additionally, the methods from [7] allow us to obtain a computer assisted proof that $\mathrm{tr}[\mathcal{S}W_{\mathrm{sep}}] \leq \frac{90.4}{100}$, for any causally separable process W_{sep}; the code to certify this value is openly available in an online repository [2]. It can be shown that $\mathrm{tr}[\mathcal{S}W_{\mathrm{SWITCH}}] = 1$, where W_{SWITCH} is the process matrix of the quantum SWITCH.

[1] R. Simon and N. Mukunda, Physics Letters A **143**, 165 (1990).
[2] M. T. Quintino, https://github.com/mtcq (2022).
[3] O. Oreshkov, F. Costa, and Č. Brukner, Nature Communications **3**, 1092 (2012), arXiv:1105.4464 [quant-ph].
[4] M. Araújo, F. Costa, and Č. Brukner, Phys. Rev. Lett. **113**, 250402 (2014), arXiv:1401.8127 [quant-ph].
[5] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, Science Advances **3**, e1602589 (2017), arXiv:1608.01683 [quant-ph].
[6] G. Chiribella, Phys. Rev. A **86**, 040301 (2012), arXiv:1109.5154 [quant-ph].
[7] J. Bavaresco, M. Murao, and M. T. Quintino, Phys. Rev. Lett. **127**, 200504 (2021), 2011.08300 [quant-ph].

Conclusions

In this thesis I have presented four experiments on photonic quantum information that explore some of the unique aspects of quantum theory.

The first experiment demonstrated a counterfactual communication protocol that circumvents the issue of the weak trace present in previous experimental proposals and demonstrations. Due to the need for several concatenated interferometers, a photonic waveguide was deemed a suitable platform for the experimental realisation. The intrinsic phase stability of integrated photonic structures obviates the need for active stabilization, and therefore allows a large number of interferometers to be concatenated. However, the rigid structure of the device means that the quantum state of the photons cannot be probed during the experiment. This prevents the implementation of weak measurements to experimentally verify the near-vanishing weak trace in Bob's laboratory. A more flexible bulk-optics realisation of a counterfactual communication protocol in which such measurements could be performed would constitute an interesting direction for future work.

The second experiment was the first implementation of a universal time-rewinding protocol for two-level quantum systems. The protocol was validated in a large parameter regime, but one restricted to unitary time evolutions. Since the only requirement on the time evolution imposed by the protocol is that it be invertible, the demonstration of time rewinding for non-unitary processes is worth exploring in future work. Additionally, while the implemented version of the protocol is probabilistic, it admits an adaptive error-correction procedure that can asymptotically bring the success probability to unity. A demonstration of at least one round of this error correction should be within experimental reach, but the issue of how to quantify the success probability of the protocol in a scenario where photon loss is the dominant source of failure remains an open question.

In the third experiment time reversal in quantum mechanics was studied from a different angle; instead of addressing the question of how to effect the time reversal, the potential of superposing two time directions of quantum evolutions was explored. Such superpositions turn out to constitute a computational resource that outperforms all temporally-ordered quantum processes in a specifically tailored task. Here the metric used for the comparison is the number of uses of a given set of quantum gates, and an important open problem is the experimental certification of this resource consumption. At the moment it is furthermore unclear if there are other, possibly more general tasks for which the advantage offered by temporally-indefinite processes is present. Finally, while the experiment exploited the symmetry properties of the experimental devices generating the time evolutions to induce the time-reversed ones, one could achieve the same results by incorporating the methods of Publication 2, though it must be noted that this approach would presently be challenging experimentally.

The fourth and final experiment was a demonstration of a quantum switch in a common-path geometry. The primary goal of the experiment was to show the utility

and applicability of the devices introduced in Chapter 3, and this was achieved by constructing a setup that outperformed similar works by a significant margin while being simpler to operate. The study and development these devices began during and was motivated by the work on Publication 2, which employed a quantum switch as part of the rewinding protocol, and this device initially suffered from phase instabilities. The connection between the different works showcases the value in translating theoretical ideas into lab, as this process often spurs, or even necessitates technical progress. In the case of the polarization gadgets presented in Chapter 3, I believe that they will see use outside the specific context of experiments on indefinite causality, as the ability to map balanced Mach-Zehnder interferometers onto a common-path geometry is a powerful and general tool.

A
Wave-plate angle formulae

Here the expressions for wave-plate angles that appear in Chapter 2 and Chapter 3 will be motivated.

A.1 SimonMukunda gadget

For the Simon–Mukunda gadget there are three wave-plate angles that need to be calculated in order to implement a unitary U. The expressions for these are given by (2.3-11), (2.3-13) and (2.3-19) in the main text. The first angle θ_q should be set such that a state $|\Psi\rangle = U^\dagger |L\rangle$ is mapped to the plane of linear polarization on the Bloch sphere by the quarter-wave plate. To find an expression for this angle it is helpful to consider the reverse process: using a QWP to map a linearly polarized state to some point on the Bloch sphere. This is illustrated in Fig. A.1, which shows how linearly polarized states are transformed by a QWP at an angle θ to the horizontal axis. As shown in Section 3.2.2 the action of a QWP can be decomposed as

$$Q(\theta) = Y(2\theta)Z(\pi/2)Y(-2\theta). \tag{A.1}$$

When acting on linearly polarized states, the first Y-rotation simply maps these states to other linear polarizations. The $Z(\pi/2)$ rotation then maps the linear plane to the $y-z$ plane, and the final $Y(2\theta)$ rotation tilts this plane about the y-axis. Therefore, every linearly polarized state transformed by a given QWP makes the same angle with the $y-z$ plane of the Bloch sphere. As shown in Fig. A.1 this angle is exactly two times the physical angle of the QWP. From this discussion it is clear that a given any state $|\psi\rangle$ on the Bloch sphere there is always a QWP angle such that the image of the linear plane under the QWP contains this state. We can therefore write:

$$|\Psi\rangle = Q(\theta)\,|\Phi\rangle_{\mathrm{L}}\,, \tag{A.2}$$

where $|\Phi\rangle_{\mathrm{L}}$ is a linearly polarized state. Since:

$$Q(90^\circ - \theta)Q(\theta) = \mathbb{1} \tag{A.3}$$

it follows that:

$$Q(90^\circ - \theta)\,|\Psi\rangle = Q(90^\circ - \theta)Q(\theta)\,|\Phi\rangle_{\mathrm{L}} = |\Phi\rangle_{\mathrm{L}}\,, \tag{A.4}$$

is in the linear plane. From Fig. A.1 it's clear that:

$$\tan 2\theta = \frac{\mathrm{Tr}\big[X\,|\Psi\rangle\langle\Psi|\big]}{\mathrm{Tr}\big[Z\,|\Psi\rangle\langle\Psi|\big]}, \tag{A.5}$$

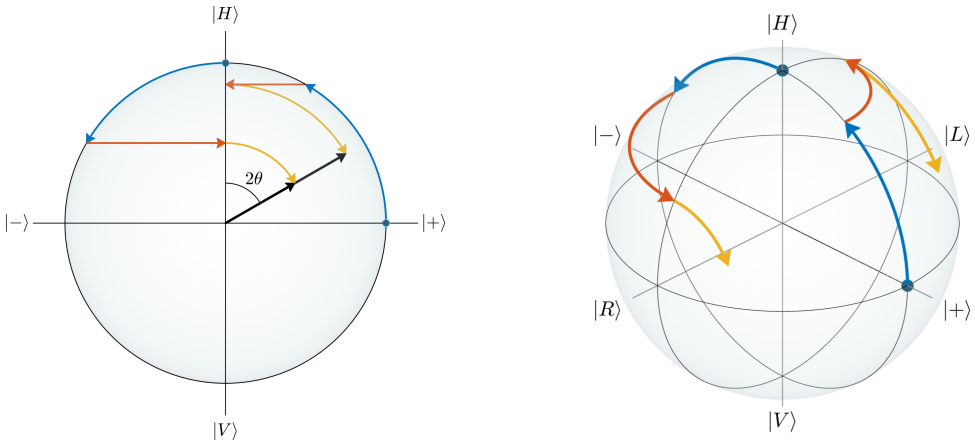


Figure A.1: **QWP angle for map to linear plane.** A quarter-wave plate at an angle θ to the horizontal axis (the positive z-axis in the figure) transforms linearly polarized states such that the projection of the final state in the linear plane is at an angle 2θ relative to the z-axis. This is illustrated for two different linearly polarized states, $|H\rangle$ and $|+\rangle$, by decomposing the QWP rotation as $Q(\theta) = Y(2\theta)Z(\pi/2)Y(-2\theta)$. By considering the inverse of the transformation in the figure one finds that the angle to map a given state to the linear plane using a QWP is given by $90° - \theta$.

since, as illustrated in Fig. 3.11, the quantities $\mathrm{Tr}[\sigma_j \rho]$ correspond to the projection of a the Bloch vector of the state ρ onto the axis j. The expression above can be inverted to find:

$$\theta_q = 90° - \theta = \frac{1}{2}\mathrm{atan2}\big(\mathrm{Tr}\big[X\,|\Psi\rangle\langle\Psi|\big], \mathrm{Tr}\big[Z\,|\Psi\rangle\langle\Psi|\big]\big). \tag{A.6}$$

The second QWP angle, φ_q, should be chosen to map a state in the linear plane to right-handed circular polarization. Let this state be

$$|\Psi_L\rangle = Q(\theta_q)\,|\Psi\rangle\,. \tag{A.7}$$

A QWP is a $\pi/2$ rotation about an axis in the linear plane of the Bloch sphere. Therefore, mapping a linear state to a circularly polarized one can be done by ensuring that the rotation axis of the transformation is orthogonal to the Bloch vector of the initial state. In order for the state to get mapped to $|R\rangle$ the rotation axis and Bloch vector need to have the correct relative handedness. As illustrated in Fig. A.2 the rotation axis defined by the QWP should be rotated by $-90°$ in the $x - z$ plane, with respect to the Bloch vector. It can also be seen that this corresponds to finding the angle with respect to the negative x-axis, instead of the positive one:

$$\tan\varphi = \mathrm{Tr}\left[\frac{Z\,|\Psi_L\rangle\langle\Psi_L|}{-X\,|\Psi_L\rangle\langle\Psi_L|}\right]. \tag{A.8}$$

The wave-plate angle is therefore:

$$\varphi_q = \frac{1}{2}\mathrm{atan2}\big(\mathrm{Tr}\big[Z\,|\Psi_L\rangle\langle\Psi_L|\big], \mathrm{Tr}\big[-X\,|\Psi_L\rangle\langle\Psi_L|\big]\big). \tag{A.9}$$

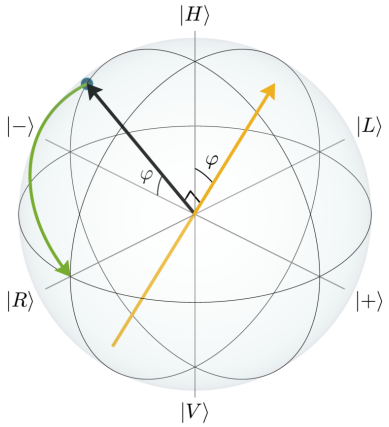


Figure A.2: **QWP angle for map to circular polarization.** In order to map a linearly polarized state $|\Psi_L\rangle$ to $|R\rangle$, the rotation axis (yellow arrow) should be orthogonal to the Bloch vector of the state (black arrow). As can be seen in the figure, the rotation axis should be chosen so that it is rotated by $-90°$ with respect to the Bloch vector, in order to ensure the correct handedness. The opposite choice of rotation axis will lead to the map $|\Psi_L\rangle \mapsto |L\rangle$.

From the definition of the atan2 function it follows that:

$$\text{atan2}(y, x) + \text{atan2}(y, -x) = \pm 180° \tag{A.10}$$

Since the QWP has a period of 180° the sign on the right-hand side does not matter, and the angle can be expressed as

$$\varphi_q = 90° - \frac{1}{2}\text{atan2}\big(\text{Tr}\big[Z\,|\Psi_L\rangle\langle\Psi_L|\big], \text{Tr}\big[X\,|\Psi_L\rangle\langle\Psi_L|\big]\big). \tag{A.11}$$

The equivalent of the relation $\text{atan}(x/y) + \text{atan}(y/x) = 90°$ for atan2 is:

$$\text{mod}\big(\text{atan2}(x, y) + \text{atan2}(y, x), 360°\big) = 90°, \tag{A.12}$$

and using this the QWP angle can finally be expressed as

$$\varphi_q = \frac{1}{2}\text{atan2}\big(\text{Tr}\big[X\,|\Psi_L\rangle\langle\Psi_L|\big], \text{Tr}\big[Z\,|\Psi_L\rangle\langle\Psi_L|\big]\big) + 45°, \tag{A.13}$$

which is the form of the equation that appears in the main text. The final angle γ_h for the HWP should be chosen to rotate a state $|\Phi\rangle = Q(\varphi_q)Q(\theta_q)U^\dagger|H\rangle$ in the linear plane to $|H\rangle$. The equation for this angle is:

$$\gamma_h = \frac{1}{4}\text{atan2}\big(\text{Tr}\big[X\,|\Phi\rangle\langle\Phi|\big], \text{Tr}\big[Z\,|\Phi\rangle\langle\Phi|\big]\big). \tag{A.14}$$

This expression comes from the facts that, as illustrated in Fig. A.3, a HWP at an angle γ to a linearly polarized state always rotates this state by 4γ on the Bloch sphere, and that on the Bloch sphere a HWP is self inverse.

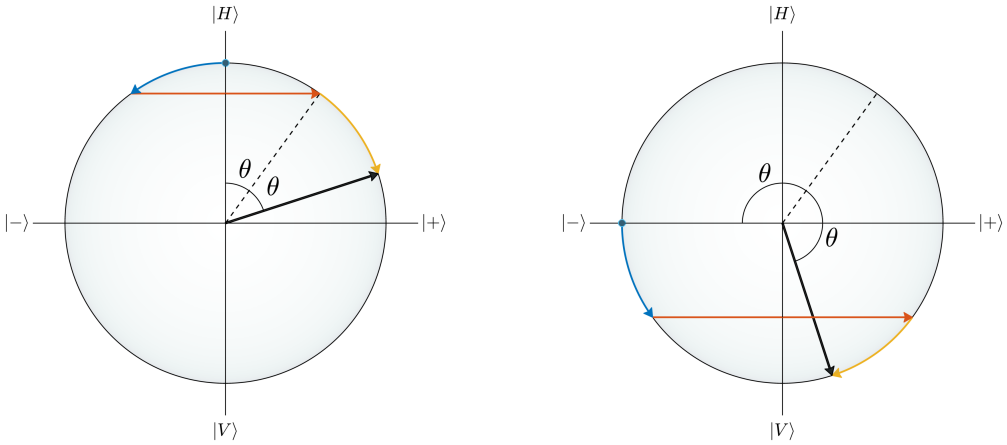


Figure A.3: **HWP angle for rotation in linear plane.** The action of a half-wave plate on a linearly polarized state is to rotate it by 2θ on the Bloch sphere, where $\theta/2$ is the physical angle between the axis of the wave plate and the polarization of the light, and θ is the angle between the Bloch vector of the corresponding state and the axis around which the wave plate rotates the state on the Bloch sphere. This is illustrated above for two different states, $|H\rangle$ and $|-\rangle$. The rotation axis is indicated by a dashed line. Similar to Fig. A.1, the action of the HWP is decomposed as $Y(\theta)Z(\pi)Y(-\theta)$ for illustrative purposes.

A.2 Reciprocal gadget

In the reciprocal gadget there are two wave-plate angles that need to be calculated. The first is a QWP angle to map a linearly polarized state to $|R\rangle$, and the second is a HWP angle to take a linearly polarized state to $|H\rangle$. These angles therefore correspond to (A.13) and (A.14) above, and do not need further explanation.

B
Coherent states in linear optics

At a few points in this thesis calculations using single-photon states are explicitly or implicitly applied to coherent states. For example, the characterisation of polarization gadgets using laser light assumes that these transformations are the same for single photons, and another example is the discussion of active phase stabilization in Chapter 7. Doing this is justified by a formal correspondence between single-photon states and coherent states in linear-optical networks, which will be derived in this Appendix.

B.1 Correspondence with single-photon states

As pointed out in Section 2.2.5, a coherent state can be represented by the displacement operator acting on vacuum:

$$e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}\,|0\rangle = |\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n(\hat{a}^\dagger)^n}{n!}\,|0\rangle\,. \tag{B.1}$$

This was not proven, but can be shown quite straightforwardly using the exponential operator identity

$$e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]}, \tag{B.2}$$

which holds whenever A and B commute with $[A,B]$. In the case of the displacement operator we have

$$[\alpha\hat{a}^\dagger, -\alpha^*\hat{a}] = -\alpha\alpha^*\hat{a}^\dagger\hat{a} + \alpha\alpha^*\hat{a}\hat{a}^\dagger = -|\alpha|^2[\hat{a}^\dagger,\hat{a}] = |\alpha|^2. \tag{B.3}$$

It is then obvious that $\alpha\hat{a}^\dagger$ and $-\alpha^*\hat{a}$ commute with the commutator above. We can therefore use (B.2) to write the displacement operator as

$$e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}} = e^{\alpha\hat{a}^\dagger}e^{-\alpha^*\hat{a}}e^{-\frac{1}{2}|\alpha|^2}. \tag{B.4}$$

It is clear that

$$e^{-\alpha^*\hat{a}}\,|0\rangle = 0, \tag{B.5}$$

and hence

$$e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}\,|0\rangle = e^{-\frac{1}{2}|\alpha|^2}e^{\alpha\hat{a}^\dagger}\,|0\rangle = e^{-\frac{1}{2}|\alpha|^2}\sum_{n}^{\infty}\frac{(\alpha\hat{a}^\dagger)^n}{n!}\,|0\rangle\,. \tag{B.6}$$

To see how coherent states are transformed by linear-optical elements, consider a unitary generated by an m-mode Hamiltonian bilinear in the creation and annihilation operators[1]:

$$U_H = e^{-iH}, \qquad H = \sum_{jk} A_{jk} \hat{a}_j^\dagger \hat{a}_k. \tag{B.7}$$

As shown in Section 2.2.7 this unitary transforms the creation operators as

$$\hat{a}^\dagger \mapsto U_H \hat{a}^\dagger U_H^\dagger = U\hat{a}^\dagger, \tag{B.8}$$

where U is an operator that can be represented by an $m \times m$ matrix. Using the fact that

$$U_H^\dagger \left|0\right\rangle = \left|0\right\rangle, \tag{B.9}$$

We can write the effect of (B.7) on a coherent state in mode k as

$$\begin{aligned} U_H \left|\alpha\right\rangle_k &= U_H e^{-\frac{1}{2}|\alpha|^2} \sum_n^\infty \frac{(\alpha \hat{a}_k^\dagger)^n}{n!} U_H^\dagger \left|0\right\rangle \\ &= e^{-\frac{1}{2}|\alpha|^2} \sum_n^\infty \frac{(\alpha U_H \hat{a}_k^\dagger U_H^\dagger)^n}{n!} \left|0\right\rangle \\ &= e^{-\frac{1}{2}|\alpha|^2} \sum_n^\infty \frac{(\alpha U \hat{a}_k^\dagger)^n}{n!} \left|0\right\rangle \\ &= e^{-\frac{1}{2}|\alpha|^2} e^{U\alpha\hat{a}_k^\dagger} \left|0\right\rangle. \end{aligned} \tag{B.10}$$

Here the second step used the fact that $U^\dagger U = \mathbb{1}$. Let

$$U\hat{a}_k^\dagger = \sum_{j=1}^m u_{jk} \hat{a}_j^\dagger. \tag{B.11}$$

Inserting this into (B.10) gives

$$\begin{aligned} U_H \left|\alpha\right\rangle_k &= e^{-\frac{1}{2}|\alpha|^2} \exp\left[\sum_{j=1}^m u_{jk} \hat{a}_j^\dagger\right] \left|0\right\rangle \\ &= e^{-\frac{1}{2}|\alpha|^2} \prod_{j=1}^m \exp\left[u_{jk} \hat{a}_j^\dagger\right] \left|0\right\rangle. \end{aligned} \tag{B.12}$$

Using the unitarity of U once more, specifically that

$$\left(U^\dagger U\right)_{ik} = \sum_j u_{ji}^* u_{jk} = \delta_{ik} \implies \sum_{j=1} |u_{jk}|^2 = 1 \tag{B.13}$$

[1] As previously stated in Section 2.2.7, all linear-optical transformations are generated by Hamiltonians of this form.

(B.12) can be re-written as

$$\begin{aligned} U_H \left|\alpha\right\rangle_k &= \prod_{j=1}^{m} e^{-\frac{1}{2}|u_{jk}\alpha|^2} \exp\left[u_{jk}\hat{a}_j^\dagger\right] \left|0\right\rangle \\ &= \bigotimes_{j=1}^{m} \left|u_{jk}\alpha\right\rangle_j . \end{aligned} \tag{B.14}$$

From this equation one sees that the amplitude of the coherent state in a given mode after the linear-optical unitary transformation corresponds exactly to the single-photon amplitude in the same mode under the same transformation. This means that the relative intensities $|u_{jk}\alpha|^2/|\alpha|^2$ are equal to the single-photon detection probabilities $|u_{jk}|^2$, which is what we wanted to show. To illustrate this with an example, consider the beamsplitter transformation:

$$U_{\mathrm{BS}} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} . \tag{B.15}$$

For a single-photon state we have that

$$\hat{a}^\dagger \left|0,0\right\rangle_{a,b} \mapsto \cos\theta \hat{a}^\dagger \left|0,0\right\rangle_{a,b} + \sin\theta \hat{b}^\dagger \left|0,0\right\rangle_{a,b} = \cos\theta \left|1,0\right\rangle_{a,b} + \sin\theta \left|0,1\right\rangle_{a,b} , \tag{B.16}$$

and for a coherent state:

$$\left|\alpha,0\right\rangle_{a,b} \mapsto \left|\alpha\cos\theta, \alpha\sin\theta\right\rangle_{a,b} . \tag{B.17}$$

This correspondence is also what justifies the natural, and somewhat obvious interpretation of a matrix such as (B.15) as describing the reflection and transmission coefficients of the beamsplitter for a classical EM-field. Similarly, this is why the Jones matrices of linearly birefringent polarization optics also describe the transformation of single-photon states.

Note that while the correspondence here was shown for an initial one-mode coherent state it also holds for m-mode states and single-photons superposed over these modes. This can be shown in the same way, or alternatively by using the fact that any m-mode coherent state can be written on the form (B.14). Any linear-optical transformation W on this state can be decomposed into a map that contains a unitary $U_H^\dagger$ that first brings the state back to a single-mode state, and then a second unitary that prepares the output state of the total transformation.

C Quantization of the electromagnetic field

This appendix clarifies a mathematical point touched on in the discussion on spontaneous parametric down-conversion in Section 2.4.2 of the main text. It then discusses the problems that arise when using the quantized free electric field in dielectric media, before briefly presenting an alternative approach.

C.1 Box quantization continuum limit

Consider a one-dimensional box with side length L and periodic boundary conditions. The basis states for this box are of the form

$$u_k(x) = \frac{1}{\sqrt{L}} e^{ikx}, \tag{C.1}$$

where

$$k = \frac{2\pi n}{L}, \qquad n \in \mathbb{Z}. \tag{C.2}$$

The basis states are orthonormal and therefore obey

$$\int_{-L/2}^{L/2} u_k^*(x) u_{k'}(x) dx = \frac{1}{L} \int_{-L/2}^{L/2} e^{i(k-k')x} dx = \delta_{k,k'}. \tag{C.3}$$

Multiplying the middle and right-hand terms with L and taking the limit $L \to \infty$ we see that

$$\lim_{L\to\infty} \int_{-L/2}^{L/2} e^{i(k-k')x} dx = 2\pi\delta(k-k') = \lim_{L\to\infty} L\delta_{k,k'}. \tag{C.4}$$

In the continuum limit we therefore have

$$\delta_{k,k'} \to \frac{2\pi}{L}\delta(k-k'). \tag{C.5}$$

A function $f(x)$ on the domain $[-L/2, L/2]$ can be written as a Fourier series [110]:

$$f(x) = \frac{1}{\sqrt{L}} \sum_k f_k e^{ikx}, \tag{C.6}$$

with Fourier coefficients

$$f_k = \frac{1}{\sqrt{L}} \int_{-L/2}^{L/2} f(x) e^{-ikx} dx. \tag{C.7}$$

Taking the continuum limit $L \to \infty$

$$\lim_{L\to\infty} \int_{-L/2}^{L/2} f(x)e^{-ikx}dx = f(k) \tag{C.8}$$

and we find that

$$f_k \to \frac{1}{\sqrt{L}} f(k). \tag{C.9}$$

For the creation and annihilation operators the equivalent substitution

$$\hat{a}_k \to \frac{1}{\sqrt{L}} \hat{a}(k) \tag{C.10}$$

together with (C.5) gives

$$[\hat{a}_k, \hat{a}_{k'}] = \delta_{k,k'} \to [\frac{1}{\sqrt{L}}\hat{a}(k), \frac{1}{\sqrt{L}}\hat{a}(k')] = \frac{2\pi}{L}\delta(k-k'). \tag{C.11}$$

In some texts the creation and annihilation operators are therefore rescaled [222]:

$$\hat{a}_k \to \sqrt{\frac{2\pi}{L}}\hat{a}(k) \tag{C.12}$$

to avoid the factor of 2π in the commutator, giving

$$[\hat{a}_k, \hat{a}_{k'}] = \delta_{k,k'} \to [\sqrt{\frac{2\pi}{L}}\hat{a}(k), \sqrt{\frac{2\pi}{L}}\hat{a}(k')] = \frac{2\pi}{L}\delta(k-k') \tag{C.13}$$

$$\implies [\hat{a}(k), \hat{a}(k')] = \delta(k-k'). \tag{C.14}$$

The equivalent expression in terms of frequency is

$$\hat{a}_\omega \to \sqrt{\frac{2\pi c}{L}}\hat{a}(\omega), \tag{C.15}$$

which is what appears in the main text.

C.2 Transverse delta function

This section will introduce a function, or distribution, known as the transverse delta function that will be used in the next section. As the name suggests, it is a function that acts like the delta function for transverse fields [110]:

$$\mathbf{F}^\perp(\mathbf{r}) = \int d^3r' \delta^\perp(\mathbf{r}-\mathbf{r}')\mathbf{F}(\mathbf{r}'), \tag{C.1}$$

where $\mathbf{F}^\perp(\mathbf{r})$ is the transverse part of the field $\mathbf{F}$:

$$\mathbf{F}(\mathbf{r}) = \mathbf{F}^\perp(\mathbf{r}) + \mathbf{F}^\parallel(\mathbf{r}) \tag{C.2}$$

$$\nabla \cdot \mathbf{F}^\perp(\mathbf{r}) = 0 \tag{C.3}$$

$$\nabla \times \mathbf{F}^\parallel(\mathbf{r}) = 0. \tag{C.4}$$

The decomposition above is known as the Hemholtz decomposition [215]. In reciprocal space $\delta^\perp$ is simply the projector onto the transverse part of the field:

$$\mathbf{F}^\perp(\mathbf{k}) = \mathbf{F}(\mathbf{k}) - \frac{\mathbf{k}\cdot\mathbf{F}(\mathbf{k})}{\|\mathbf{k}\|^2}\mathbf{k} = \delta^\perp(\mathbf{k})\mathbf{F}(\mathbf{k}), \tag{C.5}$$

which can expressed in index notation as

$$F_i^\perp(\mathbf{k}) = \left(\delta_{ij} - \frac{k_i k_j}{k^2}\right)F_j(\mathbf{k}) \tag{C.6}$$

To see that this definition indeed corresponds to a projection onto the transverse component of the field, we take the divergence of its Fourier transform:

$$\begin{aligned}\nabla\cdot\int\frac{d^3k}{(2\pi)^3}\left(\mathbf{F}(\mathbf{k}) - \frac{\mathbf{k}\cdot\mathbf{F}(\mathbf{k})}{\|\mathbf{k}\|^2}\mathbf{k}\right)e^{i\mathbf{k}\cdot\mathbf{r}} &= \int\frac{d^3k}{(2\pi)^3}\mathbf{k}\cdot\left(\mathbf{F}(\mathbf{k}) - \frac{\mathbf{k}\cdot\mathbf{F}(\mathbf{k})}{\|\mathbf{k}\|^2}\mathbf{k}\right)e^{i\mathbf{k}\cdot\mathbf{r}}\\ &= \int\frac{d^3k}{(2\pi)^3}\left(\mathbf{k}\cdot\mathbf{F}(\mathbf{k}) - \mathbf{k}\cdot\mathbf{F}(\mathbf{k})\frac{\mathbf{k}\cdot\mathbf{k}}{\|\mathbf{k}\|^2}\right)e^{i\mathbf{k}\cdot\mathbf{r}}\\ &= 0.\end{aligned} \tag{C.7}$$

Similarly, the fact that $\delta^\perp(\mathbf{r})$ acts as a delta function can be verified by Fourier transforming (C.5) twice:

$$\begin{aligned}F_i^\perp(\mathbf{r}) &= \int\frac{d^3k}{(2\pi)^3}\delta_{ij}^\perp(\mathbf{k})F_j(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}\\ &= \int\frac{d^3k}{(2\pi)^3}\delta_{ij}^\perp(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}\int d^3r' F_j(\mathbf{r}')e^{-i\mathbf{k}\cdot\mathbf{r}'}\\ &= \int d^3r'\int\frac{d^3k}{(2\pi)^3}\delta_{ij}^\perp(\mathbf{k})F_j(\mathbf{r}')e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}.\end{aligned} \tag{C.8}$$

Here the Fourier transformed functions are simply labelled by their argument. Making the identification

$$\delta_{ij}^\perp(\mathbf{r}-\mathbf{r}') = \int\frac{d^3k}{(2\pi)^3}\delta_{ij}^\perp(\mathbf{k})e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \tag{C.9}$$

recovers (C.1). For discrete momenta the transverse delta function can be defined as

$$\Delta_{ij}^\perp(\mathbf{r}) = \frac{1}{V}\sum_{\mathbf{k}}\left(\delta_{ij} - \frac{k_i k_j}{k^2}\right)e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{C.10}$$

This comes from the fact that $\sum_k e^{ikx}$ is the Fourier series of the Dirac comb:

$$\sum_{k=-\infty}^{+\infty} e^{i2\pi kx/L} = L\sum_{n=-\infty}^{+\infty}\delta(x-nL). \tag{C.11}$$

Considering a finite quantization volume where $k_i = 2\pi n_i/L_i$ gives

$$\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} = V\sum_{n_x}\sum_{n_y}\sum_{n_z}\delta(x-n_xL_x)\delta(y-n_yL_y)\delta(z-n_zL_z). \tag{C.12}$$

All the terms for which $|n_i| > 0$ lie outside the quantization volume,[1] and therefore

$$\frac{1}{V}\int_V d^3r' \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\mathbf{F}(\mathbf{r}') = \int_V d^3r' \delta(x-x')\delta(y-y')\delta(z-z')\mathbf{F}(\mathbf{r}') = \mathbf{F}(\mathbf{r}), \quad \text{(C.13)}$$

and hence

$$\frac{1}{V}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} = \Delta(\mathbf{r}-\mathbf{r}'), \quad \text{(C.14)}$$

$$\mathbf{F}^{\perp}(\mathbf{r}) = \int d^3r' \Delta^{\perp}(\mathbf{r}-\mathbf{r}')\mathbf{F}(\mathbf{r}'). \quad \text{(C.15)}$$

C.3 Quantization in nonlinear dielectrics

In Section 2.4.2 the interaction Hamiltonian for SPDC was treated using the field operators obtained by quantizing the free electric field. It has been pointed out that this approach can lead to inconsistencies [216, 217, 447], and in some cases incorrect predictions in experimentally relevant regimes [448]. To illustrate this, consider the field operators

$$\mathbf{A}(\mathbf{r},t) = \sum_{\mathbf{k},\alpha}\sqrt{\frac{\hbar}{2\epsilon_0\omega_k V}}\hat{e}_{\mathbf{k},\alpha}\Big(\hat{a}_{\mathbf{k},\alpha}(t)e^{i\mathbf{k}\cdot\mathbf{r}} + \hat{a}^{\dagger}_{\mathbf{k},\alpha}(t)e^{-i\mathbf{k}\cdot\mathbf{r}}\Big) \quad \text{(C.1)}$$

$$\mathbf{E}(\mathbf{r},t) = i\sum_{\mathbf{k},\alpha}\sqrt{\frac{\hbar\omega_k}{2\epsilon_0 V}}\hat{e}_{\mathbf{k},\alpha}\Big(\hat{a}_{\mathbf{k},\alpha}(t)e^{i\mathbf{k}\cdot\mathbf{r}} - \hat{a}^{\dagger}_{\mathbf{k},\alpha}(t)e^{-i\mathbf{k}\cdot\mathbf{r}}\Big) \quad \text{(C.2)}$$

$$\mathbf{B}(\mathbf{r},t) = i\sum_{\mathbf{k},\alpha}\sqrt{\frac{\hbar}{2\epsilon_0\omega_k V}}\mathbf{k}\times\hat{e}_{\mathbf{k},\alpha}\Big(\hat{a}_{\mathbf{k},\alpha}(t)e^{i\mathbf{k}\cdot\mathbf{r}}\Big) \quad \text{(C.3)}$$

In the Heisenberg picture the time evolution of the magnetic-field operator is given by

$$\frac{\partial\mathbf{B}}{\partial t} = \frac{i}{\hbar}[H,\mathbf{B}]. \quad \text{(C.4)}$$

Suppose that the Hamiltonian generating the time evolution is the linear Hamiltonian of a dielectric:

$$H = \frac{\epsilon_0}{2}\int \big(c^2\mathbf{B}^2 + \mathbf{E}^2 + \chi^{(1)}\mathbf{E}^2\big)d^3r. \quad \text{(C.5)}$$

Since we know that the free Hamiltonian can be directly written on a simple form:

$$H_f = \frac{\epsilon_0}{2}\int \big(c^2\mathbf{B}^2 + \mathbf{E}^2\big)d^3r = \sum_{\mathbf{k},\alpha}\hbar\omega_k\Big(\hat{a}^{\dagger}_{\mathbf{k},\alpha}\hat{a}_{\mathbf{k},\alpha} + \frac{1}{2}\Big), \quad \text{(C.6)}$$

using

$$[\hat{a}^{\dagger}\hat{a},\hat{a}^{\dagger}] = \hat{a}^{\dagger}, \qquad [\hat{a}^{\dagger}\hat{a},\hat{a}] = -\hat{a}, \quad \text{(C.7)}$$

[1] The terms with $n = \pm 1$ are delta-functions that lie on the endpoints of the interval, see Section C.4 for a discussion on this.

it is straightforward to verify that

$$\frac{i}{\hbar}[H_f, \mathbf{B}] = -\nabla \times \mathbf{E}. \tag{C.8}$$

Furthermore, because the magnetic field commutes with itself, only the electric-field term in the Hamiltonian contributes to the commutator. Assuming, for simplicity, an isotropic medium and therefore scalar linear susceptibility, we can obtain the linear dielectric Hamiltonian by simply rescaling this term

$$\mathbf{E}^2 \to (1 + \chi^{(1)})\mathbf{E}^2. \tag{C.9}$$

It therefore follows that

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{i}{\hbar}[H, \mathbf{B}] = -\nabla \times (1 + \chi^{(1)})\mathbf{E}. \tag{C.10}$$

This is similar to what we would expect from Faraday's law (2.4-8):

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}, \tag{C.11}$$

only with the additional $\chi^{(1)}$ term that we could account for by redefining the electric-field operator. However, the situation gets more troublesome if we include a nonlinear term in the Hamiltonian:

$$H = \epsilon_0 \int \Big(\frac{c^2}{2}\mathbf{B}^2 + \frac{1}{2}(\mathbf{E}^2 + \chi^{(1)}\mathbf{E}^2) + \frac{2}{3}\chi^{(2)}\mathbf{E}^3 \Big) d^3r. \tag{C.12}$$

To evaluate the effect this term has on the time evolution it is helpful to calculate the commutator of the electric- and magnetic-field operators. This is most easily done by first finding the commutator between the electric field and the vector potential:

$$\begin{aligned}[E_j(\mathbf{r}, t), A_l(\mathbf{r}', t)] = \frac{i\hbar}{2\epsilon_0 V} \sum_{\mathbf{k},\mathbf{k}'} \sum_{\alpha,\beta} (\hat{e}_{\mathbf{k},\alpha})_j (\hat{e}_{\mathbf{k}',\beta})_l \\ \times \Big[\big(\hat{a}_{\mathbf{k},\alpha} \hat{a}^\dagger_{\mathbf{k}',\beta} - \hat{a}^\dagger_{\mathbf{k}',\beta} \hat{a}_{\mathbf{k},\alpha} \big) e^{i(\mathbf{k}\cdot\mathbf{r} - \mathbf{k}'\cdot\mathbf{r}')} \\ - \big(\hat{a}^\dagger_{\mathbf{k},\alpha} \hat{a}_{\mathbf{k}',\beta} - \hat{a}_{\mathbf{k}',\beta} \hat{a}^\dagger_{\mathbf{k},\alpha} \big) e^{-i(\mathbf{k}\cdot\mathbf{r} - \mathbf{k}'\cdot\mathbf{r}')} \Big].\end{aligned} \tag{C.13}$$

Here we already neglected the obviously commuting $[\hat{a}, \hat{a}] = [\hat{a}^\dagger, \hat{a}^\dagger]$ terms. The full commutation relations of the creation and annihilation operators:

$$[\hat{a}_{\mathbf{k},\alpha}, \hat{a}_{\mathbf{k}',\beta}] = [\hat{a}^\dagger_{\mathbf{k},\alpha}, \hat{a}^\dagger_{\mathbf{k}',\beta}] = 0 \tag{C.14}$$

$$[\hat{a}_{\mathbf{k},\alpha}, \hat{a}^\dagger_{\mathbf{k}',\beta}] = \delta_{\mathbf{k},\mathbf{k}'} \delta_{\alpha,\beta} \tag{C.15}$$

impose $\mathbf{k} = \mathbf{k}'$, and since different vector components in (C.13) commute with each other we also have $j = l$. However, we need to account for the fact that in the Coulomb gauge $\mathbf{A}$ is a transverse field, and the commutator should vanish when acting on longitudinal

fields. Therefore, instead of the Kronecker delta we should use the projector onto the transverse field:

$$\delta_{ij}^{\perp} = \delta_{ij} - \frac{k_i k_j}{k^2}, \tag{C.16}$$

and using $[\hat{a}, \hat{a}^\dagger] = 1$ we get [205]

$$\begin{aligned}[E_j(\mathbf{r},t), A_l(\mathbf{r}',t)] &= \frac{i\hbar}{\epsilon_0 V}\sum_{\mathbf{k}}\Big(\delta_{jl} + \frac{k_j k_l}{k^2}\Big)e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \\ &= \frac{i\hbar}{\epsilon_0}\Delta_{jl}^{\perp}(\mathbf{r}-\mathbf{r}'),\end{aligned} \tag{C.17}$$

where $\Delta_{jl}^{\perp}(\mathbf{r})$ is the transverse delta function, and the minus sign in the second exponential of (C.13) was ignored since it doesn't affect the delta function. To find the commutator between the electric and magnetic fields we use definition of the latter in terms of the vector potential:

$$B_i = (\nabla \times \mathbf{A})_i = \epsilon_{ijk}\frac{\partial}{\partial r_j}A_k, \tag{C.18}$$

where ϵ_{ijk} is the Levi-Civita symbol. Using this relation together with (C.17) gives

$$\begin{aligned}[E_j(\mathbf{r},t), B_k(\mathbf{r}',t)] &= [E_j(\mathbf{r},t), \epsilon_{klm}\frac{\partial}{\partial r_l'}A_m(\mathbf{r}',t)] \\ &= \frac{i\hbar}{\epsilon_0}\epsilon_{klm}\frac{\partial}{\partial r_l'}\Delta_{mj}^{\perp}(\mathbf{r}-\mathbf{r}').\end{aligned} \tag{C.19}$$

Since the curl of a longitudinal field is by definition zero, the transverse delta function can be replaced with the normal delta function (C.14) without changing the commutator:

$$[E_j(\mathbf{r},t), B_k(\mathbf{r}',t)] = \frac{i\hbar}{\epsilon_0}\epsilon_{klm}\frac{\partial}{\partial r_l'}\Delta_{mj}(\mathbf{r}-\mathbf{r}'), \tag{C.20}$$

and we can now evaluate the commutator with the nonlinear term in the Hamiltonian:

$$\frac{i}{\hbar}[H^{(2)}, B_m] = \frac{2i\epsilon_0}{3\hbar}\int d^3r\ \chi_{jkl}^{(2)}E_jE_kE_lB_m - B_m\frac{2i\epsilon_0}{3\hbar}\int d^3r\ \chi_{jkl}^{(2)}E_jE_kE_l. \tag{C.21}$$

This expression can be expanded as

$$\begin{aligned}&\frac{2i\epsilon_0}{3\hbar}\int d^3r\big(\chi_{jkl}^{(2)}E_jE_kE_lB_m - \chi_{jkl}^{(2)}E_jE_kB_mE_l + \chi_{jkl}^{(2)}E_jE_kB_mE_l - \chi_{jkl}^{(2)}E_jB_mE_kE_l \\ &\qquad +\chi_{jkl}^{(2)}E_jB_mE_kE_l - B_m\chi_{jkl}^{(2)}E_jE_kE_l\big) \\ =&\frac{2i\epsilon_0}{3\hbar}\int d^3r\big(\chi_{jkl}^{(2)}E_jE_k[E_l,B_m] + \chi_{jkl}^{(2)}E_j[E_k,B_m]E_l + \chi_{jkl}^{(2)}[E_j,B_m]E_kE_l\big).\end{aligned} \tag{C.22}$$

Using the permutation symmetry of the susceptibility tensor we then get [205]

$$\begin{aligned}\frac{i}{\hbar}[H^{(2)}, B_m] &= \frac{2i\epsilon_0}{\hbar}\int d^3r\ \chi^{(2)}_{jkl}[E_j, B_m]E_kE_l \\ &= -2\epsilon_{mno}\frac{\partial}{\partial r'_n}\int d^3r\Delta_{oj}(\mathbf{r}-\mathbf{r}')\chi^{(2)}_{jkl}E_kE_l \\ &= -2\epsilon_{mnj}\frac{\partial}{\partial r'_n}\chi^{(2)}_{jkl}E_kE_l,\end{aligned} \tag{C.23}$$

and hence

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{i}{\hbar}[H, \mathbf{B}] = -\nabla\times\mathbf{E} - \nabla\times(\chi^{(1)}\mathbf{E} + 2\chi^{(2)}\mathbf{E}^2). \tag{C.24}$$

This expression is clearly different from Faraday's law, a problem first pointed out in Ref. [205], and one that a rescaling of the electric-field operator does not solve. A more general argument from [216] shows why this is the case whenever the electric field is linear in the creation operators. This is because, using the following properties of the creation and annihilation operators:

$$[\hat{a}, f(\hat{a}, \hat{a}^\dagger)] = \frac{\partial f}{\partial \hat{a}^\dagger}, \qquad [\hat{a}^\dagger, f(\hat{a}, \hat{a}^\dagger)] = -\frac{\partial f}{\partial \hat{a}}, \tag{C.25}$$

one can infer that the evolution of the magnetic-field operator under a nonlinear Hamiltonian with terms of degree $N+1$ will always be described by some function that contains terms of degree N. However, if the electric field is linear in the creation and annihilation operators so is the curl of this field, leading to the same problem as above.

Another non-physical prediction this approach makes can, as pointed out in Ref. [205], be arrived at by finding the Heisenberg evolution of the electric field under the linear Hamiltonian. Simply comparing with (C.24) without explicitly carrying out the calculation shows that

$$\frac{\partial \mathbf{E}}{\partial t} = c^2(\nabla\times\mathbf{B}). \tag{C.26}$$

Considering only the linear evolution of the magnetic field (C.10) the wave equation for the electric field can be derived, using the same steps as in Section 2.4.1. Specifically, by taking the curl of (C.24), substituting in (C.26) and using $\nabla\cdot\mathbf{E} = 0$:

$$\begin{aligned} & -\nabla\times\nabla\times(1+\chi^{(1)})\mathbf{E} = \nabla\times\frac{\partial\mathbf{B}}{\partial t} \\ \iff \quad & n^2\nabla^2\mathbf{E} - n^2\nabla(\nabla\cdot\mathbf{E}) = \frac{\partial}{\partial t}(\nabla\times\mathbf{B}) \\ \iff \quad & \nabla^2\mathbf{E} = \frac{1}{n^2c^2}\frac{\partial^2\mathbf{E}}{\partial t^2}.\end{aligned} \tag{C.27}$$

The solutions to this equation are waves that travel at $nc > c$, instead of the actual speed c/n. To avoid these problems and recover Maxwell's equations it is necessary to treat the displacement field as fundamental, instead of the electric field. This is a reflection of the fact that, in this macroscopic approach to quantization, one does not have pure excitations of the electromagnetic field inside a medium, but instead joint light-matter excitations.

The alternative description, treating the displacement field as fundamental, appears in the literature as well [448, 449, 450]. To begin, the Hamiltonian (2.4-77) needs to be expressed in terms of the **D**-field. This can be done by Taylor expanding **E** in terms of **D**:

$$E_i = \beta_{ij}^{(1)} D_j + \beta_{ijk}^{(2)} D_j D_k + \cdots \tag{C.28}$$

The tensors $\beta^{(i)}$ can be expressed in terms of the susceptibilities $\chi^{(j)}$ by perturbatively solving the Taylor expansion of **D** in terms of **E** [205]:

$$D_i = \epsilon_0[(1 + \chi_{ij}^{(1)})E_j + \chi_{ijk}^{(2)} E_j E_k + \cdots]. \tag{C.29}$$

The lowest order β term is found by truncating the Taylor series after the first term and solving for it:

$$\begin{aligned} D_i &= \epsilon_0 (I + \chi^{(1)})_{ij} E_j \\ &= \epsilon_0 (I + \chi^{(1)})_{ij} \beta_{jk}^{(1)} D_k \end{aligned} \quad \Longleftrightarrow \quad \beta^{(1)} = [\epsilon_0 (I + \chi^{(1)})]^{-1}. \tag{C.30}$$

The next term can now be found by truncating (C.29) one order higher, and substituting in the first-order term of (C.28) in the second-order term of (C.29):

$$\begin{aligned} D_j &= \epsilon_0 (I + \chi^{(1)})_{jk} E_k + \epsilon_0 \chi_{jkl}^{(2)} E_k E_l \\ &= \epsilon_0 (I + \chi^{(1)})_{jk} E_k + \epsilon_0 \chi_{jkl}^{(2)} \beta_{km}^{(1)} D_m \beta_{ln}^{(1)} D_n. \end{aligned} \tag{C.31}$$

Rearranging and applying $\beta^{(1)}$ gives

$$\begin{aligned} \beta_{ij}^{(1)} D_j - \beta_{ij}^{(1)} \epsilon_0 \chi_{jkl}^{(2)} \beta_{km}^{(1)} D_m \beta_{ln}^{(1)} D_n &= \beta_{ij}^{(1)} \epsilon_0 (I + \chi^{(1)})_{jk} E_k \\ &= E_i \end{aligned} \tag{C.32}$$

Equating terms of equal order in the expression above and (C.28) one finds:

$$\beta_{imn}^{(2)} = -\epsilon_0 \beta_{ij}^{(1)} \chi_{jkl}^{(2)} \beta_{km}^{(1)} \beta_{ln}^{(1)} \tag{C.33}$$

One then proceeds the same way for the higher-order terms. To find the Hamiltonian in terms of the **D**-field, the Taylor series (C.28) has to be substituted into (2.4-77). Doing this for the first-order term is straightforward:

$$\begin{aligned} \epsilon_0 (I + \chi^{(1)})_{ij} E_i E_j &= \epsilon_0 (I + \chi^{(1)})_{ij} \beta_{ik}^{(1)} D_k \beta_{jl}^{(1)} D_l \\ &= \epsilon_0 \beta_{ik}^{(1)} D_i D_k. \end{aligned} \tag{C.34}$$

The second-order term in the Hamiltonian receives contributions from the product of first and second-order terms inserted into (2.4-77):

$$\begin{aligned} \epsilon_0 (I + \chi^{(1)})_{ij} E_i E_j &= \epsilon_0 (I + \chi^{(1)})_{ij} (\beta_{ik}^{(1)} D_k \beta_{jlm}^{(2)} D_l D_m + \beta_{ikm}^{(2)} D_k D_m \beta_{jl}^{(1)} D_l) \\ &= \beta_{jlm}^{(2)} D_j D_l D_m + \beta_{ikm}^{(2)} D_i D_k D_m, \end{aligned} \tag{C.35}$$

as well as the first-order term of the Taylor series (C.28) substituted into the second-order term of the Hamiltonian:

$$\begin{aligned}\epsilon_0 E_j \chi^{(2)}_{jkl} E_k E_l &= \epsilon_0 \beta^{(1)}_{ij} D_j \chi^{(2)}_{jkl} \beta^{(1)}_{km} D_m \beta^{(1)}_{ln} D_n \\ &= -\beta^{(2)}_{ikl} D_i D_k D_l.\end{aligned} \tag{C.36}$$

Adding (C.35) and (C.36) with the weights from (2.4-77) gives

$$\beta^{(2)}_{ijk} D_i D_j D_k - \frac{2}{3}\beta^{(2)}_{ijk} D_i D_j D_k = \frac{1}{3}\beta^{(2)}_{ijk} D_i D_j D_k. \tag{C.37}$$

The Hamiltonian truncated after the second order therefore reads

$$H = \int \left(\frac{\mathbf{B}^2}{2\mu_0} + \frac{1}{2}\beta^{(1)}\mathbf{D}^2 + \frac{1}{3}\beta^{(2)}\mathbf{D}^3 \right) d^3 r. \tag{C.38}$$

It can be shown that the general expression for the n-th order $\mathbf{D}$ term is [216, 451]:

$$H^{(n)} = \frac{1}{n+1}\beta^{(n)}\mathbf{D}^{n+1}. \tag{C.39}$$

The Hamiltonian can be quantized by introducing the annihilation operator [205]

$$\hat{a}_{\mathbf{k},\alpha}(t) = \frac{1}{\sqrt{\hbar V}} \int d^3 r e^{-i\mathbf{k}\cdot\mathbf{r}} \hat{e}_{\mathbf{k},\alpha} \cdot \left[\sqrt{\frac{\epsilon_0 \omega_k}{2}} \mathbf{A}(\mathbf{r},t) - i \frac{1}{\sqrt{2\epsilon_0 \omega_k}} \mathbf{D}(\mathbf{r},t) \right], \tag{C.40}$$

where, as before, $\hat{e}_{\mathbf{k},\alpha}$ is a polarization vector, V is the quantization volume and $\mathbf{A}$ is the vector potential of $\mathbf{B}$. The expressions for the field operators are found by inverting the equation above, which can be done using the transverse delta function $\Delta^{\perp}(\mathbf{r})$ introduced in the previous section. To do this we multiply by (C.40) by $\hat{e}_{\mathbf{k},\alpha} e^{i\mathbf{k}\cdot\mathbf{r}'}$ and sum:

$$\sum_{\mathbf{k},\alpha} \hat{e}_{\mathbf{k},\alpha} e^{i\mathbf{k}\cdot\mathbf{r}'} \hat{a}_{\mathbf{k},\alpha}(t) = \sum_{\mathbf{k},\alpha} \hat{e}_{\mathbf{k},\alpha} e^{i\mathbf{k}\cdot\mathbf{r}'} \frac{1}{\sqrt{\hbar V}} \int d^3 r e^{-i\mathbf{k}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{k},\alpha} \cdot \mathbf{F}(\mathbf{r},t), \tag{C.41}$$

where

$$\mathbf{F}(\mathbf{r},t) = \sqrt{\frac{\epsilon_0 \omega_k}{2}} \mathbf{A}(\mathbf{r},t) - i \frac{1}{\sqrt{2\epsilon_0 \omega_k}} \mathbf{D}(\mathbf{r},t). \tag{C.42}$$

Note that, since in the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$, both the vector potential and electric displacement are transverse fields, and thus $\mathbf{F}^{\perp}(\mathbf{r},t) = \mathbf{F}(\mathbf{r},t)$. Moving the sum inside the integral in (C.41) gives

$$\begin{aligned}\sum_{\mathbf{k},\alpha} \hat{e}_{\mathbf{k},\alpha} e^{i\mathbf{k}\cdot\mathbf{r}'} \hat{a}_{\mathbf{k},\alpha}(t) &= \frac{1}{\sqrt{\hbar V}} \int d^3 r \sum_{\mathbf{k},\alpha} e^{i\mathbf{k}(\mathbf{r}'-\mathbf{r})} \hat{e}_{\mathbf{k},\alpha} \hat{e}_{\mathbf{k},\alpha} \cdot \mathbf{F}(\mathbf{r},t) \\ &= \sqrt{\frac{V}{\hbar}} \int d^3 r \Delta^{\perp}(\mathbf{r}' - \mathbf{r}) \mathbf{F}(\mathbf{r},t) \\ &= \sqrt{\frac{V}{\hbar}} \mathbf{F}(\mathbf{r}',t).\end{aligned} \tag{C.43}$$

Doing the same for the adjoint of (C.40) we find

$$\begin{aligned}\sum_{\mathbf{k},\alpha}\hat{e}_{\mathbf{k},\alpha}e^{-i\mathbf{k}\cdot\mathbf{r}'}\hat{a}^\dagger_{\mathbf{k},\alpha}(t) &= \sqrt{\frac{V}{\hbar}}\mathbf{F}^*(\mathbf{r}',t)\\ &= \sqrt{\frac{\epsilon_0\omega_k V}{2\hbar}}\mathbf{A}(\mathbf{r}',t)+i\sqrt{\frac{V}{2\hbar\epsilon_0\omega_k}}\mathbf{D}(\mathbf{r}',t).\end{aligned} \tag{C.44}$$

The fields are now obtained by adding/subtracting (C.43) and (C.44), and solving for them:

$$\mathbf{D}(\mathbf{r},t)=\sum_{\mathbf{k},\alpha}i\sqrt{\frac{\hbar\epsilon_0\omega_k}{2V}}\hat{e}_{\mathbf{k},\alpha}\Big(\hat{a}_{\mathbf{k},\alpha}(t)e^{i\mathbf{k}\cdot\mathbf{r}}-\hat{a}^\dagger_{\mathbf{k},\alpha}(t)e^{-i\mathbf{k}\cdot\mathbf{r}}\Big) \tag{C.45}$$

$$\mathbf{A}(\mathbf{r},t)=\sum_{\mathbf{k},\alpha}\sqrt{\frac{\hbar}{2\epsilon_0\omega_k V}}\hat{e}_{\mathbf{k},\alpha}\Big(\hat{a}_{\mathbf{k},\alpha}(t)e^{i\mathbf{k}\cdot\mathbf{r}}+\hat{a}^\dagger_{\mathbf{k},\alpha}(t)e^{-i\mathbf{k}\cdot\mathbf{r}}\Big). \tag{C.46}$$

Finally, the **B**-field is found by taking the curl of (C.46):

$$\begin{aligned}\mathbf{B}(\mathbf{r},t) &= \nabla\times\mathbf{A}(\mathbf{r},t)\\ &= \sum_{\mathbf{k},\alpha}i\sqrt{\frac{\hbar}{2\epsilon_0\omega_k V}}(\mathbf{k}\times\hat{e}_{\mathbf{k},\alpha})\Big(\hat{a}_{\mathbf{k},\alpha}(t)e^{i\mathbf{k}\cdot\mathbf{r}}-\hat{a}^\dagger_{\mathbf{k},\alpha}(t)e^{-i\mathbf{k}\cdot\mathbf{r}}\Big)\end{aligned} \tag{C.47}$$

It can be shown that this approach is self consistent, does satisfy Maxwell's equations, and avoids the various unphysical predictions that appear when directly quantizing the electric field. As an example, we revisit the case of propagation in a linear medium. Direct comparison with the calculations in the beginning of the section shows that

$$[D_j(\mathbf{r},t),A_l(\mathbf{r}',t)]=i\hbar\Delta^\perp_{jl}(\mathbf{r}-\mathbf{r}'), \tag{C.48}$$

and thus

$$[D_j(\mathbf{r},t),B_k(\mathbf{r}',t)]=i\hbar\epsilon_{klm}\frac{\partial}{\partial r'_l}\Delta_{mj}(\mathbf{r}-\mathbf{r}'). \tag{C.49}$$

The steps for finding the equation of motion for the magnetic-field operator are therefore exactly the same as before, and using the linear part of the Hamiltonian (C.38) we find

$$\frac{\partial\mathbf{B}}{\partial t}=\frac{i}{\hbar}[H,\mathbf{B}]=-\nabla\times\beta^{(1)}\mathbf{D}. \tag{C.50}$$

For the displacement field we similarly get

$$\frac{\partial\mathbf{D}}{\partial t}=\frac{i}{\hbar}[H,\mathbf{D}]=\frac{1}{\mu_0}\nabla\times\mathbf{B}. \tag{C.51}$$

We can now derive the wave equation for the displacement field:

$$\begin{aligned}&-\nabla\times\nabla\times\beta^{(1)}\mathbf{D}=\nabla\times\frac{\partial\mathbf{B}}{\partial t}\\ \Longleftrightarrow\qquad &\beta^{(1)}\nabla^2\mathbf{D}=\frac{\partial}{\partial t}(\nabla\times\mathbf{B}).\end{aligned} \tag{C.52}$$

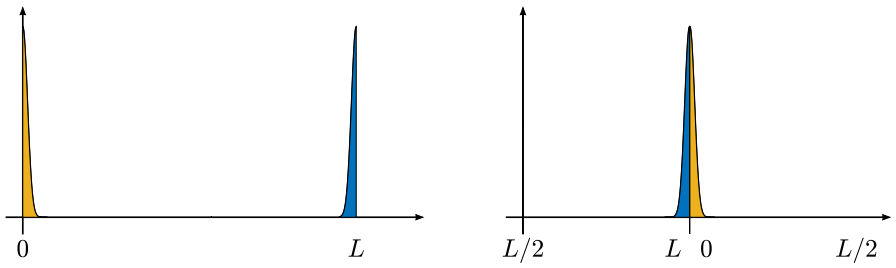


Figure C.1: **Delta function on finite interval.** Rearranging the interval so that the endpoints overlap in the center of the new interval allows the two delta functions on the endpoints, pictured here as narrow Gaussians, to act as a single delta function that can be integrated over.

Using $\beta^{(1)} = 1/(\epsilon_0 n^2)$ and (C.51) we finally get

$$\nabla^2 \mathbf{D} = \epsilon_0 \mu_0 n^2 \frac{\partial^2 \mathbf{D}}{\partial t^2} = \frac{n^2}{c^2} \frac{\partial^2 \mathbf{D}}{\partial t^2}. \tag{C.53}$$

We thus see that the excitations of the now fundamental displacement field propagate at the correct speed c/n inside the medium.

C.4 Delta function on finite interval

As shown in Section C.2, when considering a finite interval $[0, L]$ with periodic boundary conditions:

$$f(0) = f(L), \tag{C.1}$$

The dirac comb

$$\sum_n \delta(x - nL) = \text{Ш}(x) = \frac{1}{L} \sum_k e^{i2\pi kx/L} \tag{C.2}$$

acts as a delta function:

$$\int_0^L \text{Ш}(x) f(x) dx = f(x). \tag{C.3}$$

For the open interval $(0, L)$ this follows from the fact that the terms in the left-hand sum with $|n| > 0$ lie outside the interval, and the sum simply reduces to $\delta(x)$. However, terms with $n = \pm 1$ overlap with the endpoints of the closed interval for $x = L$ and $x = 0$, respectively. This results in one-sided integrals over the delta function, for example

$$\int_0^L \delta(x - L) f(x) dx, \tag{C.4}$$

which are not defined. However, due to the periodic boundary condition one can, as shown in Fig. C.1, translate the interval and combine the delta functions in the zeroth and first order terms into a single delta function. Consider

$$\begin{aligned} &\int_0^L \big[\delta(x) + \delta(x - L)\big] f(x) dx \\ &= \int_0^{L/2} \big[\delta(x) + \delta(x - L)\big] f(x) dx + \int_{L/2}^L \big[\delta(x) + \delta(x - L)\big] f(x) dx. \end{aligned} \tag{C.5}$$

The second (first) delta function in the first (second) integral always lies outside the integration interval and can therefore be dropped. Let

$$t = \begin{cases} x + \frac{L}{2} & \text{if } x \in [0, L/2) \\ x - \frac{L}{2} & \text{if } x \in [L/2, 1] \end{cases} \tag{C.6}$$

and

$$g(x) = \begin{cases} f(x + \frac{L}{2}) & \text{if } x \in [0, L/2) \\ f(x - \frac{L}{2}) & \text{if } x \in [L/2, 1]. \end{cases} \tag{C.7}$$

Performing the variable substitution gives

$$\begin{aligned} &\int_0^L \big[\delta(x) + \delta(x - L)\big] f(x) dx \\ &= \int_0^{L/2} \delta(x) f(x) dx + \int_{L/2}^L \delta(x - L) f(x) dx \\ &= \int_{[L/2, L)} \delta(t - \frac{L}{2}) f(t - \frac{L}{2}) dt + \int_0^{L/2} \delta(t - \frac{L}{2}) f(t + \frac{L}{2}) dt \\ &= \int_{[0, L)} \delta(t - \frac{L}{2}) g(t) dt = g(\frac{L}{2}) = f(0), \end{aligned} \tag{C.8}$$

which shows that the dirac comb $\text{Ш}(x)$ works as a delta function on the endpoints of the interval as well.

Bibliography

[1] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?," *Physical review*, vol. 47, no. 10, p. 777, 1935.

[2] J. S. Bell, "On the einstein podolsky rosen paradox," *Physics Physique Fizika*, vol. 1, no. 3, p. 195, 1964.

[3] A. Aspect, J. Dalibard, and G. Roger, "Experimental test of bell's inequalities using time-varying analyzers," *Physical review letters*, vol. 49, no. 25, p. 1804, 1982.

[4] R. P. Feynman, "Simulating Physics with Computers," *International Journal of Theoretical Physics*, vol. 21, pp. 467–488, June 1982.

[5] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, pp. 175–179, 1984.

[6] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM review*, vol. 41, no. 2, pp. 303–332, 1999.

[7] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, "Experimental quantum teleportation," *Nature*, vol. 390, no. 6660, pp. 575–579, 1997.

[8] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.

[9] W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland, "Quantum zeno effect," *Physical Review A*, vol. 41, no. 5, p. 2295, 1990.

[10] M. Fuwa, S. Takeda, M. Zwierz, H. M. Wiseman, and A. Furusawa, "Experimental proof of nonlocal wavefunction collapse for a single particle using homodyne measurements," *Nature Communications*, vol. 6, no. 1, pp. 1–6, 2015.

[11] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, "Experimental verification of an indefinite causal order," *Science advances*, vol. 3, no. 3, p. e1602589, 2017.

[12] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[13] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM journal of research and development*, vol. 5, no. 3, pp. 183–191, 1961.

[14] B. Schumacher, "Quantum coding," *Physical Review A*, vol. 51, no. 4, p. 2738, 1995.

[15] M. A. Nielsen and I. Chuang, "Quantum computation and quantum information," 2002.

[16] W. Gerlach and O. Stern, "Der experimentelle nachweis der richtungsquantelung im magnetfeld," *Zeitschrift für Physik*, vol. 9, no. 1, pp. 349–352, 1922.

[17] J. Von Neumann, *Mathematical foundations of quantum mechanics: New edition*. Princeton university press, 2018.

[18] J. J. Sakurai and E. D. Commins, "Modern quantum mechanics, revised edition," 1995.

[19] C. H. Bennett and S. J. Wiesner, "Communication via one-and two-particle operators on einstein-podolsky-rosen states," *Physical review letters*, vol. 69, no. 20, p. 2881, 1992.

[20] A. K. Ekert, "Quantum cryptography and bell's theorem," in *Quantum Measurements in Optics*, pp. 413–418, Springer, 1992.

[21] V. Bužek, M. Hillery, and R. F. Werner, "Optimal manipulations with qubits: Universal-not gate," *Physical Review A*, vol. 60, no. 4, p. R2626, 1999.

[22] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, "Quantum computation by adiabatic evolution," *arXiv preprint quant-ph/0001106*, 2000.

[23] R. Raussendorf and H. J. Briegel, "A one-way quantum computer," *Physical review letters*, vol. 86, no. 22, p. 5188, 2001.

[24] R. Jozsa, "An introduction to measurement based quantum computation," *NATO Science Series, III: Computer and Systems Sciences. Quantum Information Processing-From Theory to Experiment*, vol. 199, pp. 137–158, 2006.

[25] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, "Elementary gates for quantum computation," *Physical review A*, vol. 52, no. 5, p. 3457, 1995.

[26] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information," *Reviews of Modern Physics*, vol. 84, no. 2, p. 621, 2012.

[27] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," *Nature*, vol. 299, no. 5886, pp. 802–803, 1982.

[28] D. Dieks, "Communication by epr devices," *Physics Letters A*, vol. 92, no. 6, pp. 271–272, 1982.

[29] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels," *Physical review letters*, vol. 70, no. 13, p. 1895, 1993.

[30] A. Furusawa, J. L. Sørensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, and E. S. Polzik, "Unconditional quantum teleportation," *Science*, vol. 282, no. 5389, pp. 706–709, 1998.

[31] J.-W. Pan, D. Bouwmeester, H. Weinfurter, and A. Zeilinger, "Experimental entanglement swapping: entangling photons that never interacted," *Physical review letters*, vol. 80, no. 18, p. 3891, 1998.

[32] D. Gottesman and I. L. Chuang, "Quantum teleportation is a universal computational primitive," *arXiv preprint quant-ph/9908010*, 1999.

[33] D. Gottesman, "The heisenberg representation of quantum computers," *arXiv preprint quant-ph/9807006*, 1998.

[34] S. Aaronson and D. Gottesman, "Improved simulation of stabilizer circuits," *Physical Review A*, vol. 70, no. 5, p. 052328, 2004.

[35] S. Bravyi and A. Kitaev, "Universal quantum computation with ideal clifford gates and noisy ancillas," *Physical Review A*, vol. 71, no. 2, p. 022316, 2005.

[36] X. Zhou, D. W. Leung, and I. L. Chuang, "Methodology for quantum logic gate construction," *Physical Review A*, vol. 62, no. 5, p. 052316, 2000.

[37] S. Bravyi and D. Gosset, "Improved classical simulation of quantum circuits dominated by clifford gates," *Physical review letters*, vol. 116, no. 25, p. 250501, 2016.

[38] B. Eastin and E. Knill, "Restrictions on transversal encoded quantum gate sets," *Physical review letters*, vol. 102, no. 11, p. 110502, 2009.

[39] E. Knill, R. Laflamme, and G. J. Milburn, "A scheme for efficient quantum computation with linear optics," *nature*, vol. 409, no. 6816, pp. 46–52, 2001.

[40] E. Magesan, R. Blume-Kohout, and J. Emerson, "Gate fidelity fluctuations and quantum process invariants," *Physical Review A*, vol. 84, no. 1, p. 012309, 2011.

[41] F. Mezzadri, "How to generate random matrices from the classical compact groups," *arXiv preprint math-ph/0609050*, 2006.

[42] B. M. Terhal, "Bell inequalities and the separability criterion," *Physics Letters A*, vol. 271, no. 5-6, pp. 319–326, 2000.

[43] Z. Hradil, "Quantum-state estimation," *Physical Review A*, vol. 55, no. 3, p. R1561, 1997.

[44] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, "Efficient quantum state tomography," *Nature communications*, vol. 1, no. 1, pp. 1–7, 2010.

[45] J. Bavaresco, N. Herrera Valencia, C. Klöckl, M. Pivoluska, P. Erker, N. Friis, M. Malik, and M. Huber, "Measurements in two bases are sufficient for certifying high-dimensional entanglement," *Nature Physics*, vol. 14, no. 10, pp. 1032–1037, 2018.

[46] J. B. Altepeter, E. R. Jeffrey, and P. G. Kwiat, "Photonic state tomography," *Advances in Atomic, Molecular, and Optical Physics*, vol. 52, pp. 105–159, 2005.

[47] D. F. James, P. G. Kwiat, W. J. Munro, and A. G. White, "On the measurement of qubits," in *Asymptotic Theory of Quantum Statistical Inference: Selected Papers*, pp. 509–538, World Scientific, 2005.

[48] J. Řeháček, Z. Hradil, and M. Ježek, "Iterative algorithm for reconstruction of entangled states," *Physical Review A*, vol. 63, no. 4, p. 040303, 2001.

[49] W. F. Stinespring, "Positive functions on c*-algebras," *Proceedings of the American Mathematical Society*, vol. 6, no. 2, pp. 211–216, 1955.

[50] D. Kretschmann, D. Schlingemann, and R. F. Werner, "The information-disturbance tradeoff and the continuity of stinespring's representation," *IEEE transactions on information theory*, vol. 54, no. 4, pp. 1708–1717, 2008.

[51] M.-D. Choi, "Completely positive linear maps on complex matrices," *Linear algebra and its applications*, vol. 10, no. 3, pp. 285–290, 1975.

[52] I. Bengtsson and K. Życzkowski, *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge university press, 2017.

[53] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, "Witnessing causal nonseparability," *New Journal of Physics*, vol. 17, no. 10, p. 102001, 2015.

[54] O. Oreshkov, F. Costa, and Č. Brukner, "Quantum correlations with no causal order," *Nature communications*, vol. 3, no. 1, pp. 1–8, 2012.

[55] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, "Quantum computations without definite causal structure," *Physical Review A*, vol. 88, no. 2, p. 022318, 2013.

[56] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, "Indefinite causal order in a quantum switch," *Physical review letters*, vol. 121, no. 9, p. 090503, 2018.

[57] C. Branciard, M. Araújo, A. Feix, F. Costa, and Č. Brukner, "The simplest causal inequalities and their violation," *New Journal of Physics*, vol. 18, no. 1, p. 013008, 2015.

[58] T. Purves and A. J. Short, "Quantum theory cannot violate a causal inequality," *Physical Review Letters*, vol. 127, no. 11, p. 110402, 2021.

[59] J. Wechs, C. Branciard, and O. Oreshkov, "Existence of processes violating causal inequalities on time-delocalised subsystems," *arXiv preprint arXiv:2201.11832*, 2022.

[60] D. W. Leung, "Choi's proof as a recipe for quantum process tomography," *Journal of Mathematical Physics*, vol. 44, no. 2, pp. 528–533, 2003.

[61] J. B. Altepeter, D. Branning, E. Jeffrey, T. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White, "Ancilla-assisted quantum process tomography," *Physical Review Letters*, vol. 90, no. 19, p. 193601, 2003.

[62] K. Mattle, H. Weinfurter, P. G. Kwiat, and A. Zeilinger, "Dense coding in experimental quantum communication," *Physical Review Letters*, vol. 76, no. 25, p. 4656, 1996.

[63] C. E. Shannon, "Communication theory of secrecy systems," *The Bell system technical journal*, vol. 28, no. 4, pp. 656–715, 1949.

[64] A. C. Elitzur and L. Vaidman, "Quantum mechanical interaction-free measurements," *Foundations of Physics*, vol. 23, no. 7, pp. 987–997, 1993.

[65] P. Kwiat, H. Weinfurter, T. Herzog, A. Zeilinger, and M. A. Kasevich, "Interaction-free measurement," *Physical Review Letters*, vol. 74, no. 24, p. 4763, 1995.

[66] H. Salih, Z.-H. Li, M. Al-Amri, and M. S. Zubairy, "Protocol for direct counterfactual quantum communication," *Physical review letters*, vol. 110, no. 17, p. 170502, 2013.

[67] L. Vaidman, "Comment on "protocol for direct counterfactual quantum communication"," *Physical Review Letters*, vol. 112, no. 20, p. 208901, 2014.

[68] H. Salih, Z.-H. Li, M. Al-Amri, and M. S. Zubairy, "Salih et al. reply," *Physical Review Letters*, vol. 112, no. 20, p. 208902, 2014.

[69] L. Vaidman, "Comment on "direct counterfactual transmission of a quantum state"," *Physical Review A*, vol. 93, no. 6, p. 066301, 2016.

[70] R. B. Griffiths, "Particle path through a nested mach-zehnder interferometer," *Physical Review A*, vol. 94, no. 3, p. 032115, 2016.

[71] D. Arvidsson-Shukur, A. Gottfries, and C. Barnes, "Evaluation of counterfactuality in counterfactual communication protocols," *Physical Review A*, vol. 96, no. 6, p. 062316, 2017.

[72] L. Vaidman, "Analysis of counterfactuality of counterfactual communication protocols," *Physical Review A*, vol. 99, no. 5, p. 052127, 2019.

[73] L. Vaidman, "Past of a quantum particle," *Physical Review A*, vol. 87, no. 5, p. 052104, 2013.

[74] Y. Aharonov, D. Z. Albert, and L. Vaidman, "How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100," *Physical review letters*, vol. 60, no. 14, p. 1351, 1988.

[75] B. Tamir and E. Cohen, "Introduction to weak measurements and weak values," *Quanta*, vol. 2, no. 1, pp. 7–17, 2013.

[76] A. Peres, *Quantum theory: concepts and methods*, vol. 72. Springer, 1997.

[77] L. Vaidman, "Tracing the past of a quantum particle," *Physical Review A*, vol. 89, no. 2, p. 024102, 2014.

[78] M. A. Alonso and A. N. Jordan, "Can a dove prism change the past of a single photon?," *Quantum Studies: Mathematics and Foundations*, vol. 2, no. 3, pp. 255–261, 2015.

[79] Q. Duprey and A. Matzkin, "Null weak values and the past of a quantum particle," *Physical Review A*, vol. 95, no. 3, p. 032110, 2017.

[80] D. Sokolovski, "Comment on "null weak values and the past of a quantum particle"," *Physical Review A*, vol. 97, no. 4, p. 046102, 2018.

[81] L. Vaidman and I. Tsutsui, "When photons are lying about where they have been," *Entropy*, vol. 20, no. 7, p. 538, 2018.

[82] U. Peleg and L. Vaidman, "Comment on "past of a quantum particle revisited"," *Physical Review A*, vol. 99, no. 2, p. 026103, 2019.

[83] Y. Aharonov and L. Vaidman, "Properties of a quantum system during the time interval between two measurements," *Physical Review A*, vol. 41, no. 1, p. 11, 1990.

[84] Y. Aharonov and L. Vaidman, "The two-state vector formalism: an updated review," *Time in quantum mechanics*, pp. 399–447, 2008.

[85] D. R. M. Arvidsson-Shukur and C. H. W. Barnes, "Quantum counterfactual communication without a weak trace," *Physical Review A*, vol. 94, no. 6, p. 062303, 2016.

[86] T.-G. Noh *et al.*, "Counterfactual quantum cryptography," *Physical review letters*, vol. 103, no. 23, p. 230501, 2009.

[87] J. J. Pla, K. Y. Tan, J. P. Dehollain, W. H. Lim, J. J. Morton, F. A. Zwanenburg, D. N. Jamieson, A. S. Dzurak, and A. Morello, "High-fidelity readout and control of a nuclear spin qubit in silicon," *Nature*, vol. 496, no. 7445, pp. 334–338, 2013.

[88] M. Veldhorst, C. Yang, J. Hwang, W. Huang, J. Dehollain, J. Muhonen, S. Simmons, A. Laucht, F. Hudson, K. M. Itoh, *et al.*, "A two-qubit logic gate in silicon," *Nature*, vol. 526, no. 7573, pp. 410–414, 2015.

[89] A. P. Place, L. V. Rodgers, P. Mundada, B. M. Smitham, M. Fitzpatrick, Z. Leng, A. Premkumar, J. Bryon, A. Vrajitoarea, S. Sussman, *et al.*, "New material platform for superconducting transmon qubits with coherence times exceeding 0.3 milliseconds," *Nature communications*, vol. 12, no. 1, pp. 1–6, 2021.

[90] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, *et al.*, "Superconducting quantum circuits at the surface code threshold for fault tolerance," *Nature*, vol. 508, no. 7497, pp. 500–503, 2014.

[91] A. Kandala, K. X. Wei, S. Srinivasan, E. Magesan, S. Carnevale, G. Keefe, D. Klaus, O. Dial, and D. McKay, "Demonstration of a high-fidelity cnot gate for fixed-frequency transmons with engineered z z suppression," *Physical Review Letters*, vol. 127, no. 13, p. 130501, 2021.

[92] V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics: Volume 4*, vol. 4. Butterworth-Heinemann, 1982.

[93] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, "Cavity optomechanics," *Reviews of Modern Physics*, vol. 86, no. 4, p. 1391, 2014.

[94] R. Riedinger, S. Hong, R. A. Norte, J. A. Slater, J. Shang, A. G. Krause, V. Anant, M. Aspelmeyer, and S. Gröblacher, "Non-classical correlations between single photons and phonons from a mechanical oscillator," *Nature*, vol. 530, no. 7590, pp. 313–316, 2016.

[95] M. Giustina, M. A. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán, *et al.*, "Significant-loophole-free test of bell's theorem with entangled photons," *Physical review letters*, vol. 115, no. 25, p. 250401, 2015.

[96] P. Walther, K. J. Resch, T. Rudolph, E. Schenck, H. Weinfurter, V. Vedral, M. Aspelmeyer, and A. Zeilinger, "Experimental one-way quantum computing," *Nature*, vol. 434, no. 7030, pp. 169–176, 2005.

[97] Y.-W. Cho, G. Campbell, J. Everett, J. Bernu, D. Higginbottom, M. Cao, J. Geng, N. Robins, P. Lam, and B. Buchler, "Highly efficient optical quantum memory with long coherence time in cold atoms," *Optica*, vol. 3, no. 1, pp. 100–107, 2016.

[98] D. A. Steck, "Quantum and atom optics," 2007.

[99] C. J. Foot, *Atomic physics*, vol. 7. OUP Oxford, 2004.

[100] M. E. Peskin and D. V. Schroeder, "An introduction to quantum field theory," 1995.

[101] H. Bruus and K. Flensberg, *Many-body quantum theory in condensed matter physics: an introduction*. OUP Oxford, 2004.

[102] V. Parigi, A. Zavatta, M. Kim, and M. Bellini, "Probing quantum commutation rules by addition and subtraction of single photons to/from a light field," *Science*, vol. 317, no. 5846, pp. 1890–1893, 2007.

[103] A. Zavatta, V. Parigi, M. Kim, H. Jeong, and M. Bellini, "Experimental demonstration of the bosonic commutation relation via superpositions of quantum operations on thermal light fields," *Physical Review Letters*, vol. 103, no. 14, p. 140406, 2009.

[104] R. Loudon, *The quantum theory of light*. OUP Oxford, 2000.

[105] A. M. Fox, M. Fox, *et al.*, *Quantum optics: an introduction*, vol. 15. Oxford university press, 2006.

[106] B. E. Saleh and M. C. Teich, *Fundamentals of photonics*. john Wiley & sons, 2019.

[107] R. H. Brown and R. Q. Twiss, "Correlation between photons in two coherent beams of light," *Nature*, vol. 177, no. 4497, pp. 27–29, 1956.

[108] M. O. Scully and M. S. Zubairy, "Quantum optics," 1999.

[109] J. B. Brask, "Gaussian states and operations–a quick reference," *arXiv preprint arXiv:2102.05748*, 2021.

[110] J. Garrison and R. Chiao, *Quantum optics*. OUP Oxford, 2008.

[111] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.*, vol. 40, pp. 749–759, Jun 1932.

[112] A. Kenfack and K. Życzkowski, "Negativity of the wigner function as an indicator of non-classicality," *Journal of Optics B: Quantum and Semiclassical Optics*, vol. 6, no. 10, p. 396, 2004.

[113] M. Gu, C. Weedbrook, N. C. Menicucci, T. C. Ralph, and P. van Loock, "Quantum computing with continuous-variable clusters," *Physical Review A*, vol. 79, no. 6, p. 062318, 2009.

[114] S. L. Braunstein and P. Van Loock, "Quantum information with continuous variables," *Reviews of modern physics*, vol. 77, no. 2, p. 513, 2005.

[115] K. E. Cahill and R. J. Glauber, "Density operators and quasiprobability distributions," *Physical Review*, vol. 177, no. 5, p. 1882, 1969.

[116] U. Leonhardt, *Measuring the quantum state of light*, vol. 22. Cambridge university press, 1997.

[117] A. I. Lvovsky, "Squeezed light," *Photonics: Scientific Foundations, Technology and Applications*, vol. 1, pp. 121–163, 2015.

[118] C. M. Caves, "Quantum-mechanical noise in an interferometer," *Physical Review D*, vol. 23, no. 8, p. 1693, 1981.

[119] M. Tse, H. Yu, N. Kijbunchoo, A. Fernandez-Galiana, P. Dupej, L. Barsotti, C. Blair, D. Brown, S. Dwyer, A. Effler, *et al.*, "Quantum-enhanced advanced ligo detectors in the era of gravitational-wave astronomy," *Physical Review Letters*, vol. 123, no. 23, p. 231107, 2019.

[120] A. Biswas and G. S. Agarwal, "Nonclassicality and decoherence of photon-subtracted squeezed states," *Physical Review A*, vol. 75, no. 3, p. 032104, 2007.

[121] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, "Linear optical quantum computing with photonic qubits," *Reviews of modern physics*, vol. 79, no. 1, p. 135, 2007.

[122] S. Aaronson and A. Arkhipov, "The Computational Complexity of Linear Optics," *arXiv e-prints*, p. arXiv:1011.3245, Nov. 2010.

[123] M. C. Tichy, "Interference of identical particles from entanglement to boson-sampling," *Journal of Physics B: Atomic, Molecular and Optical Physics*, vol. 47, no. 10, p. 103001, 2014.

[124] M. Tillmann, S.-H. Tan, S. E. Stoeckl, B. C. Sanders, H. De Guise, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, "Generalized multiphoton quantum interference," *Physical Review X*, vol. 5, no. 4, p. 041015, 2015.

[125] V. Shchesnovich, "Partial indistinguishability theory for multiphoton experiments in multiport devices," *Physical Review A*, vol. 91, no. 1, p. 013844, 2015.

[126] M. C. Tichy, "Sampling of partially distinguishable bosons and the relation to the multidimensional permanent," *Physical Review A*, vol. 91, no. 2, p. 022316, 2015.

[127] J. J. Renema, A. Menssen, W. R. Clements, G. Triginer, W. S. Kolthammer, and I. A. Walmsley, "Efficient classical algorithm for boson sampling with partially distinguishable photons," *Physical review letters*, vol. 120, no. 22, p. 220502, 2018.

[128] R. Ghosh and L. Mandel, "Observation of nonclassical effects in the interference of two photons," *Physical Review Letters*, vol. 59, no. 17, p. 1903, 1987.

[129] C.-K. Hong, Z.-Y. Ou, and L. Mandel, "Measurement of subpicosecond time intervals between two photons by interference," *Physical review letters*, vol. 59, no. 18, p. 2044, 1987.

[130] J. L. O'brien, A. Furusawa, and J. Vučković, "Photonic quantum technologies," *Nature Photonics*, vol. 3, no. 12, pp. 687–695, 2009.

[131] F. Flamini, N. Spagnolo, and F. Sciarrino, "Photonic quantum information processing: a review," *Reports on Progress in Physics*, vol. 82, no. 1, p. 016001, 2018.

[132] S. Slussarenko and G. J. Pryde, "Photonic quantum information processing: A concise review," *Applied Physics Reviews*, vol. 6, no. 4, p. 041303, 2019.

[133] J.-i. Yoshikawa, S. Yokoyama, T. Kaji, C. Sornphiphatphong, Y. Shiozawa, K. Makino, and A. Furusawa, "Invited article: Generation of one-million-mode continuous-variable cluster state by unlimited time-domain multiplexing," *APL photonics*, vol. 1, no. 6, p. 060801, 2016.

[134] W. Asavanant, Y. Shiozawa, S. Yokoyama, B. Charoensombutamon, H. Emura, R. N. Alexander, S. Takeda, J.-i. Yoshikawa, N. C. Menicucci, H. Yonezawa, *et al.*, "Generation of time-domain-multiplexed two-dimensional cluster state," *Science*, vol. 366, no. 6463, pp. 373–376, 2019.

[135] M. V. Larsen, X. Guo, C. R. Breum, J. S. Neergaard-Nielsen, and U. L. Andersen, "Deterministic generation of a two-dimensional cluster state," *Science*, vol. 366, no. 6463, pp. 369–372, 2019.

[136] D. Gottesman, A. Kitaev, and J. Preskill, "Encoding a qubit in an oscillator," *Physical Review A*, vol. 64, no. 1, p. 012310, 2001.

[137] I. Tzitrin, J. E. Bourassa, N. C. Menicucci, and K. K. Sabapathy, "Progress towards practical qubit computation using approximate gottesman-kitaev-preskill codes," *Physical Review A*, vol. 101, no. 3, p. 032315, 2020.

[138] Y. Yu, F. Ma, X.-Y. Luo, B. Jing, P.-F. Sun, R.-Z. Fang, C.-W. Yang, H. Liu, M.-Y. Zheng, X.-P. Xie, *et al.*, "Entanglement of two quantum memories via fibres over dozens of kilometres," *Nature*, vol. 578, no. 7794, pp. 240–245, 2020.

[139] H.-S. Zhong, Y. Li, W. Li, L.-C. Peng, Z.-E. Su, Y. Hu, Y.-M. He, X. Ding, W. Zhang, H. Li, *et al.*, "12-photon entanglement and scalable scattershot boson sampling with optimal entangled-photon pairs from parametric down-conversion," *Physical review letters*, vol. 121, no. 25, p. 250505, 2018.

[140] Z.-D. Li, R. Zhang, X.-F. Yin, L.-Z. Liu, Y. Hu, Y.-Q. Fang, Y.-Y. Fei, X. Jiang, J. Zhang, L. Li, *et al.*, "Experimental quantum repeater without quantum memory," *Nature photonics*, vol. 13, no. 9, pp. 644–648, 2019.

[141] A. Rossi, G. Vallone, A. Chiuri, F. De Martini, and P. Mataloni, "Multipath entanglement of two photons," *Physical review letters*, vol. 102, no. 15, p. 153902, 2009.

[142] J. C. Matthews, A. Politi, A. Stefanov, and J. L. O'brien, "Manipulation of multiphoton entanglement in waveguide quantum circuits," *Nature Photonics*, vol. 3, no. 6, pp. 346–350, 2009.

[143] P. J. Shadbolt, M. R. Verde, A. Peruzzo, A. Politi, A. Laing, M. Lobino, J. C. Matthews, M. G. Thompson, and J. L. O'Brien, "Generating, manipulating and measuring entanglement and mixture with a reconfigurable photonic circuit," *Nature Photonics*, vol. 6, no. 1, pp. 45–49, 2012.

[144] I. Marcikic, H. De Riedmatten, W. Tittel, H. Zbinden, and N. Gisin, "Long-distance teleportation of qubits at telecommunication wavelengths," *Nature*, vol. 421, no. 6922, pp. 509–513, 2003.

[145] I. Marcikic, H. De Riedmatten, W. Tittel, H. Zbinden, M. Legré, and N. Gisin, "Distribution of time-bin entangled qubits over 50

km of optical fiber," *Physical Review Letters*, vol. 93, no. 18, p. 180502, 2004.

[146] K. Y. Bliokh and F. Nori, "Transverse and longitudinal angular momenta of light," *Physics Reports*, vol. 592, pp. 1–38, 2015.

[147] K. Y. Bliokh, F. J. Rodríguez-Fortuño, F. Nori, and A. V. Zayats, "Spin–orbit interactions of light," *Nature Photonics*, vol. 9, no. 12, pp. 796–808, 2015.

[148] B. Brecht, D. V. Reddy, C. Silberhorn, and M. G. Raymer, "Photon temporal modes: a complete framework for quantum information science," *Physical Review X*, vol. 5, no. 4, p. 041017, 2015.

[149] U. L. Andersen, J. S. Neergaard-Nielsen, P. Van Loock, and A. Furusawa, "Hybrid discrete-and continuous-variable quantum information," *Nature Physics*, vol. 11, no. 9, pp. 713–719, 2015.

[150] S. Walborn, S. Pádua, and C. Monken, "Hyperentanglement-assisted bell-state analysis," *Physical Review A*, vol. 68, no. 4, p. 042313, 2003.

[151] J. T. Barreiro, N. K. Langford, N. A. Peters, and P. G. Kwiat, "Generation of hyperentangled photon pairs," *Physical review letters*, vol. 95, no. 26, p. 260501, 2005.

[152] X.-L. Wang, Y.-H. Luo, H.-L. Huang, M.-C. Chen, Z.-E. Su, C. Liu, C. Chen, W. Li, Y.-Q. Fang, X. Jiang, *et al.*, "18-qubit entanglement with six photons' three degrees of freedom," *Physical review letters*, vol. 120, no. 26, p. 260502, 2018.

[153] Y.-H. Luo, M.-C. Chen, M. Erhard, H.-S. Zhong, D. Wu, H.-Y. Tang, Q. Zhao, X.-L. Wang, K. Fujii, L. Li, *et al.*, "Quantum teleportation of physical qubits into logical code spaces," *Proceedings of the National Academy of Sciences*, vol. 118, no. 36, p. e2026250118, 2021.

[154] R. Fickler, R. Lapkiewicz, W. N. Plick, M. Krenn, C. Schaeff, S. Ramelow, and A. Zeilinger, "Quantum entanglement of high angular momenta," *Science*, vol. 338, no. 6107, pp. 640–643, 2012.

[155] M. Krenn, M. Huber, R. Fickler, R. Lapkiewicz, S. Ramelow, and A. Zeilinger, "Generation and confirmation of a (100× 100)-dimensional entangled quantum system," *Proceedings of the National Academy of Sciences*, vol. 111, no. 17, pp. 6243–6247, 2014.

[156] M. Malik, M. Erhard, M. Huber, M. Krenn, R. Fickler, and A. Zeilinger, "Multi-photon entanglement in high dimensions," *Nature Photonics*, vol. 10, no. 4, pp. 248–252, 2016.

[157] M. Erhard, R. Fickler, M. Krenn, and A. Zeilinger, "Twisted photons: new quantum perspectives in high dimensions," *Light: Science & Applications*, vol. 7, no. 3, pp. 17146–17146, 2018.

[158] S. Bartolucci, P. M. Birchall, M. Gimeno-Segovia, E. Johnston, K. Kieling, M. Pant, T. Rudolph, J. Smith, C. Sparrow, and M. D. Vidrighin, "Creation of entangled photonic states using linear optics," *arXiv preprint arXiv:2106.13825*, 2021.

[159] S. Bartolucci, P. Birchall, H. Bombin, H. Cable, C. Dawson, M. Gimeno-Segovia, E. Johnston, K. Kieling, N. Nickerson, M. Pant, *et al.*, "Fusion-based quantum computation," *Nature Communications*, vol. 14, no. 1, p. 912, 2023.

[160] R. Simon and N. Mukunda, "Universal su (2) gadget for polarization optics," *Physics Letters A*, vol. 138, no. 9, pp. 474–480, 1989.

[161] R. Simon and N. Mukunda, "Minimal three-component su (2) gadget for polarization optics," *Physics Letters A*, vol. 143, no. 4-5, pp. 165–169, 1990.

[162] P. Kurzyński and A. Wójcik, "Quantum walk as a generalized measuring device," *Physical review letters*, vol. 110, no. 20, p. 200404, 2013.

[163] Z. Bian, J. Li, H. Qin, X. Zhan, R. Zhang, B. C. Sanders, and P. Xue, "Realization of single-qubit positive-operator-valued measurement via a one-dimensional photonic quantum walk," *Physical review letters*, vol. 114, no. 20, p. 203602, 2015.

[164] Y.-y. Zhao, N.-k. Yu, P. Kurzyński, G.-y. Xiang, C.-F. Li, and G.-C. Guo, "Experimental realization of generalized qubit measurements based on quantum walks," *Physical Review A*, vol. 91, no. 4, p. 042101, 2015.

[165] Z. Hou, J.-F. Tang, J. Shang, H. Zhu, J. Li, Y. Yuan, K.-D. Wu, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, "Deterministic realization of collective measurements via photonic quantum walks," *Nature communications*, vol. 9, no. 1, pp. 1–7, 2018.

[166] T. C. Ralph, N. K. Langford, T. Bell, and A. White, "Linear optical controlled-not gate in the coincidence basis," *Physical Review A*, vol. 65, no. 6, p. 062324, 2002.

[167] N. Kiesel, C. Schmid, U. Weber, R. Ursin, and H. Weinfurter, "Linear optics controlled-phase gate made simple," *Physical review letters*, vol. 95, no. 21, p. 210505, 2005.

[168] K. Kieling, J. O'Brien, and J. Eisert, "On photonic controlled phase gates," *New Journal of Physics*, vol. 12, no. 1, p. 013003, 2010.

[169] M. Michler, K. Mattle, H. Weinfurter, and A. Zeilinger, "Interferometric bell-state analysis," *Physical Review A*, vol. 53, no. 3, p. R1209, 1996.

[170] N. Lütkenhaus, J. Calsamiglia, and K.-A. Suominen, "Bell measurements for teleportation," *Physical Review A*, vol. 59, no. 5, p. 3295, 1999.

[171] J. Calsamiglia and N. Lütkenhaus, "Maximum efficiency of a linear-optical bell-state analyzer," *Applied Physics B*, vol. 72, no. 1, pp. 67–71, 2001.

[172] F. Ewert and P. van Loock, "3/4-efficient bell measurement with passive linear optics and unentangled ancillae," *Physical review letters*, vol. 113, no. 14, p. 140403, 2014.

[173] T. Pittman, B. Jacobs, and J. Franson, "Probabilistic quantum logic operations using polarizing beam splitters," *Physical Review A*, vol. 64, no. 6, p. 062311, 2001.

[174] S. Gasparoni, J.-W. Pan, P. Walther, T. Rudolph, and A. Zeilinger, "Realization of a photonic controlled-not gate sufficient for quantum computation," *Physical review letters*, vol. 93, no. 2, p. 020504, 2004.

[175] Z. Zhao, A.-N. Zhang, Y.-A. Chen, H. Zhang, J.-F. Du, T. Yang, and J.-W. Pan, "Experimental demonstration of a nondestructive controlled-not quantum gate for two independent photon qubits," *Physical review letters*, vol. 94, no. 3, p. 030501, 2005.

[176] J. Zeuner, A. N. Sharma, M. Tillmann, R. Heilmann, M. Gräfe, A. Moqanaki, A. Szameit, and P. Walther, "Integrated-optics heralded controlled-not gate for polarization-encoded qubits," *npj Quantum Information*, vol. 4, no. 1, pp. 1–7, 2018.

[177] D. E. Browne and T. Rudolph, "Resource-efficient linear optical quantum computation," *Physical Review Letters*, vol. 95, no. 1, p. 010501, 2005.

[178] K. Kieling, T. Rudolph, and J. Eisert, "Percolation, renormalization, and quantum computing with nondeterministic gates," *Physical Review Letters*, vol. 99, no. 13, p. 130501, 2007.

[179] M. Gimeno-Segovia, P. Shadbolt, D. E. Browne, and T. Rudolph, "From three-photon greenberger-horne-zeilinger states to ballistic universal quantum computation," *Physical review letters*, vol. 115, no. 2, p. 020502, 2015.

[180] M. Pant, D. Towsley, D. Englund, and S. Guha, "Percolation thresholds for photonic quantum computing," *Nature communications*, vol. 10, no. 1, pp. 1–11, 2019.

[181] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, "Invited review article: Single-photon sources and detectors," *Review of scientific instruments*, vol. 82, no. 7, p. 071101, 2011.

[182] M. Oxborrow and A. G. Sinclair, "Single-photon sources," *Contemporary Physics*, vol. 46, no. 3, pp. 173–206, 2005.

[183] B. Lounis and M. Orrit, "Single-photon sources," *Reports on Progress in Physics*, vol. 68, no. 5, p. 1129, 2005.

[184] A. Migdall, S. V. Polyakov, J. Fan, and J. C. Bienfang, *Single-photon generation and detection: physics and applications*. Academic Press, 2013.

[185] P. Senellart, G. Solomon, and A. White, "High-performance semiconductor quantum-dot single-photon sources," *Nature nanotechnology*, vol. 12, no. 11, pp. 1026–1039, 2017.

[186] A. Kuhn, M. Hennrich, and G. Rempe, "Deterministic single-photon source for distributed quantum networking," *Physical review letters*, vol. 89, no. 6, p. 067901, 2002.

[187] M. Hijlkema, B. Weber, H. P. Specht, S. C. Webster, A. Kuhn, and G. Rempe, "A single-photon server with just one atom," *Nature Physics*, vol. 3, no. 4, pp. 253–255, 2007.

[188] M. Keller, B. Lange, K. Hayasaka, W. Lange, and H. Walther, "A calcium ion in a cavity as a controlled single-photon source," *New Journal of Physics*, vol. 6, no. 1, p. 95, 2004.

[189] H. Barros, A. Stute, T. Northup, C. Russo, P. Schmidt, and R. Blatt, "Deterministic single-photon source from a single ion," *New Journal of Physics*, vol. 11, no. 10, p. 103004, 2009.

[190] M. Steiner, A. Hartschuh, R. Korlacki, and A. J. Meixner, "Highly efficient, tunable single photon source based on single molecules," *Applied physics letters*, vol. 90, no. 18, p. 183122, 2007.

[191] L. Zhang, Y.-J. Yu, L.-G. Chen, Y. Luo, B. Yang, F.-F. Kong, G. Chen, Y. Zhang, Q. Zhang, Y. Luo, *et al.*, "Electrically driven single-photon emission from an isolated single molecule," *Nature communications*, vol. 8, no. 1, pp. 1–7, 2017.

[192] C. Kurtsiefer, S. Mayer, P. Zarda, and H. Weinfurter, "Stable solid-state source of single photons," *Physical review letters*, vol. 85, no. 2, p. 290, 2000.

[193] H. Bernien, L. Childress, L. Robledo, M. Markham, D. Twitchen, and R. Hanson, "Two-photon quantum interference from separate nitrogen vacancy centers in diamond," *Physical Review Letters*, vol. 108, no. 4, p. 043604, 2012.

[194] S. Castelletto, B. Johnson, V. Ivády, N. Stavrias, T. Umeda, A. Gali, and T. Ohshima, "A silicon carbide room-temperature single-photon source," *Nature materials*, vol. 13, no. 2, pp. 151–156, 2014.

[195] A. Lohrmann, B. Johnson, J. McCallum, and S. Castelletto, "A review on single photon sources in silicon carbide," *Reports on Progress in Physics*, vol. 80, no. 3, p. 034502, 2017.

[196] N. Somaschi, V. Giesz, L. De Santis, J. Loredo, M. P. Almeida, G. Hornecker, S. L. Portalupi, T. Grange, C. Anton, J. Demory, *et al.*, "Near-optimal single-photon sources in the solid state," *Nature Photonics*, vol. 10, no. 5, pp. 340–345, 2016.

[197] X. Ding, Y. He, Z.-C. Duan, N. Gregersen, M.-C. Chen, S. Unsleber, S. Maier, C. Schneider, M. Kamp, S. Höfling, *et al.*, "On-demand single photons with high extraction efficiency and near-unity indistinguishability from a resonantly driven quantum dot in a micropillar," *Physical review letters*, vol. 116, no. 2, p. 020401, 2016.

[198] J. Liu, R. Su, Y. Wei, B. Yao, S. F. C. d. Silva, Y. Yu, J. Iles-Smith, K. Srinivasan, A. Rastelli, J. Li, *et al.*, "A solid-state source of strongly entangled photon pairs with high brightness and indistinguishability," *Nature nanotechnology*, vol. 14, no. 6, pp. 586–593, 2019.

[199] N. Tomm, A. Javadi, N. O. Antoniadis, D. Najer, M. C. Löbl, A. R. Korsch, R. Schott, S. R. Valentin, A. D. Wieck, A. Ludwig, *et al.*, "A bright and fast source of coherent single photons," *Nature Nanotechnology*, vol. 16, no. 4, pp. 399–403, 2021.

[200] A. Kuhn and D. Ljunggren, "Cavity-based single-photon sources," *Contemporary Physics*, vol. 51, no. 4, pp. 289–313, 2010.

[201] P. Lodahl, S. Mahmoodian, and S. Stobbe, "Interfacing single photons and single quantum dots with photonic nanostructures," *Reviews of Modern Physics*, vol. 87, no. 2, p. 347, 2015.

[202] R. Byer and S. Harris, "Power and bandwidth of spontaneous parametric emission," *Physical Review*, vol. 168, no. 3, p. 1064, 1968.

[203] D. C. Burnham and D. L. Weinberg, "Observation of simultaneity in parametric production of optical photon pairs," *Physical Review Letters*, vol. 25, no. 2, p. 84, 1970.

[204] Y. Shih and C. O. Alley, "New type of einstein-podolsky-rosen-bohm experiment using pairs of light quanta produced by optical parametric down conversion," *Physical Review Letters*, vol. 61, no. 26, p. 2921, 1988.

[205] M. Hillery, "An introduction to the quantum theory of nonlinear optics," *arXiv preprint arXiv:0901.3439*, 2009.

[206] J. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, "Interactions between light waves in a nonlinear dielectric," *Physical review*, vol. 127, no. 6, p. 1918, 1962.

[207] R. W. Boyd, *Nonlinear optics*. Elsevier, 2003.

[208] P. E. Powers and J. W. Haus, *Fundamentals of nonlinear optics*. CRC press, 2017.

[209] F. A. Laudenbach, "Engineering spectrally pure quantum states with spdc using periodically poled crystals and pulsed laser sources," Master's thesis, University of Vienna, 2015.

[210] L. Mandel and E. Wolf, *Optical coherence and quantum optics*. Cambridge university press, 1995.

[211] T. Pittman, B. Jacobs, and J. Franson, "Single photons on pseudodemand from stored parametric down-conversion," *Physical Review A*, vol. 66, no. 4, p. 042303, 2002.

[212] A. L. Migdall, D. Branning, and S. Castelletto, "Tailoring single-photon and multiphoton probabilities of a single-photon on-demand source," *Physical Review A*, vol. 66, no. 5, p. 053805, 2002.

[213] F. Kaneda and P. G. Kwiat, "High-efficiency single-photon generation via large-scale active time multiplexing," *Science advances*, vol. 5, no. 10, p. eaaw8586, 2019.

[214] J. D. Jackson, *Classical electrodynamics*. American Association of Physics Teachers, 1999.

[215] D. J. Griffiths, "Introduction to electrodynamics," 2005.

[216] N. Quesada and J. Sipe, "Why you should not use the electric field to quantize in nonlinear optics," *Optics letters*, vol. 42, no. 17, pp. 3443–3446, 2017.

[217] M. Hillery and L. D. Mlodinow, "Quantization of electrodynamics in nonlinear dielectric media," *Physical Review A*, vol. 30, no. 4, p. 1860, 1984.

[218] Y. Shen, "Quantum statistics of nonlinear optics," *Physical Review*, vol. 155, no. 3, p. 921, 1967.

[219] J. Tucker and D. F. Walls, "Quantum theory of the traveling-wave frequency converter," *Physical Review*, vol. 178, no. 5, p. 2036, 1969.

[220] P. Mosley and P. Mosley, *Generation of heralded single photons in pure quantum states*. PhD thesis, University of Oxford, 2007.

[221] A. B. U'Ren, *Multi-photon state engineering for quantum information processing applications*. PhD thesis, 2004.

[222] K. Blow, R. Loudon, S. J. Phoenix, and T. Shepherd, "Continuum fields in quantum optics," *Physical Review A*, vol. 42, no. 7, p. 4102, 1990.

[223] G. D. Mahan, *Quantum mechanics in a nutshell*. Princeton University Press, 2009.

[224] A. Christ, B. Brecht, W. Mauerer, and C. Silberhorn, "Theory of quantum frequency conversion and type-ii parametric down-conversion in the high-gain regime," *New Journal of Physics*, vol. 15, no. 5, p. 053038, 2013.

[225] N. Quesada and J. Sipe, "Time-ordering effects in the generation of entangled photons using nonlinear optical processes," *Physical review letters*, vol. 114, no. 9, p. 093903, 2015.

[226] Z.-Y. J. Ou, *Multi-photon quantum interference*, vol. 43. Springer, 2007.

[227] D. S. Hum and M. M. Fejer, "Quasi-phasematching," *Comptes Rendus Physique*, vol. 8, no. 2, pp. 180–198, 2007.

[228] F. Steinlechner, M. Gilaberte, M. Jofre, T. Scheidl, J. P. Torres, V. Pruneri, and R. Ursin, "Efficient heralding of polarization-entangled photons from type-0 and type-ii spontaneous parametric downconversion in periodically poled ktiopo 4," *JOSA B*, vol. 31, no. 9, pp. 2068–2076, 2014.

[229] S. P. Neumann, M. Selimovic, M. Bohmann, and R. Ursin, "Experimental entanglement generation for quantum key distribution beyond 1 gbit/s," *arXiv preprint arXiv:2107.07756*, 2021.

[230] A. Anwar, C. Perumangatt, F. Steinlechner, T. Jennewein, and A. Ling, "Entangled photon-pair sources based on three-wave mixing in bulk crystals," *Review of Scientific Instruments*, vol. 92, no. 4, p. 041101, 2021.

[231] P. J. Mosley, J. S. Lundeen, B. J. Smith, and I. A. Walmsley, "Conditional preparation of single photons using parametric downconversion: a recipe for purity," *New Journal of Physics*, vol. 10, no. 9, p. 093011, 2008.

[232] F. Laudenbach, H. Hübel, M. Hentschel, P. Walther, and A. Poppe, "Modelling parametric down-conversion yielding spectrally pure photon pairs," *Optics express*, vol. 24, no. 3, pp. 2712–2727, 2016.

[233] R.-B. Jin, R. Shimizu, K. Wakui, H. Benichi, and M. Sasaki, "Widely tunable single photon source with high purity at telecom wavelength," *Optics express*, vol. 21, no. 9, pp. 10659–10666, 2013.

[234] S. Merkouche, V. Thiel, and B. J. Smith, "Spectrally resolved four-photon interference of time-frequency-entangled photons," *Physical Review A*, vol. 105, no. 2, p. 023708, 2022.

[235] S. Merkouche, V. Thiel, A. O. Davis, and B. J. Smith, "Heralding multiple photonic pulsed bell pairs via frequency-resolved entanglement swapping," *Physical Review Letters*, vol. 128, no. 6, p. 063602, 2022.

[236] A. M. Brańczyk, A. Fedrizzi, T. M. Stace, T. C. Ralph, and A. G. White, "Engineered optical nonlinearity for quantum light sources," *Optics express*, vol. 19, no. 1, pp. 55–65, 2011.

[237] A. Dosseva, Ł. Cincio, and A. M. Brańczyk, "Shaping the joint spectrum of down-converted photons through optimized custom poling," *Physical Review A*, vol. 93, no. 1, p. 013801, 2016.

[238] F. Graffitti, D. Kundys, D. T. Reid, A. M. Brańczyk, and A. Fedrizzi, "Pure down-conversion photons through sub-coherence-

length domain engineering," *Quantum Science and Technology*, vol. 2, no. 3, p. 035001, 2017.

[239] C. Chen, C. Bo, M. Y. Niu, F. Xu, Z. Zhang, J. H. Shapiro, and F. N. Wong, "Efficient generation and characterization of spectrally factorable biphotons," *Optics express*, vol. 25, no. 7, pp. 7300–7312, 2017.

[240] A. Pickston, F. Graffitti, P. Barrow, C. L. Morrison, J. Ho, A. M. Brańczyk, and A. Fedrizzi, "Optimised domain-engineered crystals for pure telecom photon sources," *Optics Express*, vol. 29, no. 5, pp. 6991–7002, 2021.

[241] Y. Chen, S. Ecker, L. Chen, F. Steinlechner, M. Huber, and R. Ursin, "Temporal distinguishability in hong-ou-mandel interference for harnessing high-dimensional frequency entanglement," *npj Quantum Information*, vol. 7, no. 1, pp. 1–7, 2021.

[242] S. Ecker, P. Sohr, L. Bulla, M. Huber, M. Bohmann, and R. Ursin, "Experimental single-copy entanglement distillation," *Physical Review Letters*, vol. 127, no. 4, p. 040506, 2021.

[243] E. Bocquillon, C. Couteau, M. Razavi, R. Laflamme, and G. Weihs, "Coherence measures for heralded single-photon sources," *Physical Review A*, vol. 79, no. 3, p. 035801, 2009.

[244] M. Bashkansky, I. Vurgaftman, A. C. Pipino, and J. Reintjes, "Significance of heralding in spontaneous parametric down-conversion," *Physical Review A*, vol. 90, no. 5, p. 053825, 2014.

[245] C. Kurtsiefer, M. Oberparleiter, and H. Weinfurter, "High-efficiency entangled photon pair collection in type-ii parametric fluorescence," *Physical Review A*, vol. 64, no. 2, p. 023802, 2001.

[246] F. A. Bovino, P. Varisco, A. M. Colla, G. Castagnoli, G. Di Giuseppe, and A. V. Sergienko, "Effective fiber-coupling of entangled photons for quantum communication," *Optics Communications*, vol. 227, no. 4-6, pp. 343–348, 2003.

[247] R. Andrews, E. Pike, and S. Sarkar, "Optimal coupling of entangled photons into single-mode optical fibers," *Optics Express*, vol. 12, no. 14, pp. 3264–3269, 2004.

[248] D. Ljunggren and M. Tengner, "Optimal focusing for maximal collection of entangled narrow-band photon pairs into single-mode fibers," *Physical Review A*, vol. 72, no. 6, p. 062301, 2005.

[249] P. Kolenderski, W. Wasilewski, and K. Banaszek, "Modeling and optimization of photon pair sources based on spontaneous parametric down-conversion," *Physical Review A*, vol. 80, no. 1, p. 013811, 2009.

[250] R. S. Bennink, "Optimal collinear gaussian beams for spontaneous parametric down-conversion," *Physical Review A*, vol. 81, no. 5, p. 053805, 2010.

[251] P. B. Dixon, D. Rosenberg, V. Stelmakh, M. E. Grein, R. S. Bennink, E. A. Dauler, A. J. Kerman, R. J. Molnar, and F. N. Wong, "Heralding efficiency and correlated-mode coupling of near-ir fiber-coupled photon pairs," *Physical Review A*, vol. 90, no. 4, p. 043804, 2014.

[252] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, and Y. Shih, "New high-intensity source of polarization-entangled photon pairs," *Physical Review Letters*, vol. 75, no. 24, p. 4337, 1995.

[253] R. S. Bennink, Y. Liu, D. D. Earl, and W. P. Grice, "Spatial distinguishability of photons produced by spontaneous parametric down-conversion with a focused pump," *Physical Review A*, vol. 74, no. 2, p. 023802, 2006.

[254] A. Villar, A. Lohrmann, and A. Ling, "Experimental entangled photon pair generation using crystals with parallel optical axes," *Optics express*, vol. 26, no. 10, pp. 12396–12402, 2018.

[255] G. Weihs, *Ein Experiment zum Test der Bellschen Ungleichung unter Einsteinscher Lokalität*. na, 1999.

[256] C. Zhang, Y.-F. Huang, B.-H. Liu, C.-F. Li, and G.-C. Guo, "Spontaneous parametric down-conversion sources for multiphoton experiments," *Advanced Quantum Technologies*, vol. 4, no. 5, p. 2000132, 2021.

[257] S. Takeuchi, "Beamlike twin-photon generation by use of type ii parametric downconversion," *Optics Letters*, vol. 26, no. 11, pp. 843–845, 2001.

[258] Y.-H. Kim, "Quantum interference with beamlike type-ii spontaneous parametric down-conversion," *Physical Review A*, vol. 68, no. 1, p. 013804, 2003.

[259] O. Kwon, Y.-W. Cho, and Y.-H. Kim, "Single-mode coupling efficiencies of type-ii spontaneous parametric down-conversion:

collinear, noncollinear, and beamlike phase matching," *Physical Review A*, vol. 78, no. 5, p. 053825, 2008.

[260] X.-L. Wang, L.-K. Chen, W. Li, H.-L. Huang, C. Liu, C. Chen, Y.-H. Luo, Z.-E. Su, D. Wu, Z.-D. Li, *et al.*, "Experimental ten-photon entanglement," *Physical review letters*, vol. 117, no. 21, p. 210502, 2016.

[261] H. Kim, O. Kwon, and H. S. Moon, "Pulsed sagnac source of polarization-entangled photon pairs in telecommunication band," *Scientific reports*, vol. 9, no. 1, pp. 1–7, 2019.

[262] M. H. Rubin, D. N. Klyshko, Y. Shih, and A. Sergienko, "Theory of two-photon entanglement in type-ii optical parametric down-conversion," *Physical Review A*, vol. 50, no. 6, p. 5122, 1994.

[263] B.-S. Shi and A. Tomita, "Generation of a pulsed polarization entangled photon pair using a sagnac interferometer," *Physical Review A*, vol. 69, no. 1, p. 013803, 2004.

[264] T. Kim, M. Fiorentino, and F. N. Wong, "Phase-stable source of polarization-entangled photons using a polarization sagnac interferometer," *Physical Review A*, vol. 73, no. 1, p. 012316, 2006.

[265] A. Fedrizzi, T. Herbst, A. Poppe, T. Jennewein, and A. Zeilinger, "A wavelength-tunable fiber-coupled source of narrowband entangled photons," *Optics Express*, vol. 15, no. 23, pp. 15377–15386, 2007.

[266] Y. Chen, S. Ecker, S. Wengerowsky, L. Bulla, S. K. Joshi, F. Steinlechner, and R. Ursin, "Polarization entanglement by time-reversed hong-ou-mandel interference," *Physical Review Letters*, vol. 121, no. 20, p. 200502, 2018.

[267] T. E. Stuart, J. A. Slater, F. Bussières, and W. Tittel, "Flexible source of nondegenerate entangled photons based on a two-crystal sagnac interferometer," *Physical Review A*, vol. 88, no. 1, p. 012301, 2013.

[268] M. Chekhova, O. Ivanova, V. Berardi, and A. Garuccio, "Spectral properties of three-photon entangled states generated via three-photon parametric down-conversion in a χ (3) medium," *Physical Review A*, vol. 72, no. 2, p. 023818, 2005.

[269] A. Dot, A. Borne, B. Boulanger, K. Bencheikh, and J. A. Levenson, "Quantum theory analysis of triple photons generated by a χ (3) process," *Physical Review A*, vol. 85, no. 2, p. 023809, 2012.

[270] C. Okoth, A. Cavanna, N. Y. Joly, and M. V. Chekhova, "Seeded and unseeded high-order parametric down-conversion," *Physical Review A*, vol. 99, no. 4, p. 043809, 2019.

[271] F. A. Domínguez-Serna, A. B. U'Ren, and K. Garay-Palmett, "Third-order parametric down-conversion: A stimulated approach," *Physical Review A*, vol. 101, no. 3, p. 033813, 2020.

[272] A. Cavanna, J. Hammer, C. Okoth, E. Ortiz-Ricardo, H. Cruz-Ramirez, K. Garay-Palmett, A. B. U'Ren, M. H. Frosz, X. Jiang, N. Y. Joly, *et al.*, "Progress toward third-order parametric down-conversion in optical fibers," *Physical Review A*, vol. 101, no. 3, p. 033840, 2020.

[273] M. Banic, M. Liscidini, and J. Sipe, "Resonant and non-resonant integrated third order parametric down-conversion," *arXiv preprint arXiv:2204.09159*, 2022.

[274] I. Alonso Calafell, L. A. Rozema, D. Alcaraz Iranzo, A. Trenti, P. K. Jenke, J. D. Cox, A. Kumar, H. Bieliaiev, S. Nanot, C. Peng, *et al.*, "Giant enhancement of third-harmonic generation in graphene–metal heterostructures," *Nature Nanotechnology*, vol. 16, no. 3, pp. 318–324, 2021.

[275] H. Hübel, D. R. Hamel, A. Fedrizzi, S. Ramelow, K. J. Resch, and T. Jennewein, "Direct generation of photon triplets using cascaded photon-pair sources," *Nature*, vol. 466, no. 7306, pp. 601–603, 2010.

[276] D. R. Hamel, L. K. Shalm, H. Hübel, A. J. Miller, F. Marsili, V. B. Verma, R. P. Mirin, S. W. Nam, K. J. Resch, and T. Jennewein, "Direct generation of three-photon polarization entanglement," *Nature Photonics*, vol. 8, no. 10, pp. 801–807, 2014.

[277] P. Walther, M. Aspelmeyer, and A. Zeilinger, "Heralded generation of multiphoton entanglement," *Physical Review A*, vol. 75, no. 1, p. 012313, 2007.

[278] C. Śliwa and K. Banaszek, "Conditional preparation of maximal polarization entanglement," *Physical Review A*, vol. 67, no. 3, p. 030101, 2003.

[279] S. Barz, G. Cronenberg, A. Zeilinger, and P. Walther, "Heralded generation of entangled photon pairs," *Nature photonics*, vol. 4, no. 8, pp. 553–556, 2010.

[280] X. Zou, L. J. Wang, and L. Mandel, "Induced coherence and indistinguishability in optical

interference," *Physical review letters*, vol. 67, no. 3, p. 318, 1991.

[281] M. Krenn, A. Hochrainer, M. Lahiri, and A. Zeilinger, "Entanglement by path identity," *Physical review letters*, vol. 118, no. 8, p. 080401, 2017.

[282] A. Hochrainer, M. Lahiri, M. Erhard, M. Krenn, and A. Zeilinger, "Quantum indistinguishability by path identity and with undetected photons," *Reviews of Modern Physics*, vol. 94, no. 2, p. 025007, 2022.

[283] T. Feng, X. Zhang, Y. Tian, and Q. Feng, "On-chip multiphoton entangled states by path identity," *International Journal of Theoretical Physics*, vol. 58, no. 11, pp. 3726–3733, 2019.

[284] J. Kysela, M. Erhard, A. Hochrainer, M. Krenn, and A. Zeilinger, "Path identity as a source of high-dimensional entanglement," *Proceedings of the National Academy of Sciences*, vol. 117, no. 42, pp. 26118–26122, 2020.

[285] X. Gu, L. Chen, and M. Krenn, "Quantum experiments and hypergraphs: Multiphoton sources for quantum interference, quantum computation, and quantum entanglement," *Physical Review A*, vol. 101, no. 3, p. 033816, 2020.

[286] T. Herzog, J. Rarity, H. Weinfurter, and A. Zeilinger, "Frustrated two-photon creation via interference," *Physical review letters*, vol. 72, no. 5, p. 629, 1994.

[287] A. M. Brańczyk, "Hong-ou-mandel interference," *arXiv preprint arXiv:1711.00080*, 2017.

[288] A. Lyons, G. C. Knee, E. Bolduc, T. Roger, J. Leach, E. M. Gauger, and D. Faccio, "Attosecond-resolution hong-ou-mandel interferometry," *Science advances*, vol. 4, no. 5, p. eaap9416, 2018.

[289] A. J. Brady and S. Haldar, "Frame dragging and the hong-ou-mandel dip: Gravitational effects in multiphoton interference," *Physical Review Research*, vol. 3, no. 2, p. 023024, 2021.

[290] A. M. Steinberg, P. G. Kwiat, and R. Y. Chiao, "Dispersion cancellation and high-resolution time measurements in a fourth-order optical interferometer," *Physical Review A*, vol. 45, no. 9, p. 6659, 1992.

[291] C. K. Hitzenberger, A. Baumgartner, W. Drexler, and A. F. Fercher, "Dispersion effects in partial coherence interferometry: implications for intraocular ranging," *Journal of Biomedical Optics*, vol. 4, no. 1, pp. 144–151, 1999.

[292] A. F. Abouraddy, M. B. Nasr, B. E. Saleh, A. V. Sergienko, and M. C. Teich, "Quantum-optical coherence tomography with dispersion cancellation," *Physical Review A*, vol. 65, no. 5, p. 053817, 2002.

[293] M. B. Nasr, B. E. Saleh, A. V. Sergienko, and M. C. Teich, "Demonstration of dispersion-canceled quantum-optical coherence tomography," *Physical review letters*, vol. 91, no. 8, p. 083601, 2003.

[294] M. Okano, R. Okamoto, A. Tanaka, S. Ishida, N. Nishizawa, and S. Takeuchi, "Dispersion cancellation in high-resolution two-photon interference," *Physical Review A*, vol. 88, no. 4, p. 043845, 2013.

[295] M. B. Nasr, D. P. Goode, N. Nguyen, G. Rong, L. Yang, B. M. Reinhard, B. E. Saleh, and M. C. Teich, "Quantum optical coherence tomography of a biological sample," *Optics Communications*, vol. 282, no. 6, pp. 1154–1159, 2009.

[296] S. Ramelow, L. Ratschbacher, A. Fedrizzi, N. Langford, and A. Zeilinger, "Discrete tunable color entanglement," *Physical review letters*, vol. 103, no. 25, p. 253601, 2009.

[297] Y. Chen, M. Fink, F. Steinlechner, J. P. Torres, and R. Ursin, "Hong-ou-mandel interferometry on a biphoton beat note," *npj Quantum Information*, vol. 5, no. 1, pp. 1–6, 2019.

[298] G. Lifante, *Integrated photonics: fundamentals*. John Wiley & Sons, 2003.

[299] C. A. Brackett, "Dense wavelength division multiplexing networks: Principles and applications," *IEEE Journal on Selected areas in Communications*, vol. 8, no. 6, pp. 948–964, 1990.

[300] D.-K. Yang and S.-T. Wu, *Fundamentals of liquid crystal devices*. John Wiley & Sons, 2014.

[301] M. Schadt and W. Helfrich, "Voltage-dependent optical activity of a twisted nematic liquid crystal," *Applied Physics Letters*, vol. 18, no. 4, pp. 127–128, 1971.

[302] K. C. Kao and G. A. Hockham, "Dielectric-fibre surface waveguides for optical frequencies," in *Proceedings of the Institution of Electrical Engineers*, vol. 113, pp. 1151–1158, IET, 1966.

[303] J. E. Toney, *Lithium niobate photonics*. Artech House, 2015.

[304] I. Krasnokutska, J.-L. J. Tambasco, X. Li, and A. Peruzzo, "Ultra-low loss photonic circuits in lithium niobate on insulator," *Optics express*, vol. 26, no. 2, pp. 897–904, 2018.

[305] S. Y. Siew, E. J. H. Cheung, H. Liang, A. Bettiol, N. Toyoda, B. Alshehri, E. Dogheche, and A. J. Danner, "Ultra-low loss ridge waveguides on lithium niobate via argon ion milling and gas clustered ion beam smoothening," *Optics Express*, vol. 26, no. 4, pp. 4421–4430, 2018.

[306] E. L. Wooten, K. M. Kissa, A. Yi-Yan, E. J. Murphy, D. A. Lafaw, P. F. Hallemeier, D. Maack, D. V. Attanasio, D. J. Fritz, G. J. McBrien, *et al.*, "A review of lithium niobate modulators for fiber-optic communications systems," *IEEE Journal of selected topics in Quantum Electronics*, vol. 6, no. 1, pp. 69–82, 2000.

[307] C. Wang, M. Zhang, X. Chen, M. Bertrand, A. Shams-Ansari, S. Chandrasekhar, P. Winzer, and M. Lončar, "Integrated lithium niobate electro-optic modulators operating at cmos-compatible voltages," *Nature*, vol. 562, no. 7725, pp. 101–104, 2018.

[308] M. He, M. Xu, Y. Ren, J. Jian, Z. Ruan, Y. Xu, S. Gao, S. Sun, X. Wen, L. Zhou, *et al.*, "High-performance hybrid silicon and lithium niobate mach–zehnder modulators for 100 gbit s- 1 and beyond," *Nature Photonics*, vol. 13, no. 5, pp. 359–364, 2019.

[309] M. Chou, J. Hauden, M. Arbore, and M. Fejer, "1.5-μm-band wavelength conversion based on difference-frequency generation in linbo 3 waveguides with integrated coupling structures," *Optics letters*, vol. 23, no. 13, pp. 1004–1006, 1998.

[310] M. Chou, K. Parameswaran, M. M. Fejer, and I. Brener, "Multiple-channel wavelength conversion by use of engineered quasi-phase-matching structures in linbo 3 waveguides," *Optics letters*, vol. 24, no. 16, pp. 1157–1159, 1999.

[311] C. Becker, T. Oesselke, J. Pandavenes, R. Ricken, K. Rochhausen, G. Schreiber, W. Sohler, H. Suche, R. Wessel, S. Balsamo, *et al.*, "Advanced ti: Er: Linbo/sub 3/waveguide lasers," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 6, no. 1, pp. 101–113, 2000.

[312] C. Vannahme, H. Suche, S. Reza, R. Ricken, V. Quiring, and W. Sohler, "Integrated optical ti: Linbo3 ring resonator for rotation rate sensing," *Proc. 13th ECIO*, vol. 2007, 2007.

[313] M. Bazzan and C. Sada, "Optical waveguides in lithium niobate: Recent developments and applications," *Applied Physics Reviews*, vol. 2, no. 4, p. 040603, 2015.

[314] G. Poberaj, H. Hu, W. Sohler, and P. Guenter, "Lithium niobate on insulator (lnoi) for micro-photonic devices," *Laser & photonics reviews*, vol. 6, no. 4, pp. 488–503, 2012.

[315] L. Cai, Y. Wang, and H. Hu, "Low-loss waveguides in a single-crystal lithium niobate thin film," *Optics letters*, vol. 40, no. 13, pp. 3013–3016, 2015.

[316] R. A. Logan and F. K. Reinhart, "Optical waveguides in gaas–algaas epitaxial layers," *Journal of Applied Physics*, vol. 44, no. 9, pp. 4172–4176, 1973.

[317] C. P. Dietrich, A. Fiore, M. G. Thompson, M. Kamp, and S. Höfling, "Gaas integrated quantum photonics: Towards compact and multi-functional quantum photonic integrated circuits," *Laser & Photonics Reviews*, vol. 10, no. 6, pp. 870–894, 2016.

[318] J. Wang, A. Santamato, P. Jiang, D. Bonneau, E. Engin, J. W. Silverstone, M. Lermer, J. Beetz, M. Kamp, S. Höfling, *et al.*, "Gallium arsenide (gaas) quantum photonic waveguide circuits," *Optics Communications*, vol. 327, pp. 49–55, 2014.

[319] R. Baets, A. Z. Subramanian, S. Clemmen, B. Kuyken, P. Bienstman, N. Le Thomas, G. Roelkens, D. Van Thourhout, P. Helin, and S. Severi, "Silicon photonics: silicon nitride versus silicon-on-insulator," in *Optical Fiber Communication Conference*, pp. Th3J–1, Optical Society of America, 2016.

[320] D. J. Blumenthal, R. Heideman, D. Geuzebroek, A. Leinse, and C. Roeloffzen, "Silicon nitride in silicon photonics," *Proceedings of the IEEE*, vol. 106, no. 12, pp. 2209–2231, 2018.

[321] B. Jalali and S. Fathpour, "Silicon photonics," *Journal of lightwave technology*, vol. 24, no. 12, pp. 4600–4615, 2006.

[322] R. Soref, "The past, present, and future of silicon photonics," *IEEE Journal of selected topics in quantum electronics*, vol. 12, no. 6, pp. 1678–1687, 2006.

[323] J. Leuthold, C. Koos, and W. Freude, "Nonlinear silicon photonics," *Nature photonics*, vol. 4, no. 8, pp. 535–544, 2010.

[324] K. M. Davis, K. Miura, N. Sugimoto, and K. Hirao, "Writing waveguides in glass with a femtosecond laser," *Optics letters*, vol. 21, no. 21, pp. 1729–1731, 1996.

[325] K. Miura, J. Qiu, H. Inouye, T. Mitsuyu, and K. Hirao, "Photowritten optical waveguides in various glasses with ultrashort pulse laser," *Applied Physics Letters*, vol. 71, no. 23, pp. 3329–3331, 1997.

[326] C. B. Schaffer, A. Brodeur, J. F. García, and E. Mazur, "Micromachining bulk glass by use of femtosecond laser pulses with nanojoule energy," *Optics letters*, vol. 26, no. 2, pp. 93–95, 2001.

[327] W. Bogaerts, R. Baets, P. Dumon, V. Wiaux, S. Beckx, D. Taillaert, B. Luyssaert, J. Van Campenhout, P. Bienstman, and D. Van Thourhout, "Nanophotonic waveguides in silicon-on-insulator fabricated with cmos technology," *Journal of Lightwave Technology*, vol. 23, no. 1, p. 401, 2005.

[328] J. Cardenas, C. B. Poitras, J. T. Robinson, K. Preston, L. Chen, and M. Lipson, "Low loss etchless silicon photonic waveguides," *Optics express*, vol. 17, no. 6, pp. 4752–4757, 2009.

[329] W. Bogaerts, S. K. Selvaraja, P. Dumon, J. Brouckaert, K. De Vos, D. Van Thourhout, and R. Baets, "Silicon-on-insulator spectral filters fabricated with cmos technology," *IEEE journal of selected topics in quantum electronics*, vol. 16, no. 1, pp. 33–44, 2010.

[330] D.-X. Xu, J. H. Schmid, G. T. Reed, G. Z. Mashanovich, D. J. Thomson, M. Nedeljkovic, X. Chen, D. Van Thourhout, S. Keyvaninia, and S. K. Selvaraja, "Silicon photonic integration platform—have we found the sweet spot?," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 4, pp. 189–205, 2014.

[331] M. R. Watts, "Adiabatic microring resonators," *Optics Letters*, vol. 35, no. 19, pp. 3231–3233, 2010.

[332] R. Soref and B. Bennett, "Electrooptical effects in silicon," *IEEE journal of quantum electronics*, vol. 23, no. 1, pp. 123–129, 1987.

[333] H. Sun, Q. Qiao, Q. Guan, and G. Zhou, "Silicon photonic phase shifters and their applications: A review," *Micromachines*, vol. 13, no. 9, p. 1509, 2022.

[334] G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. Thomson, "Silicon optical modulators," *Nature photonics*, vol. 4, no. 8, pp. 518–526, 2010.

[335] N. C. Harris, Y. Ma, J. Mower, T. Baehr-Jones, D. Englund, M. Hochberg, and C. Galland, "Efficient, compact and low loss thermo-optic phase shifter in silicon," *Optics express*, vol. 22, no. 9, pp. 10487–10493, 2014.

[336] S. Liu, J. Feng, Y. Tian, H. Zhao, L. Jin, B. Ouyang, J. Zhu, and J. Guo, "Thermo-optic phase shifters based on silicon-on-insulator platform: State-of-the-art and a review," *Frontiers of Optoelectronics*, vol. 15, no. 1, pp. 1–21, 2022.

[337] A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu, and M. Paniccia, "A high-speed silicon optical modulator based on a metal–oxide–semiconductor capacitor," *Nature*, vol. 427, no. 6975, pp. 615–618, 2004.

[338] L. Liao, D. Samara-Rubio, M. Morse, A. Liu, D. Hodge, D. Rubin, U. D. Keil, and T. Franck, "High speed silicon mach-zehnder modulator," *Optics express*, vol. 13, no. 8, pp. 3129–3135, 2005.

[339] L. Liao, A. Liu, D. Rubin, J. Basak, Y. Chetrit, H. Nguyen, R. Cohen, N. Izhaky, and M. Paniccia, "40 gbit/s silicon optical modulator for high-speed applications," *Electronics letters*, vol. 43, no. 22, pp. 1196–1197, 2007.

[340] T.-K. Liang and H. K. Tsang, "Role of free carriers from two-photon absorption in raman amplification in silicon-on-insulator waveguides," *Applied physics letters*, vol. 84, no. 15, pp. 2745–2747, 2004.

[341] C. A. Husko, A. S. Clark, M. J. Collins, A. De Rossi, S. Combrié, G. Lehoucq, I. H. Rey, T. F. Krauss, C. Xiong, and B. J. Eggleton, "Multi-photon absorption limits to heralded single photon sources," *Scientific reports*, vol. 3, no. 1, pp. 1–8, 2013.

[342] S. Tanzilli, A. Martin, F. Kaiser, M. P. De Micheli, O. Alibart, and D. B. Ostrowsky, "On the genesis and evolution of integrated quantum optics," *Laser & Photonics Reviews*, vol. 6, no. 1, pp. 115–143, 2012.

[343] J. W. Silverstone, J. Wang, D. Bonneau, P. Sibson, R. Santagati, C. Erven, J. O'Brien, and M. Thompson, "Silicon quantum photonics," in *2016 International Con-*

ference on Optical MEMS and Nanophotonics (OMN), pp. 1–2, IEEE, 2016.

[344] D. Thomson, A. Zilkie, J. E. Bowers, T. Komljenovic, G. T. Reed, L. Vivien, D. Marris-Morini, E. Cassan, L. Virot, J.-M. Fédéli, *et al.*, "Roadmap on silicon photonics," *Journal of Optics*, vol. 18, no. 7, p. 073003, 2016.

[345] A. W. Elshaari, W. Pernice, K. Srinivasan, O. Benson, and V. Zwiller, "Hybrid integrated quantum photonic circuits," *Nature Photonics*, vol. 14, no. 5, pp. 285–298, 2020.

[346] J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, "Integrated photonic quantum technologies," *Nature Photonics*, vol. 14, no. 5, pp. 273–284, 2020.

[347] G. Moody, V. J. Sorger, D. J. Blumenthal, P. W. Juodawlkis, W. Loh, C. Sorace-Agaskar, A. E. Jones, K. C. Balram, J. C. Matthews, A. Laing, *et al.*, "2022 roadmap on integrated quantum photonics," *Journal of Physics: Photonics*, vol. 4, no. 1, p. 012501, 2022.

[348] A. Peruzzo, M. Lobino, J. C. Matthews, N. Matsuda, A. Politi, K. Poulios, X.-Q. Zhou, Y. Lahini, N. Ismail, K. Wörhoff, *et al.*, "Quantum walks of correlated photons," *Science*, vol. 329, no. 5998, pp. 1500–1503, 2010.

[349] A. Crespi, R. Osellame, R. Ramponi, V. Giovannetti, R. Fazio, L. Sansoni, F. De Nicola, F. Sciarrino, and P. Mataloni, "Anderson localization of entangled photons in an integrated quantum walk," *Nature Photonics*, vol. 7, no. 4, pp. 322–328, 2013.

[350] F. Caruso, A. Crespi, A. G. Ciriolo, F. Sciarrino, and R. Osellame, "Fast escape of a quantum walker from an integrated photonic maze," *Nature communications*, vol. 7, no. 1, pp. 1–7, 2016.

[351] J. Wang, S. Paesani, Y. Ding, R. Santagati, P. Skrzypczyk, A. Salavrakos, J. Tura, R. Augusiak, L. Mančinska, D. Bacco, *et al.*, "Multidimensional quantum entanglement with large-scale integrated optics," *Science*, vol. 360, no. 6386, pp. 285–291, 2018.

[352] A. Javadi, I. Söllner, M. Arcari, S. L. Hansen, L. Midolo, S. Mahmoodian, G. Kiršanskė, T. Pregnolato, E. Lee, J. Song, *et al.*, "Single-photon non-linear optics with a quantum dot in a waveguide," *Nature communications*, vol. 6, no. 1, pp. 1–5, 2015.

[353] S. Sun, H. Kim, Z. Luo, G. S. Solomon, and E. Waks, "A single-photon switch and transistor enabled by a solid-state quantum memory," *Science*, vol. 361, no. 6397, pp. 57–60, 2018.

[354] H. L. Jeannic, A. Tiranov, J. Carolan, T. Ramos, Y. Wang, M. H. Appel, S. Scholz, A. D. Wieck, A. Ludwig, N. Rotenberg, *et al.*, "Dynamical photon–photon interaction mediated by a quantum emitter," *Nature Physics*, pp. 1–5, 2022.

[355] S. Liu, O. A. D. Sandberg, M. L. Chan, B. Schrinski, Y. Anyfantaki, R. B. Nielsen, R. G. Larsen, A. Skalkin, Y. Wang, L. Midolo, *et al.*, "Violation of bell inequality by photon scattering on a two-level emitter," *arXiv preprint arXiv:2306.12801*, 2023.

[356] R. Horn, P. Abolghasem, B. J. Bijlani, D. Kang, A. Helmy, and G. Weihs, "Monolithic source of photon pairs," *Physical review letters*, vol. 108, no. 15, p. 153605, 2012.

[357] E. Engin, D. Bonneau, C. M. Natarajan, A. S. Clark, M. G. Tanner, R. H. Hadfield, S. N. Dorenbos, V. Zwiller, K. Ohira, N. Suzuki, *et al.*, "Photon pair generation in a silicon micro-ring resonator with reverse bias enhancement," *Optics express*, vol. 21, no. 23, pp. 27826–27834, 2013.

[358] R. S. Daveau, K. C. Balram, T. Pregnolato, J. Liu, E. H. Lee, J. D. Song, V. Verma, R. Mirin, S. W. Nam, L. Midolo, *et al.*, "Efficient fiber-coupled single-photon source based on quantum dots in a photonic-crystal waveguide," *Optica*, vol. 4, no. 2, pp. 178–184, 2017.

[359] J. Benedikter, H. Kaupp, T. Hümmer, Y. Liang, A. Bommer, C. Becher, A. Krueger, J. M. Smith, T. W. Hänsch, and D. Hunger, "Cavity-enhanced single-photon source based on the silicon-vacancy center in diamond," *Physical Review Applied*, vol. 7, no. 2, p. 024031, 2017.

[360] M. K. Bhaskar, D. D. Sukachev, A. Sipahigil, R. E. Evans, M. J. Burek, C. T. Nguyen, L. J. Rogers, P. Siyushev, M. H. Metsch, H. Park, *et al.*, "Quantum nonlinear optics with a germanium-vacancy color center in a nanoscale diamond waveguide," *Physical review letters*, vol. 118, no. 22, p. 223603, 2017.

[361] T. Kashiwazaki, N. Takanashi, T. Yamashima, T. Kazama, K. Enbutsu, R. Kasahara, T. Umeki, and A. Furusawa, "Continuous-wave 6-db-squeezed light

with 2.5-thz-bandwidth from single-mode ppln waveguide," *APL Photonics*, vol. 5, no. 3, p. 036104, 2020.

[362] N. J. Martinez, M. Gehl, C. T. Derose, A. L. Starbuck, A. T. Pomerene, A. L. Lentine, D. C. Trotter, and P. S. Davids, "Single photon detection in a waveguide-coupled ge-on-si lateral avalanche photodiode," *Optics express*, vol. 25, no. 14, pp. 16130–16139, 2017.

[363] W. H. Pernice, C. Schuck, O. Minaeva, M. Li, G. Goltsman, A. Sergienko, and H. Tang, "High-speed and high-efficiency travelling wave single-photon detectors embedded in nanophotonic circuits," *Nature communications*, vol. 3, no. 1, pp. 1–10, 2012.

[364] T. Gerrits, N. Thomas-Peter, J. C. Gates, A. E. Lita, B. J. Metcalf, B. Calkins, N. A. Tomlin, A. E. Fox, A. L. Linares, J. B. Spring, *et al.*, "On-chip, photon-number-resolving, telecommunication-band detectors for scalable photonic information processing," *Physical Review A*, vol. 84, no. 6, p. 060301, 2011.

[365] L. Sansoni, F. Sciarrino, G. Vallone, P. Mataloni, A. Crespi, R. Ramponi, and R. Osellame, "Polarization entangled state measurement on a chip," *Physical review letters*, vol. 105, no. 20, p. 200503, 2010.

[366] A. Crespi, R. Ramponi, R. Osellame, L. Sansoni, I. Bongioanni, F. Sciarrino, G. Vallone, and P. Mataloni, "Integrated photonic quantum gates for polarization qubits," *Nature communications*, vol. 2, no. 1, pp. 1–6, 2011.

[367] G. Corrielli, A. Crespi, R. Geremia, R. Ramponi, L. Sansoni, A. Santinelli, P. Mataloni, F. Sciarrino, and R. Osellame, "Rotated waveplates in integrated waveguide optics," *Nature communications*, vol. 5, no. 1, pp. 1–6, 2014.

[368] F. Eltes, G. E. Villarreal-Garcia, D. Caimi, H. Siegwart, A. A. Gentile, A. Hart, P. Stark, G. D. Marshall, M. G. Thompson, J. Barreto, *et al.*, "An integrated optical modulator operating at cryogenic temperatures," *Nature Materials*, vol. 19, no. 11, pp. 1164–1168, 2020.

[369] P. Sibson, J. E. Kennard, S. Stanisic, C. Erven, J. L. O'Brien, and M. G. Thompson, "Integrated silicon photonics for high-speed quantum key distribution," *Optica*, vol. 4, no. 2, pp. 172–177, 2017.

[370] K. Okamoto, *Fundamentals of optical waveguides*. Elsevier, 2021.

[371] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, "Experimental realization of any discrete unitary operator," *Physical review letters*, vol. 73, no. 1, p. 58, 1994.

[372] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, "Optimal design for universal multiport interferometers," *Optica*, vol. 3, no. 12, pp. 1460–1465, 2016.

[373] J. Dalibard and C. Cohen-Tannoudji, "Laser cooling below the doppler limit by polarization gradients: simple theoretical models," *JOSA B*, vol. 6, no. 11, pp. 2023–2045, 1989.

[374] J. Tinbergen, "Astronomical polarimetry," *Astronomical Polarimetry*, 2005.

[375] J. N. Damask, *Polarization optics in telecommunications*, vol. 101. Springer Science & Business Media, 2004.

[376] V. S. Asadchy, M. S. Mirmoosa, A. Diaz-Rubio, S. Fan, and S. A. Tretyakov, "Tutorial on electromagnetic nonreciprocity and its origins," *Proceedings of the IEEE*, vol. 108, no. 10, pp. 1684–1727, 2020.

[377] L. Rayleigh, "On the magnetic rotation of light and the second law of thermodynamics," *Nature*, vol. 64, no. 1667, pp. 577–578, 1901.

[378] M. Mansuripur, "Reciprocity in classical linear optics," *Optics and Photonics News*, vol. 9, no. 7, pp. 53–58, 1998.

[379] S. Pancharatnam, "Generalized theory of interference and its applications," in *Proceedings of the Indian Academy of Sciences-Section A*, vol. 44, pp. 398–417, Springer, 1956.

[380] M. V. Berry, "The adiabatic phase and pancharatnam's phase for polarized light," *Journal of Modern Optics*, vol. 34, no. 11, pp. 1401–1407, 1987.

[381] R. Bhandari, "Polarization of light and topological phases," *Physics Reports*, vol. 281, no. 1, pp. 1–64, 1997.

[382] M. V. Berry, "Quantal phase factors accompanying adiabatic changes," *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, vol. 392, no. 1802, pp. 45–57, 1984.

[383] Y. Aharonov and J. Anandan, "Phase change during a cyclic quantum evolution," *Physical Review Letters*, vol. 58, no. 16, p. 1593, 1987.

[384] R. Bhandari, "Synthesis of general polarization transformers. a geometric phase approach," *Physics Letters A*, vol. 138, no. 9, pp. 469–473, 1989.

[385] K. Shiraishi, F. Tajima, and S. Kawakami, "Compact faraday rotator for an optical isolator using magnets arranged with alternating polarities," *Optics letters*, vol. 11, no. 2, pp. 82–84, 1986.

[386] M. Martinelli, "A universal compensator for polarization changes induced by birefringence on a retracing beam," *Optics Communications*, vol. 72, no. 6, pp. 341–344, 1989.

[387] G. García-Pérez, M. A. Rossi, B. Sokolov, F. Tacchino, P. K. Barkoutsos, G. Mazzola, I. Tavernelli, and S. Maniscalco, "Learning to measure: Adaptive informationally complete generalized measurements for quantum algorithms," *PRX Quantum*, vol. 2, no. 4, p. 040342, 2021.

[388] W. Chow, J. Gea-Banacloche, L. Pedrotti, V. Sanders, W. Schleich, and M. Scully, "The ring laser gyro," *Reviews of Modern Physics*, vol. 57, no. 1, p. 61, 1985.

[389] N. C. Harris, G. R. Steinbrecher, M. Prabhu, Y. Lahini, J. Mower, D. Bunandar, C. Chen, F. N. Wong, T. Baehr-Jones, M. Hochberg, *et al.*, "Quantum transport simulations in a programmable nanophotonic processor," *Nature Photonics*, vol. 11, no. 7, pp. 447–452, 2017.

[390] V. Saggio, B. E. Asenbeck, A. Hamann, T. Strömberg, P. Schiansky, V. Dunjko, N. Friis, N. C. Harris, M. Hochberg, D. Englund, *et al.*, "Experimental quantum speed-up in reinforcement learning agents," *Nature*, vol. 591, no. 7849, pp. 229–233, 2021.

[391] N. C. Harris, *Programmable nanophotonics for quantum information processing and artificial intelligence*. PhD thesis, Massachusetts Institute of Technology, 2017.

[392] M. Prabhu, C. Roques-Carmes, Y. Shen, N. Harris, L. Jing, J. Carolan, R. Hamerly, T. Baehr-Jones, M. Hochberg, V. Čeperić, *et al.*, "A recurrent ising machine in a photonic integrated circuit," *arXiv preprint arXiv:1909.13877*, 2019.

[393] L. Faustini and G. Martini, "Bend loss in single-mode fibers," *Journal of lightwave technology*, vol. 15, no. 4, pp. 671–679, 1997.

[394] R. T. Schermer and J. H. Cole, "Improved bend loss formula verified for optical fiber by simulation and experiment," *IEEE Journal of Quantum Electronics*, vol. 43, no. 10, pp. 899–909, 2007.

[395] D. Trillo, B. Dive, and M. Navascués, "Translating uncontrolled systems in time," *Quantum*, vol. 4, p. 374, 2020.

[396] D. Trillo, B. Dive, and M. Navascués, "A universal quantum rewinding protocol with an arbitrarily high probability of success," *arXiv preprint arXiv:2205.01131*, 2022.

[397] E. L. Hahn, "Spin echoes," *Physical review*, vol. 80, no. 4, p. 580, 1950.

[398] L. Viola and S. Lloyd, "Dynamical suppression of decoherence in two-state quantum systems," *Physical Review A*, vol. 58, no. 4, p. 2733, 1998.

[399] G. S. Uhrig, "Keeping a quantum bit alive by optimized π-pulse sequences," *Physical Review Letters*, vol. 98, no. 10, p. 100504, 2007.

[400] I. S. Sardharwalla, T. S. Cubitt, A. W. Harrow, and N. Linden, "Universal refocusing of systematic quantum noise," *arXiv preprint arXiv:1602.07963*, 2016.

[401] V. Dobrovitski, G. De Lange, D. Riste, and R. Hanson, "Bootstrap tomography of the pulses for quantum control," *Physical review letters*, vol. 105, no. 7, p. 077601, 2010.

[402] M. Navascués, "Resetting uncontrolled quantum systems," *Physical Review X*, vol. 8, no. 3, p. 031008, 2018.

[403] Z.-D. Li, X.-F. Yin, Z. Wang, L.-Z. Liu, R. Zhang, Y.-Z. Zhang, X. Jiang, J. Zhang, L. Li, N.-L. Liu, *et al.*, "Photonic realization of quantum resetting," *Optica*, vol. 7, no. 7, pp. 766–770, 2020.

[404] M. Gong, F. Xu, Z.-D. Li, Z. Wang, Y.-Z. Zhang, Y. Wu, S. Li, Y. Zhao, S. Wang, C. Zha, *et al.*, "Verification of a resetting protocol for an uncontrolled superconducting qubit," *npj Quantum Information*, vol. 6, no. 1, pp. 1–9, 2020.

[405] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, "Reversing unknown quantum transformations: Universal quantum circuit for inverting general unitary operations," *Physical Review Letters*, vol. 123, no. 21, p. 210502, 2019.

[406] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, "Experimental superposition of orders of quantum gates," *Nature communications*, vol. 6, no. 1, pp. 1–6, 2015.

[407] T. D. Jennewein, *Quantum communication and teleportation experiments using entangled photon pairs*. na, 2002.

[408] S. Milz, F. A. Pollock, and K. Modi, "An introduction to operational quantum dynamics," *Open Systems & Information Dynamics*, vol. 24, no. 04, p. 1740016, 2017.

[409] M. J. Todd, "Semidefinite optimization," *Acta Numerica*, vol. 10, pp. 515–560, 2001.

[410] M. Navascués, S. Pironio, and A. Acín, "Bounding the set of quantum correlations," *Physical Review Letters*, vol. 98, no. 1, p. 010401, 2007.

[411] M. T. Quintino, T. Vértesi, and N. Brunner, "Joint measurability, einstein-podolsky-rosen steering, and bell nonlocality," *Physical review letters*, vol. 113, no. 16, p. 160402, 2014.

[412] M. Horodecki, P. Horodecki, and R. Horodecki, "General teleportation channel, singlet fraction, and quasidistillation," *Physical Review A*, vol. 60, no. 3, p. 1888, 1999.

[413] M. T. Quintino and D. Ebler, "Deterministic transformations between unitary operations: Exponential advantage with adaptive quantum circuits and the power of indefinite causality," *Quantum*, vol. 6, p. 679, 2022.

[414] G. Chiribella and Z. Liu, "Quantum operations with indefinite time direction," *Communications Physics*, vol. 5, no. 1, pp. 1–8, 2022.

[415] M. Araújo, A. Feix, F. Costa, and Č. Brukner, "Quantum circuits cannot control unknown operations," *New Journal of Physics*, vol. 16, no. 9, p. 093026, 2014.

[416] X.-Q. Zhou, T. C. Ralph, P. Kalasuwan, M. Zhang, A. Peruzzo, B. P. Lanyon, and J. L. O'brien, "Adding control to arbitrary unknown quantum operations," *Nature communications*, vol. 2, no. 1, pp. 1–8, 2011.

[417] T. Colnaghi, G. M. D'Ariano, S. Facchini, and P. Perinotti, "Quantum computation with programmable connections between gates," *Physics Letters A*, vol. 376, no. 45, pp. 2940–2943, 2012.

[418] G. Chiribella, "Perfect discrimination of no-signalling channels via quantum superposition of causal structures," *Physical Review A*, vol. 86, no. 4, p. 040301, 2012.

[419] M. Araújo, F. Costa, and Č. Brukner, "Computational advantage from quantum-controlled ordering of gates," *Physical review letters*, vol. 113, no. 25, p. 250402, 2014.

[420] P. A. Guérin, A. Feix, M. Araújo, and Č. Brukner, "Exponential communication complexity advantage from quantum superposition of the direction of communication," *Physical review letters*, vol. 117, no. 10, p. 100502, 2016.

[421] K. Wei, N. Tischler, S.-R. Zhao, Y.-H. Li, J. M. Arrazola, Y. Liu, W. Zhang, H. Li, L. You, Z. Wang, *et al.*, "Experimental quantum switching for exponentially superior quantum communication complexity," *Physical review letters*, vol. 122, no. 12, p. 120504, 2019.

[422] M. J. Renner and Č. Brukner, "Computational advantage from a quantum superposition of qubit gate orders," *Physical Review Letters*, vol. 128, no. 23, p. 230503, 2022.

[423] G. Rubino, G. Manzano, and Č. Brukner, "Quantum superposition of thermodynamic evolutions with opposing time's arrows," *Communications Physics*, vol. 4, no. 1, pp. 1–10, 2021.

[424] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, "Probabilistic exact universal quantum circuits for transforming unitary operations," *Physical Review A*, vol. 100, no. 6, p. 062339, 2019.

[425] Q. Dong, M. T. Quintino, A. Soeda, and M. Murao, "The quantum switch is uniquely defined by its action on unitary operations," *arXiv preprint arXiv:2106.00034*, 2021.

[426] M. Campisi, P. Hänggi, and P. Talkner, "Colloquium: Quantum fluctuation relations: Foundations and applications," *Reviews of Modern Physics*, vol. 83, no. 3, p. 771, 2011.

[427] J. Bavaresco, M. Murao, and M. T. Quintino, "Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination," *Physical review letters*, vol. 127, no. 20, p. 200504, 2021.

[428] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Theoretical framework for quantum networks," *Physical Review A*, vol. 80, no. 2, p. 022339, 2009.

[429] M. Ziman, "Process positive-operator-valued measure: A mathematical framework for the description of process tomography experiments," *Physical Review A*, vol. 77, no. 6, p. 062112, 2008.

[430] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American statistical association*, vol. 58, no. 301, pp. 13–30, 1963.

[431] J. Bavaresco, M. Araújo, Č. Brukner, and M. T. Quintino, "Semi-device-independent certification of indefinite causal order," *Quantum*, vol. 3, p. 176, 2019.

[432] H. Cao, J. Bavaresco, N.-N. Wang, L. A. Rozema, C. Zhang, Y.-F. Huang, B.-H. Liu, C.-F. Li, G.-C. Guo, and P. Walther, "Experimental semi-device-independent certification of indefinite causal order," *arXiv preprint arXiv:2202.05346*, 2022.

[433] S. Yoshida, A. Soeda, and M. Murao, "Reversing unknown qubit-unitary operation, deterministically and exactly," *arXiv preprint arXiv:2209.02907*, 2022.

[434] O. Oreshkov, "Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics," *Quantum*, vol. 3, p. 206, 2019.

[435] M. Fellous-Asiani, R. Mothe, L. Bresque, H. Dourdent, P. A. Camati, A. A. Abbott, A. Auffèves, and C. Branciard, "Comparing the quantum switch and its simulations with energetically-constrained operations," *arXiv preprint arXiv:2208.01952*, 2022.

[436] K. Goswami and J. Romero, "Experiments on quantum causality," *AVS Quantum Science*, vol. 2, no. 3, p. 037101, 2020.

[437] Y. Guo, X.-M. Hu, Z.-B. Hou, H. Cao, J.-M. Cui, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, and G. Chiribella, "Experimental transmission of quantum information using a superposition of causal orders," *Physical review letters*, vol. 124, no. 3, p. 030502, 2020.

[438] G. Rubino, L. A. Rozema, D. Ebler, H. Kristjánsson, S. Salek, P. A. Guérin, A. A. Abbott, C. Branciard, Č. Brukner, G. Chiribella, *et al.*, "Experimental quantum communication enhancement by superposing trajectories," *Physical Review Research*, vol. 3, no. 1, p. 013093, 2021.

[439] H. Cao, N.-N. Wang, Z. Jia, C. Zhang, Y. Guo, B.-H. Liu, Y.-F. Huang, C.-F. Li, and G.-C. Guo, "Quantum simulation of indefinite causal order induced quantum refrigeration," *Physical Review Research*, vol. 4, no. 3, p. L032029, 2022.

[440] P. Yin, X. Zhao, Y. Yang, Y. Guo, W.-H. Zhang, G.-C. Li, Y.-J. Han, B.-H. Liu, J.-S. Xu, G. Chiribella, *et al.*, "Experimental super-heisenberg quantum metrology with indefinite gate order," *arXiv preprint arXiv:2303.17223*, 2023.

[441] R. W. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. Ford, A. Munley, and H. Ward, "Laser phase and frequency stabilization using an optical resonator," *Applied Physics B*, vol. 31, pp. 97–105, 1983.

[442] E. D. Black, "An introduction to pound–drever–hall laser frequency stabilization," *American journal of physics*, vol. 69, no. 1, pp. 79–87, 2001.

[443] M. Antesberger, M. T. Quintino, P. Walther, and L. A. Rozema, "Higher-order process matrix tomography of a passively-stable quantum switch," *arXiv preprint arXiv:2305.19386*, 2023.

[444] K. Goswami, Y. Cao, G. Paz-Silva, J. Romero, and A. White, "Increasing communication capacity via superposition of order," *Physical Review Research*, vol. 2, no. 3, p. 033292, 2020.

[445] M. M. Taddei, J. Cariñe, D. Martínez, T. García, N. Guerrero, A. A. Abbott, M. Araújo, C. Branciard, E. S. Gómez, S. P. Walborn, *et al.*, "Computational advantage from the quantum superposition of multiple temporal orders of photonic gates," *PRX Quantum*, vol. 2, no. 1, p. 010320, 2021.

[446] T. Nagata, R. Okamoto, J. L. O'brien, K. Sasaki, and S. Takeuchi, "Beating the standard quantum limit with four-entangled photons," *Science*, vol. 316, no. 5825, pp. 726–729, 2007.

[447] M. G. Raymer, "Quantum theory of light in a dispersive structured linear dielectric: a macroscopic hamiltonian tutorial treatment," *Journal of Modern Optics*, vol. 67, no. 3, pp. 196–212, 2020.

[448] N. Quesada, L. Helt, M. Menotti, M. Liscidini, and J. Sipe, "Beyond photon pairs—nonlinear quantum photonics in the high-gain regime: a tutorial," *Advances in Optics and Photonics*, vol. 14, no. 3, pp. 291–403, 2022.

[449] J. Schneeloch, S. H. Knarr, D. F. Bogorin, M. L. Levangie, C. C. Tison, R. Frank, G. A. Howland, M. L. Fanto, and P. M. Alsing, "Introduction to the absolute brightness and number statistics in spontaneous parametric down-conversion," *Journal of Optics*, vol. 21, no. 4, p. 043501, 2019.

[450] F. Dell'Anno, S. De Siena, and F. Illuminati, "Multiphoton quantum optics and quantum state engineering," *Physics reports*, vol. 428, no. 2-3, pp. 53–168, 2006.

[451] P. D. Drummond and M. Hillery, *The quantum theory of nonlinear optics*. Cambridge University Press, 2014.

Acknowledgements

The journey to completing a PhD can be a long one, and my own would not have been possible, or as enjoyable, without the people I met along the way, many of whom not only helped me grow as a scientist but as a person as well. The first person I would like to thank is Prof. Philip Walther for inviting me to join his group, and also for giving me the freedom to pursue my own ideas. Many of these turned out to be dead ends, but some of them constitute part of, or were instrumental in bringing about, the results presented in this thesis.

As an experimentalist you rarely work alone, and two people who deserve special mention are Peter and Valeria, with whom I spent countless hours and many late night in the lab, sharing both frustration and success.

Peter was always willing to listen to and engage with my latest question or idea. Sharing an eye for details as well as an often necessary stubbornness made working together feel very natural, and I can't recall a single instance where we ultimately disagreed on what the right approach to problem in the lab was. He would often do the things I was less inclined to, and do them more efficiently than I would have, and I could not have completed this thesis without him. His willingness to just go for things led to many moments I will never forget, such as when we fully disassembled the Mira not being sure if we could put it back together again, or our less scientific pursuits such as making the office door knob and the Mira mode-lock email-notification bot.

Valeria and I started around the same time, and I remember when we first met on my very first day in Vienna during the CoQuS interviews. Waiting in the hall after having seen the labs we discussed which one we'd prefer to work in, and we both said the second floor lab. Knowing that she already had a position in the group I resigned myself to a working different lab. Little did I know that we would end up sharing not only a lab, but many fun times and become good friends. No one else brought more laughter and enthusiasm to the group. I will also never forget my disbelief when you managed to break mode lock by touching a Babinet-Soleil.

There are many other members of the group I wish to thank as well. Jonas, for helping me when I first started, always giving the correct practical advice which I frequently ignored, and for becoming a good friend with whom I shared many great moments, especially in the Austrian mountains. Giulia, initially for indulging my questions and curiosity about her work, and later for all the great discussions and fun times together, for inviting me to collaborate on ideas outside my own projects, and for all our talks. Franci, for being himself and teaching me about the wonders of Neapolitan food, as well as good deal about physics. Our group retreats and get-togethers would not have been the same without you and your antics, and I'm glad to have gotten to know you as a friend.

I want to thank Bob for forcing me to write things down instead of immediately erasing them from the board, for being a great springboard for ideas, and for all our fun chats. Chiara, for welcoming me into 'her' lab, and for all the amazing hikes. Marie-Christine for being my guide to Austrian culture, for always helping me with German and for all the lunches we shared. Lee for always taking the time to discuss when I came to him with questions about physics, and for maintaining the cryo and nanowires.

Jenke, for the days starting at 11 o'clock, the taco evenings together, all the laughs and the unforgettable sailing trips.

Additionally I want to give my thanks to the quantum-group administrative staff at the University of Vienna for their support. First and foremost Petra, for always stepping up and trying to help, but also Christiane and Soli for all their help.

I want to thank Prof. Akira Furusawa for hosting me in his group at the University of Tokyo. I treasured the opportunity to learn about and work on an experiment in his lab. I was inspired by his deep knowledge of and passion for physics, as well as his research vision, and I am grateful for all the time I got to spend picking his brain. I also want to thank Marco for displaying a saintly patience when answering my questions about quantum information theory. I don't think I will ever meet a nicer theorist to collaborate with, and I will always remember the hidden karaoke bar you took me to in Toyko, even if I won't remember the way.

There are many people who helped me in the process of writing this thesis, not only through proof reading, but also by answering or helping me answer various questions, and I am thankful towards them all. Marco for all his help with precise details, Peter for his attentive comments, Christopher for his feedback on nonlinear optics, Matt, Iris, Franci, Lee, Vale, Bori and Juan for their comments, and the TEX community on Stackexchange for the solutions to most of my problems.

Finally, I want to thank Federica, the only physicist with whom I don't speak about physics, for her friendship. Melina, for encouraging me to accept the offer and move to Vienna back when we lived in Lund. My sister, for her support and good advice, and my parents for their unquestioning support.