universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„New quantum effects in space and time: complex networks, projectiles, and time translations“

verfasst von / submitted by

David Trillo Fernández

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:	UA 796 605 411
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Physik
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Časlav Brukner Miguel Navascués, PhD

Frankly, foundations of quantum
theory are still as enigmatic as before.

B.S.Tsirelson

Abstract

This thesis studies three new non-relativistic quantum effects related to space and time.

First, we determine that quantum theory with complex numbers predicts experimental results which are impossible in quantum theory with real numbers, as long as our notions of causality are unchanged. For this, we propose an experiment analogous to a Bell experiment, but in a more complicated causal network.

Second, we study a quantum particle in a line, and determine that there is a possibility that it can be observed in a distant region with a higher probability than a classical particle with the same momentum distribution.

Finally, we characterize all the time translations that can be probabilistically applied to an unknown quantum system of fixed dimension, and show that it is possible to rewind qubits with a high probability of success.

Zusammenfassung

In dieser Doktorarbeit werden drei neue nicht-relativistische Quanteneffekten erforscht, die mit Raum und Zeit zusammenhängen.

Erstens stellen wir fest, dass die experimentellen Vorhersagen der Quantentheorie von dem Zahlkörper des Hilbertraums abhängen, solange wir unsere Konzept von Kausalität nicht ändern. Dafür stellen wir ein Experiment vor, dass den Bell-Test zu einem komplexeren Netzwerk verallgemeinert. Dieses Experiment erlaubt es, komplexwertige und reellwertige Quantentheorie gegeneinander zu testen.

Zweitens untersuchen wir ein Quantenteilchen in einer Dimension. Wir zeigen, dass es möglich ist, es in einer entfernten Region mit höherer Wahrscheinlichkeit zu beobachten, als jedes klassische Teilchen mit derselbe Impulswahrscheinlichkeitsverteilung.

Zum Schluss charakterisieren wir alle Zeittranslationen, die probabilistisch auf ein unbekanntes Quantensystem mit fester Dimension angewendet werden können. Wir zeigen, dass es möglich ist, Qubits mit hoher Erfolgswahrscheinlichkeit zurückzuspulen.

Acknowledgements

Countless people have helped me throughout my PhD in some way or another. I'm grateful to all of you for having been a part of my life these last years - more than a small acknowledgements section could ever reflect. Also, I'm sure I forgot to mention some people, and for that I'm deeply sorry!

I would like to begin by thanking my supervisor Miguel Navascués for his constant support. His creativity, boundless curiosity and intelligence have been a never-ending source of inspiration, from which I believe I've been able to grow and surpass my limits. I've always been way too comfortable dealing with pure mathematicians, and Miguel has taught me the value of being grounded in the real world while doing research. I can also not emphasize enough how important the relaxed attitude and positive energy he brings every day to the table has been for me. This thesis would not have been possible without him.

I would also not be here without the constant support and encouragement of Juan León and Carlos Palazuelos, my mentors all throughout my studies. I would also like to thank Časlav Brukner for his encouragement, mentoring and help with the weirdest of requests. I am furthermore indebted to Valerio Scarani and Marcin Pawłowski for agreeing to review my thesis under such stringent time constraints. Thank you.

I would also like to thank the ÖAW for funding me through a DOC Fellowship, and their admin people, who have been nothing but helpful. I could also never forget to thank all the admin people at IQOQI Vienna who have made my life easier while working there, and the ones at the DSPL for all their help during the final stages of the writing. I have to especially thank Orkideh for the warm welcome to Vienna, which immediately made me feel at ease in what was otherwise a terrible week, and Urna, for relieving us of our cleaning duties with amazing skill and presence. On a similar note I'd like to thank Quique for accompanying and supporting me at the beginning of my trip.

My working life has been great thanks to all the amazing people that I've met throughout the years. First, all the people from the Navascués group, who not only have taught me a lot, but with whom I've also had a lot of fun: Ben Dive, Edgar Aguilar, Mirjam Weilenmann, Thinh Le, Andrew Garner, Mateus Araújo, Paolo Abiuso, Huan-Yu Ku. My only regret is that we never got a perfect score in Geoguessr. Also quite important to me have been my fellow physics students Martin Renner, Veronika Baumann, Rui Vasconcelos, Misha Maslov, Max Lerchbaumer, Jaroslav Kysela, Philippe Allard, Marius Krumm, Paul Juschitz, Alexander Zahrer. Thank you for all the fun times we have had together. I would also like to thank Costantino Budroni, Ämin Baumeler, Markus P. Müller, Luis Cortés Barbado, Robert Kindler for helping make IQOQI a nice friendly place to be in; and the people at the faculty of mathematics Eric Stenhede, Pilar Páez, Giancarlo Castellano, Johannes Droschl, Willi Kepplinger and his kangaroo Franz Josef, Steffen Plunder, Kevin Weigl, who adopted me as one of their own.

My personal life in Vienna has also been quite pleasant thanks to all the amazing people that I've met. My most heartfelt thanks go also to all the new friends who've made me feel at home and with whom I've lived countless adventures: Leon, Simon, Caro, Julian, Christoph, Bernhard, Anatol, Ilja, Quirin, Julian, Dominik, Mark. Special thanks to my DnD and CoC gang for endless hours of fun: Lukas, Emanuel, Anna, Julia, Konsti, Nicholas, Alex, Konrad. I would also like to thank my trumpet teacher, Johannes

Probst, for his infinite patience and understanding. I wish you the best for your own PhD!

I also want to thank the old friends who still get excited every time I visit home: Víctor, Jorge, and everyone at the Lewis Carroll association, old and new. Ione, Sara, Esther, Alex, Lucía, Antonio, Blanca, César, Fen, Ann. Thank you for staying my friends, even though I'm so far away. I miss you always.

During the great confinement, even though it is quite cliché, my view on life and the human condition did change quite dramatically. I am particularly grateful to all the friends who were there for me at the time. Javi, it was very fun to learn Lean with you. Paco, thanks for teaching me how to play Go. David, Alejandra, Marcos, Jaime, I had a blast playing board games. Pedro, I'll miss our endless study sessions.

I would also like to thank everyone that has visited me in Vienna, or that I have visited somewhere: Aitor Azemar, Alberto Ruiz de Alarcón, Daniel Zeitz, Ana, Cosmin and Laura, Med, Carlos, Dani, my parents and my grandma.

Let our friendships never end.

Contents

Chapter 1

Introduction

Our notions of space and time changed drastically at the beginning of last century, in two very different ways. On one hand, from general relativity we get the idea that space and time stand on equal footing, and that we should instead think of a space-time Lorentzian manifold. On the other hand, in quantum mechanics we throw away completely the notion that particles are well-localized, and instead think of their position as an observable that we can measure. However, there is no good notion of a time observable which encompasses all time-related measurements. Instead one has to carefully explain every time what "time" one is talking about: evolution time, time of arrival, a ticking clock, etc. Thus space and time are not on equal footing at all.

Combining these two (at first glance quite different) viewpoints is quite a tough challenge, but crucial for quantum gravity. Thus, studying each one independently is also of great value. In this thesis, we focus on space and time in non-relativistic quantum physics. We introduce three new quantum effects which deepen our understanding of their role in this theory.

In Chapter 2 we begin by studying space and time in the most minimal way: with causal networks. A causal network is a directed acyclic graph which specifies which parties in an experiment hold causal influence over other parties, as determined by the finiteness of the speed of light. We study causal networks that go beyond the one needed for Bell experiments. We determine that in certain causal networks the usual theory of quantum mechanics predicts experimental outcomes which are not reproducible by real quantum mechanics.

In Chapter 3 we study the movement of particles accross space in a simplified scenario. Namely, we look at the probability of observing a one-dimensional system in a distant region $[a, \infty) \subset \mathbb{R}$ after we have determined that it originally was localized in region $[0, L] \subset \mathbb{R} \setminus [a, \infty)$. We show that there are quantum systems for which this probability is greater than for any classical system with the same momentum distribution. This is a new quantum effect, similar to tunneling, which we call *quantum projectiles*. We determine that it is mathematically equivalent to some other effect called *quantum backflow*, where a particle with positive momentum is seen to traverse backwards. We give new bounds on the maximum advantage for such effects.

Finally, in Chapter 4 we study evolution time on finite-dimensional quantum systems. We are mainly interested in the so-called *time translations*, which are maps taking a system to another point in its evolution curve. We characterize all the possible time translations that can be probabilistically implemented in an unknown quantum system of a fixed dimension. For qubits, we furthermore prove that there are protocols which implement a "rewinding" with an arbitrarily high probability of success.

1. Content and structure of the thesis

This thesis is made up of four papers, which we have in part reorganized and expanded without changing the results. Chapter 2 is based on [RTW+21], Chapter 3 on [TLN22], and Chapter 4 on [TDN20; TDN23].

The goal was to make the reading experience more manageable, and perhaps also serve as a lighter introduction to these papers, as much as time has allowed me. Some explanations have been expanded and made more pedagogical, but some other parts have also been summarized or removed for the sake of exposition, so that one should also check the original papers for a more complete picture. A comparatively small part of the thesis is taken verbatim from those papers. If this endeavor proves to be useful to at least one person, I will be satisfied.

I have tried to make many of the symbols clickable. If you are reading this thesis online you may notice that you can click on some unhighlighted parts such as the following: "goto introduction". Doing so will usually take you to their definition, or at least to a part of the thesis where you can read more about it. This is an idea I've taken from [Mar21], and I think it should be a more widespread practice. It would certainly have made reading many papers much easier if they had been written this way.

I have also strived to make the thesis readable for a wider audience than the target-audience of each paper, and more rigorous than usual. However, at some point a line must be drawn, so I assume that the reader is familiar with functional analysis and quantum mechanics, and some results are not formulated with full rigor. In the next section I include a small guide of concepts and notation that I will use throughout the thesis.

2. Basic definitions and notation

2.1. Functional analysis. Given a vector space V over a field $\mathbb{K}$, we denote by $\mathrm{L}(V)$ the set of $\mathbb{K}$-linear maps from V to V, also called operators. If V is equipped with a topology, this set may include maps which are not continuous. By $\mathrm{GL}(V)$ we denote the subspace of $\mathrm{L}(V)$ consisting on maps which are invertible.

A Hilbert space is always denoted by $\mathcal{H}$ or $\mathcal{K}$ and, unless explicitely stated, may have finite or infinite dimension. It is always separable, although most theorems work for non-separable Hilbert spaces as well. It is always considered to be a vector space over the complex numbers, except in Chapter 2, where Hilbert spaces over real numbers are also considered. Elements of a Hilbert space are denoted by "kets" $|\psi\rangle$, and elements of the dual via Riesz theorem by "bras" $\langle\psi|$, so that the scalar product of $|\psi\rangle, |\varphi\rangle$ is $\langle\psi, \varphi\rangle \equiv \langle\psi|\varphi\rangle$. The scalar product is always linear in the second variable. This is dubbed the Dirac "bra-ket" notation.

If $\mathcal{H}$ is a Hilbert space, we denote by $\mathrm{B}(\mathcal{H})$ the vector space of continuous linear maps from $\mathcal{H}$ to $\mathcal{H}$. The natural topology in this space is given by the operator norm

$$\|A\| := \sup_{|\psi\rangle\in\mathcal{H}} \frac{\|A\,|\psi\rangle\|_{\mathcal{H}}}{\||\psi\rangle\|_{\mathcal{H}}}.$$

An operator A is of trace-class, $A \in \mathcal{T}(\mathcal{H})$, if for every orthonormal basis $\{|\psi_n\rangle\}_{n=1}^{\infty}$ of $\mathcal{H}$,

$$\mathrm{tr}(A) := \sum_{n=1}^{\infty} \langle\psi_n|\, A\,|\psi_n\rangle$$

is well-defined and independent of the choice of basis. In the subspace $\mathcal{T}(\mathcal{H})$ the natural norm is the trace norm, defined as $\|A\|_1 := \mathrm{tr}\left(\sqrt{AA^\dagger}\right)$.

We frequently consider unbounded maps. Such operators cannot be defined on all of $\mathcal{H}$ but only on a subspace, which we require to be dense in $\mathcal{H}$. Whenever $\mathcal{H}$ has unbounded operators, we consider $\mathrm{L}(\mathcal{H})$ to be the set of unbounded operators. Note that it is not a vector space, since the sum of two operators might not have a dense domain. The subspace where an operator A is defined is called its domain of definition, and denoted by $\mathrm{Dom}(A)$. Equality between unbounded operators A, B means that $\mathrm{Dom}(A) = \mathrm{Dom}(B)$ and $A\,|\psi\rangle = B\,|\psi\rangle$ for all $|\psi\rangle \in \mathrm{Dom}(A)$.

The adjoint of an operator $(A, \mathrm{Dom}(A))$ is $(A^\dagger, \mathrm{Dom}(A^\dagger))$, where

$$\mathrm{Dom}(A^\dagger) := \{|\varphi\rangle \in \mathcal{H} \mid \langle\varphi|\, A \text{ is bounded in } \mathrm{Dom}(A)\}$$

and for all $|\varphi\rangle \in \mathrm{Dom}(A^\dagger)$, $A^\dagger\,|\varphi\rangle$ is the unique vector $|\varphi'\rangle$ such that $\langle\varphi|\, A\,|\psi\rangle = \langle\varphi'|\psi\rangle$ for all $\psi \in \mathrm{Dom}(A^\dagger)$. An operator is Hermitian or self-adjoint, if $A = A^\dagger$.

We avoid the notation A^* for either the adjoint or the complex conjugation, in order to avoid confusion. Instead, we denote complex conjugation with a line, as in $\overline{z}$.

The spectrum of an operator $A \in \mathrm{L}(\mathcal{H})$ is the set of complex numbers $\lambda \in \mathbb{C}$ such that $(A - \lambda\mathbb{1}_{\mathcal{H}})$ does not have a continuous inverse. It is denoted by $\sigma(A)$. Self-adjoint operators have real spectrum.

A recurring Hilbert space that we are going to use is the space of square-integrable functions of some measure space. It is defined as

$$L^2(X, d\mu) := \left\{ f : X \to \mathbb{C} \text{ measurable} \;\middle|\; \int_X |f(x)|^2 d\mu < \infty \right\}.$$

When the measure is understood from context, we omit it. $\mathbb{R}$ is always equipped with the usual topology, and the Borel σ-algebra.

The Borel σ-algebra of a topological space X is the σ-algebra generated by the open sets of the topology. It is denoted by $\mathcal{B}(X)$.

Whenever we have a linear map $X : V \to W$, we can obtain a linear map

$$X_i : V_1 \otimes \cdots V_{i-1} \otimes V \otimes V_{i+1} \otimes \cdots \otimes V_n \longrightarrow V_1 \otimes \cdots \otimes V_{i-1} \otimes W \otimes V_{i+1} \otimes \cdots \otimes V_n$$

for arbitrary vector spaces V_i, which are usually deduced from context, as

$$X_i := \mathbb{1}_{V_1} \otimes \cdots \otimes \mathbb{1}_{V_{i-1}} \otimes X \otimes \mathbb{1}_{V_{i+1}} \otimes \cdots \otimes \mathbb{1}_{V_n}.$$

Given a tensor product of Hilbert spaces $\mathcal{H}_A \otimes \mathcal{H}_B$, the partial trace is a linear map

$$\begin{aligned} \mathrm{tr}_B : \mathrm{L}(\mathcal{H}_A \otimes \mathcal{H}_B) &\to \mathrm{L}(\mathcal{H}_A) \equiv \mathrm{L}(\mathcal{H}_A) \otimes \mathbb{K} \\ L &\mapsto \sum_{e \in \mathcal{B}} (\mathbb{1}_{\mathcal{H}_A} \otimes \langle e|) L (\mathbb{1}_{\mathcal{H}_A} \otimes |e\rangle), \end{aligned}$$

where $\mathcal{B}$ is an orthonormal basis of $\mathcal{H}_B$, and $\mathbb{K}$ is the base field ($\mathbb{R}$ or $\mathbb{C}$). This is only defined for trace-class operators. We also define the partial trace of a vector as the partial trace on the projection on the subspace generated by said vector. That is, $\mathrm{tr}_B(|\psi\rangle_{AB}) \equiv \mathrm{tr}_B(|\psi\rangle\langle\psi|)$.

2.2. Quantum information. We always work in units where $\hbar = 1$. Certain operators and states of finite-dimensional Hilbert spaces are special for quantum information, and have the following notation.

The Pauli matrices are

$$\sigma_x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Their eigenvectors are respectively denoted by $|0\rangle, |1\rangle$ for σ_z (also called the computational basis),

$$|+\rangle := \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |-\rangle := \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

for σ_x, and

$$|i\rangle := \frac{|0\rangle + i\,|1\rangle}{\sqrt{2}}, \quad |-i\rangle := \frac{|0\rangle - i\,|1\rangle}{\sqrt{2}},$$

for σ_y. The Hadamard gate is defined as $H|0\rangle = |+\rangle, H|1\rangle = |-\rangle$. The Bell basis is given by

$$|\phi^+\rangle := \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \; |\phi^-\rangle := \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \; |\psi^+\rangle := \frac{|10\rangle + |01\rangle}{\sqrt{2}}, \; |\psi^-\rangle := \frac{|10\rangle - |01\rangle}{\sqrt{2}}.$$

These states in density matrix form are denoted by capital letters, as

$$\Phi^+ := |\phi^+\rangle\langle\phi^+|, \; \Phi^- := |\phi^-\rangle\langle\phi^-|, \; \Psi^+ := |\psi^+\rangle\langle\psi^+|, \; \Psi^- := |\psi^-\rangle\langle\psi^-|.$$

The `SWAP` gate is defined as the unitary map in $\mathcal{H} \otimes \mathcal{H}$ that performs an exchange or "swap". That is, $\mathtt{SWAP}(|\psi\rangle \otimes |\phi\rangle) := |\phi\rangle \otimes |\psi\rangle$.

A Completely Positive Trace-Preserving (CPTP) map, also called a *quantum channel*, is a positive map $\Lambda : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})$ such that $\Lambda \otimes \mathbb{1}_{\mathcal{K}}$ is also positive for all $\mathcal{K}$, and $\mathrm{tr}(\Lambda(\rho)) = \mathrm{tr}(\rho)$ for all $\rho \in \mathcal{T}(\mathcal{H})$.

2.3. Miscellaneous. The natural numbers $\mathbb{N}$ do not include 0. The monoid of natural numbers together with 0 is denoted $\mathbb{N}_0$.

Given a ring R, we denote by $\mathrm{M}_n(R)$ the R-algebra of $n \times n$ matrices with entries in R with the usual scalar product, matrix multiplication and addition. If $R = \mathbb{C}$, we denote by U_n the $n \times n$ unitary matrices. If $R = \mathbb{R}$, we denote by O_n the $n \times n$ orthogonal matrices. For $R = \mathbb{C}$, we call GL_n the $n \times n$ invertible matrices. Finally, SL_n is the subspace of GL_n of matrices with determinant 1.

$\mathbb{C}^\times$ is the multiplicative group of complex numbers except 0.

An alphabet, usually denoted by Σ, is a set of symbols. Each symbol is called a *letter*. A *word* of length $n \in \mathbb{N}$ with letters in Σ is a string of n letters. The empty word is a word of length 0. The set of words of length n is denoted by Σ^n. We also denote

$$\Sigma^{\leq n} := \bigsqcup_{i=0}^{n} \Sigma^n, \quad \Sigma^* := \bigsqcup_{i=0}^{\infty} \Sigma^n.$$

We can also impose some equivalence relation in Σ^*. The classes of equivalence are called *reduced words*. In such a context, we redefine $\Sigma^{\leq n}$ to be the set of reduced words with a representative of length $\leq n$.

The algebra of polynomials on n variables is denoted $\mathbb{C}[x_1, ..., x_n]$. Sometimes we have an involution $*$ acting on $x_1, ..., x_n$. This makes $\mathbb{C}[x_1, ..., x_n]$ a *-algebra defining $*$ on any polynomial as an anti-linear and multiplicative extension of $*$ acting on the variables. When this is the case, we denote it by $\mathbb{C}^*[x_1, ..., x_n]$. The algebra of noncommutative polynomials on n variables is denoted $\mathbb{C}[X_1, ..., X_n]$. That is, non-commutative variables are denoted by capital letters. This is nothing but formal linear combinations of words in $X_1, ..., X_n$. In this case, we exclude the empty word. That is, there are no degree zero monomials. This is because the coefficients are in $\mathbb{C}$, while we are interested in evaluating the polynomials in different $\mathbb{C}$-algebras.

Chapter 2

Real Quantum Theory

This Chapter has been published as

- M.-O. Renou, D. Trillo, M. Weilenmann, T. P. Le, A. Tavakoli, N. Gisin, A. Acín, M. Navascués.
  "Quantum theory based on real numbers can be experimentally falsified"
  Nature 600, 625–629 (2021),

which is Reference [RTW+21], and to which all authors contributed equally.

The paper has been adapted for inclusion in this thesis, as follows. Section 2.2 corresponds to Appendix A of [RTW+21]. Section 3.3 corresponds to Appendix D of [RTW+21]. Section 4 is the main contribution and corresponds to Appendices E through H of [RTW+21]. The aforementioned Sections have been adapted to fit into the context of the thesis. The other Sections of this Chapter introduce some mathematical and physical preliminaries and contain no original work.

1. Introduction

Since the introduction of the Schrödinger equation, people have wondered why complex numbers appear in our formulation of quantum mechanics. Until that moment, it seemed that complex numbers were just a computational tool in wave mechanics, and even Schrödinger himself objected to their use and tried to remove them from the formulation of the theory. Many reasons have since then been given for the necessity of complex numbers the most common one being the breakdown of tomographic locality within real quantum mechanics [KM19]. But these are never experimentally checkable, only of a philosophical nature.

In this Chapter, we show that there is a feasible experiment that falsifies real quantum mechanics. To do this, we go beyond the usual Bell experiments [Bel64], and propose an experiment in a non-trivial tripartite causal network. The standard Bell experiment works as follows: two parties, Alice and Bob, are space-like separated. That is, any action of Alice is independent of any action of Bob. However, they share a resource. That is, two systems which in the past were together. This resource can be quantum

or classical, and there is a difference in the experimental results that Alice and Bob can obtain depending on the case. Such a scenario is diagrammatized as in Figure 1

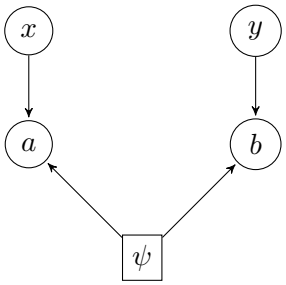

Figure 1. The causal structure of a Bell experiment. Here, Alice and Bob share a common quantum state ψ, and are able to perform measurements according to individual classical settings x, y, resulting in classical outcomes a, b. The results of an experiment are conditional probability distributions $\mathrm{P}(a, b|x, y)$, also called behaviours. Depending on whether ψ is quantum or classical, the set of possible behaviours is different.

It turns out, that in such a scenario it is impossible to experimentally detect a difference between real and complex quantum mechanics. However, by increasing the complexity of our causal network, as first introduced by [Fri12] in the setting of quantum mechanics, we can start detecting more subtle effects. In particular, we identify a simple scenario for which the predictions of real quantum mechanics are different than those in complex quantum mechanics, forever settling the debate over which theory we should consider.

This Chapter is organized as follows. In Section 2 we introduce the mathematical framework of the standard quantum theory with complex numbers. Then, in Section 3, we introduce the theory of real quantum mechanics, and prove than in all easy causal networks it can simulate complex quantum mechanics. Finally, in Section 4 we prove that, in the bilocality scenario, there is an experiment which gives different results in real and complex quantum mechanics.

2. Nonrelativistic Quantum Mechanics

2.1. The Hilbert space formulation of Quantum Mechanics. When forced to produce a set of postulates for QM, people will usually start writing some list of statements resembling the following:

(1) To a physical system A, there corresponds a separable Hilbert space $\mathcal{H}_A$, and a particular state of A is given by a unit vector $\psi \in \mathcal{H}_A$.

(2) There is an operator $H \in \mathrm{L}(\mathcal{H}_A)$, called the Hamiltonian, such that the state of a system changes in time as prescribed by the Schrödinger equation $i\partial_t\psi = H\psi$.

(3) To a measurable quantity f of the system A, there corresponds a self-adjoint operator $\hat{f} \in \mathrm{L}(\mathcal{H}_A)$ such that measuring f when A is in state ψ yields a probability distribution of outcomes with moments $\mathbb{E}_\psi\left[f^n\right] = \langle\psi|\left(\hat{f}\right)^n|\psi\rangle$.

These postulates form the basis of an *operational theory*, in which one identifies three stages to any experiment: a preparation phase, where a particular system is prepared in a particular state; an evolution phase, where the system exists and evolves in time;

and a measurement phase, where some property of the system is measured. The three stated postulates thus provide a mathematical framework in which to model any such experiment: one need only to find an adecuate Hilbert space and operators to describe their preparation, evolution and measurement stages.

In this Chapter we will not deal with the dynamical evolution given by Postulate (2). As it is customary in quantum information, we set $H \equiv 0$ and evolve a system discretely if we need to by applying unitaries.

Furthermore, one must include in the list of postulates some compositionality rule. Since we can only do experiments with small parts of the universe, we need a way to combine our descriptions of individual systems to get a description of a combined system. The usual way to do this in QM is the following:

(4) The Hilbert space of a combined system AB is the tensor product of the Hilbert spaces of system A and system B. That is, $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$.

An independent preparation procedure for systems A and B then yields a tensor product state $\psi \otimes \varphi \in \mathcal{H}_A \otimes \mathcal{H}_B$.

A local operation on system A is given by an operator of the form $\hat{f} \otimes \mathbb{1}_B \in \mathrm{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$.

This Postulate has been the subject of intense discussion. There is an alternative Postulate where instead of tensor products we consider commutation relations, as follows:

(iv) There is a Hilbert space $\mathcal{H}_{AB}$ for the combined system AB. All operators that correspond to local operations on system A commute with those that correspond to local operations on system B.

Until recently, it was not known whether these two postulates gave equivalent predictions in Bell nonlocality experiments. In [JNV+20], it is finally shown that they give rise to different theories, and therefore one must choose one.

2.2. The fourth postulate in Algebraic Quantum Field Theory. Many defenders of Postulate (iv) argue from the point of view of Quantum Field Theory (QFT). Even though any QFT is by construction (special) relativistic, in this section we argue that Postulate (4) can still be recovered in some cases, and therefore shouldn't be immediately dismissed.

Algebraic Quantum Field Theory (AQFT) is famous for proposing the fact that the local operations are the ones that commute with each other. More precisely, one of the possible starting points of AQFT is a correspondence from certain (open, bounded, contractible) regions of spacetime $\mathcal{O}$ in Minkowski space $\mathcal{M}$ to algebras of operators $\mathcal{A}(\mathcal{O})$ acting on a single Hilbert space $\mathcal{H}$ [Haa12]. This algebra of operators is meant to represent the operations that an experimenter might be able to perform in spacetime region $\mathcal{O}$, and is usually taken to be a C*-algebra. The axioms of AQFT are about the properties of the map $\mathcal{O} \to \mathcal{A}(\mathcal{O})$, and we need discuss only a few.

(I) A local state in the spacetime region $\mathcal{O}$ corresponds to a unital linear functional $\psi : \mathcal{A}(\mathcal{M}) \to \mathbb{C}$ such that $\psi(xx^*) \geq 0$ for all $x \in \mathcal{A}(\mathcal{O})$.

(IV) Given two space-like separated regions of spactime $\mathcal{O}_A$ and $\mathcal{O}_B$, the corresponding operator algebras $\mathcal{A}(\mathcal{O}_A)$ and $\mathcal{A}(\mathcal{O}_B)$ commute.

Postulate (IV) is the analogous of Postulate (iv) in AQFT, where furthermore there is a single universal Hilbert space $\mathcal{H}$ on which all local operations act. There are also several other Postulates about the structure of the correspondence $\mathcal{O} \to \mathcal{A}(\mathcal{O})$ so that it behaves as we expect with respect to inclusion of spacetime regions, and so on. For example, $\mathcal{A}(\mathcal{O}) \subset \mathcal{A}(\mathcal{M})$, and all algebras have a common unit element 1.

It is sometimes found that Postulate (I) admits states which are too general, and several axioms to define a set of "physical" states have been tried over time. Of particular interest is the question: what states can an experimenter prepare? Certainly any experimenter only has control over a small region of spacetime $\mathcal{O}_A$. And in order to prepare a particular state in $\mathcal{O}_A$ it is not unreasonable to require that some of the region $\mathcal{O}_A$ has to be shielded against external influences or devoted to experimental apparata. Therefore, the final state lives in a strictly smaller region $\mathcal{O}$.

Mathematically, we think of a map $T : \mathcal{A}(\mathcal{M}) \to \mathcal{A}(\mathcal{M})$ that performs this shielding against external influence. The fact that the experimenter only has access to the local region $\mathcal{O}$ translates to the property that $T(y) = T(1)y$ for all y in the commutant of $\mathcal{A}(\mathcal{O})$ in $\mathcal{A}(\mathcal{M})$, which we denote by $\mathcal{A}(\mathcal{O})'$. Indeed, if someone is performing some operation y in a spacelike separated region of spacetime, then $y \in \mathcal{A}(\mathcal{O})'$ by Postulate (IV), and a local shielding T should have no effect on y. However, we can allow a global effect on the common unit element - we say that T is *weakly localized*. Under very standard quantum mechanical assumptions ($\mathcal{A}(\mathcal{M}) = \mathrm{B}(\mathcal{H})$, $\mathcal{A}(\mathcal{O})$ is a von Neumann algebra on $\mathcal{H}$, T is normal and positive) it can be proven [Wer87] that weakly localized operations are precisely the ones that admit the form

$$T : x \mapsto \sum c_i^* x c_i$$

for $c_i \in \mathcal{A}(\mathcal{O})$, $x \in \mathrm{B}(\mathcal{H})$.

Therefore, we say that a local state ψ on $\mathcal{O}$ can be prepared in the region $\mathcal{O}_A$ if there exists a weakly localized operation in a region $\mathcal{O}_A$ such that $\overline{\mathcal{O}} \subset \mathcal{O}_A$ and $T(x) = \psi(x)T(1)$ for all $x \in \mathcal{A}(\mathcal{O})$. With these definitions and assumptions, we have the following:

Theorem 2.2.1 ([Wer87])**.** *There is a local state on $\mathcal{O}$ prepared in the region $\mathcal{O}_A$ if and only if there exists a type I factor $\mathcal{N}$ such that $\mathcal{A}(\mathcal{O}) \subset \mathcal{N} \subset \mathcal{A}(\mathcal{O}_A)$.*

A von Neumann algebra $\mathcal{N}$ in $\mathcal{H}$ is a type I factor, by one of the many equivalent definitions, if there exists a unitary $U : \mathcal{H} \to \mathcal{H}_A \otimes \mathcal{H}_B$ such that $\mathcal{N} = U^*(\mathrm{B}(\mathcal{H}_A) \otimes \mathbb{1}_B)U$. It therefore follows that, if we have two spacelike separated parties preparing local states in regions $\mathcal{O}_1$ and $\mathcal{O}_2$, we have by Postulate (IV) and Theorem 2.2.1 that $\mathcal{A}(\mathcal{O}_1) \subset \mathcal{N} \subset \mathcal{A}(\mathcal{O}_A)$, and $\mathcal{A}(\mathcal{O}_2)' \subset \mathcal{A}(\mathcal{O}_A)' \subset \mathcal{N}' \subset \mathcal{A}(\mathcal{O}_1)'$. Therefore, there is a unitary transformation such that

$$U\mathcal{A}(\mathcal{O}_1)U^* \subset \mathrm{B}(\mathcal{H}_A) \otimes \mathbb{1}_B, \quad U\mathcal{A}(\mathcal{O}_2)U^* \subset \mathbb{1}_A \otimes \mathrm{B}(\mathcal{H}_B),$$

and we recover the tensor product structure of Postulate (4).

In general, a correspondence $\mathcal{O} \to \mathcal{A}(\mathcal{O})$, where $\mathcal{A}(\mathcal{O})$ is a von Neumann algebra, is said to have the *split property* if, whenever $\overline{\mathcal{O}_1} \subset \mathcal{O}_2$, there is a type I von Neumann algebra $\mathcal{N}$ such that $\mathcal{A}(\mathcal{O}_1) \subset \mathcal{N} \subset \mathcal{O}_2$. This is usually taken as an extra axiom for AQFT, but can also be recovered from thermodynamic considerations, from energy bounds, and is satisfied in several well-known models such as free scalar fields, the Dirac field and some

interacting models in $1+1$d (see [FR20], and references within). Therefore, even though there is a difference between Postulates (4) and (*iv*) in the context of Bell nonlocality, we conclude that the physical one is Postulate (4) and that (*iv*) is purely of mathematical and computational interest. Alternatively, for the more conservative reader, we claim that Postulate (4) is a *good enough approximation* to study the composition of systems, and therefore causal structures.

2.3. More on the Postulates of Quantum Mechanics.

2.3.1. *Projection-Valued Measures.* In general, Postulate (3) as stated is not enough to obtain a probability distribution for the possible outcomes of a measurement [**Ex. 3.14**; Sie17]. It is sufficient, however, if there is a finite number of possible outcomes, and this is all we will look at in this Chapter. Nevertheless, in any situation, there is a canonical way of choosing such a probability distribution, which is equally -or even more- important as Postulate (3).

Definition 2.3.1 (PVM)**.** Let Σ be a σ-algebra on a set Λ. Let $\mathcal{H}$ be a Hilbert space. A Projector-Valued Measure (PVM) on (Λ, Σ) with values in $\mathrm{B}(\mathcal{H})$ is a map $\mu : \Sigma \to \mathrm{B}(\mathcal{H})$ such that

- For all $E \in \Sigma$, $\mu(E)$ is a projection.
- $\mu(\emptyset) = 0$, and $\mu(\Sigma) = \mathbb{1}_{\mathcal{H}}$.
- For all $|\psi\rangle \in \mathcal{H}$, and disjoint $E_i \in \Sigma$, $\mu\left(\bigsqcup_{i=1}^{\infty} E_i\right)|\psi\rangle = \sum_{j=1}^{\infty} \mu(E_j)\,|\psi\rangle$.
- For all $E, F \in \Sigma$, $\mu(E \cap F) = \mu(E)\mu(F)$.

From a PVM and a state $|\psi\rangle \in \mathcal{H}$ we can extract a proper probability measure on (Λ, Σ), as

$$\mu_\psi : E \mapsto \langle\psi|\,\mu(E)\,|\psi\rangle\,.$$

So now we just need a way to obtain a PVM from a self-adjoint operator. This is given by the spectral theorem.

Lemma 2.3.2 (Integration with respect to a PVM)**.** *Let μ be a PVM on a measurable set (Λ, Σ) with values on $\mathrm{B}(\mathcal{H})$. Let $f : \Lambda \to \mathbb{C}$ be a measurable function. Let*

$$\mathcal{H}_f := \left\{ |\psi\rangle \in \mathcal{H} \mid \int_\Lambda |f(\lambda)|^2 d\mu_\psi < \infty \right\}.$$

Then there is a unique operator $\int_\Lambda f d\mu$ with domain $\mathcal{H}_f$ such that

$$\langle\psi| \left(\int_\Lambda f d\mu \right) |\psi\rangle = \int_\Lambda f(\lambda) d\mu_\psi.$$

Furthermore, $\mathcal{H}_f$ is dense in $\mathcal{H}$ and $\int_\Lambda f d\mu$ is self-adjoint whenever f is real-valued.

Theorem 2.3.3 (Spectral theorem)**.** *Let $A \in \mathrm{L}(\mathcal{H})$ be a self-adjoint operator. There is a unique PVM μ^A on $(\sigma(A), \mathcal{B}(\sigma(A)))$ with values in $\mathrm{B}(\mathcal{H})$ such that*

$$\int_{\sigma(A)} \lambda d\mu^A = A.$$

Furthermore, for any $n \in \mathbb{N}$, we have

$$\int_{\sigma(A)} \lambda^n d\mu^A = A^n.$$

Proof. See, for example, Proposition 10.1 and Theorem 10.4 of [Hal13]. □

Using this, we can rewrite Postulate (3) in an equivalent manner, which is generally more useful.

(3') To a measurable quantity f of the system A with possible measurement outcomes in $\Lambda \subset \mathbb{R}$, corresponds a PVM μ^f on $(X, \mathcal{B}(X))$ with values in $\mathrm{B}(\mathcal{H})$ such that the probability measure of the outcomes of the measurement is given by μ^f_ψ.

Indeed, supose we have such a PVM. Then the operator

$$\hat{f} := \int_\Lambda \lambda d\mu^f$$

is self-adjoint by Lemma 2.3.2. Furthermore, we have that

$$\langle\psi| (\hat{f})^n |\psi\rangle = \langle\psi| \int_\Lambda \lambda^n d\mu^f |\psi\rangle = \int_\Lambda \lambda^n d\mu^f_\psi = \mathbb{E}_\psi [f^n],$$

so that Postulate (3) holds.

On the other hand, if we have a self-adjoint operator $\hat{f}$, by the spectral theorem we have a PVM $\mu^{\hat{f}}$ defined on its spectrum Λ, such that

$$\int_\Lambda \lambda d\mu^{\hat{f}} = \hat{f}.$$

Therefore,

$$\mathbb{E}_\psi [f^n] = \int_\Lambda \lambda^n d\mu^{\hat{f}}_\psi = \langle\psi| \int_\Lambda \lambda^n d\mu^{\hat{f}} |\psi\rangle = \langle\psi| \hat{f}^n |\psi\rangle,$$

so Postulate (3') makes sense. However, Postulates (3) and (3') are not exactly equivalent, because if the self-adjoint operator is not bounded, then the proability measure that has moments $\langle\psi| \hat{f}^n |\psi\rangle$ is in general not unique. Postulate (3') simply gives a prescription to choose one. However, it is the most natural choice, and therefore we shall work with Postulate (3').

The spectral theorem has one more important consequence, and it is that it allows us to define an action of measurable functions on self-adjoint operators:

Definition 2.3.4 (Functional calculus)**.** Let A a self-adjoint operator acting on $\mathcal{H}$, and f be a measurable function on $(\sigma(A), \mathcal{B}(\sigma(A)))$. We define

$$f(A) := \int_{\sigma(A)} f(\lambda) d\mu^A.$$

Note that it coincides with the usual definition of $f(A)$ if f is a polynomial, and thus already defined.

Remark 2.3.5 (Finite PVMs)**.** In this Chapter we will be concerned only with experiments that have a finite number of outcomes. In this case, all these theorems take on a much simpler form. Indeed, if there is only a finite set of outcomes $\Lambda = \{\lambda_1, ..., \lambda_n\}$, then we see from the Definition that a PVM is a set $\{P_i\}_{i=1}^n$ of as many orthogonal projectors of $\mathrm{B}(\mathcal{H})$ as outcomes, and they sum to the identity. Each P_i corresponds to $\mu(\lambda_i)$, and they determine the PVM uniquely. Furthermore, a self-adjoint operator A with a finite

spectrum Λ is defined on all of $\mathcal{H}$ by Lemma 2.3.2 and we have from the uniqueness in that same lemma, that

$$A = \int_\Lambda \lambda d\mu^A = \sum_{i=1}^n \lambda_i P_i.$$

Indeed,

$$\langle\psi| \sum_{i=1}^n \lambda_i P_i |\psi\rangle = \sum_{i=1}^n \lambda_i \langle\psi| \mu(\lambda_i) |\psi\rangle = \int_\Lambda \lambda d\mu_\psi.$$

For the rest of this Chapter, we will only consider measurements with a finite set of outcomes.

2.3.2. *Mixed states.* Suppose we have a composite state on two systems, which are space-like separated. That is, a vector in $\mathcal{H}_A \otimes \mathcal{H}_B$. When we have two parties, Alice and Bob, performing a measurement, we assume from Postulate (4) that the operator they are measuring is a self-adjoint map of the form $A \otimes B \in \mathrm{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$. However, we want to be able to see things from the point of view of each of the parties. Until Alice meets with Bob at some point in the future, she only sees a partial probability distribution of the outcomes she measures (namely, the spectrum of $A \otimes \mathbb{1}$ which is the same as $\sigma(A)$). In particular, we consider an outcome of such an experiment as a combined outcome of Alice and Bob, and thus write

$$\mathrm{P}(a, b|x, y) = \langle\psi| A_a^x \otimes B_b^y |\psi\rangle,$$

where $\{A_a^x\}_a$ are the PVMs associated to A^x for all x, and $\{B_b^y\}_b$ are the PVMs corresponding to B^y for all y.

Since Alice only has access to a partial probability measure, there are many situations in which we would also like to assign a state to Alice which reproduces this measure as μ_ψ^A. If Alice and Bob obtain the behavior $\mathrm{P}(a, b|x, y)$ as a result of a measurement, then that means that Alice must locally see the partial probability distribution

$$\mathrm{P}_A(a|x) = \sum_{b \in \sigma(B^y)} \mathrm{P}(a, b|x, y) = \sum_{b \in \sigma(B^y)} \langle\psi| A_a^x \otimes B_y^x |\psi\rangle = \langle\psi| A_a^x \otimes \mathbb{1}_{\mathcal{H}_B} |\psi\rangle \tag{1}$$

What this means under our postulates is a vector $|\varphi\rangle \in \mathcal{H}_A$ such that for any PVM of Alice, we recover the correct behavior $\mathrm{P}_A(a|x) = \langle\varphi| A_x^a |\varphi\rangle$. However, this doesn't always exist.

Example 2.3.6. Let $|\psi\rangle := (|00\rangle + |11\rangle)/\sqrt{2} \in \mathbb{C}^2 \otimes \mathbb{C}^2$. There is no vector $|\varphi\rangle \in \mathbb{C}^2$ such that $\langle\varphi| A |\varphi\rangle = \langle\psi| A \otimes \mathbb{1} |\psi\rangle$ for all projectors $A \in \mathrm{L}(\mathbb{C}^2)$.

Indeed, expand $|\varphi\rangle = a|0\rangle + b|1\rangle$. Then,

$$\langle\varphi| |0\rangle\langle 0| |\varphi\rangle = |a|^2 = \frac{1}{2}, \quad \langle\varphi| |1\rangle\langle 1| |\varphi\rangle = |b|^2 = \frac{1}{2}.$$

However,

$$\langle\varphi| |+\rangle\langle +| |\varphi\rangle = \frac{|a+b|^2}{2} = \frac{1}{2}, \quad \langle\varphi| |-\rangle\langle -| |\varphi\rangle = \frac{|a-b|^2}{2} = \frac{1}{2}.$$

And no two complex numbers a, b can satisfy this four equations simultaneously.

This means that if we want to associate a local state to Alice we have to generalize the notion of state itself. And there is an easy way to do this. Note from Eq. 1 that

$$\mathrm{P}_A(a|x) = \mathrm{tr}((A_a^x \otimes \mathbb{1}_{\mathcal{H}_B})\,|\psi\rangle\langle\psi|).$$

Therefore, by defining $\rho := \mathrm{tr}_B(|\psi\rangle\langle\psi|)$ we have that

$$P_A(a|x) = \mathrm{tr}(A_a^x \rho).$$

Now, from the rules of probability, we arrive at the usual definition for a state

Definition 2.3.7. A state on a Hilbert space $\mathcal{H}$ is a positive operator of trace one. We denote the set of states as $\mathcal{S}(\mathcal{H})$.

Note that since a positive operator satisfies $\langle x|\,\rho\,|x\rangle \geq 0$, the trace is always well-defined. That is, it doesn't depend on the choice of orthonormal basis.

The states we have been using until now, vectors in the Hilbert space, correspond to projectors onto a 1-dimensional subspace: $\rho_\psi := |\psi\rangle\langle\psi|$. These are called *pure states*, while a general state is sometimes named a *mixed state* or a *density matrix*. Pure states are particularly nice, and it turns out that by enlarging the Hilbert space we are working with we can think of any mixed state as the restriction of a pure state to a subsystem:

Theorem 2.3.8 (Purification)**.** *Let $\rho \in \mathcal{S}(\mathcal{H})$. Then, there exists a vector $|\psi\rangle = \mathcal{H} \otimes \mathcal{H}$ such that $\rho = \mathrm{tr}_{\mathcal{H}}(|\psi\rangle\langle\psi|)$.*

Proof. Let $\{|\psi_i\rangle\}_i$ be an orthonormal basis of $\mathcal{H}$. Since

$$1 = \mathrm{tr}(\rho) = \sum_i \langle\psi_i|\,\rho\,|\psi_i\rangle\,,$$

in particular it must be that $\langle\psi_i|\,\rho\,|\psi_i\rangle = 0$ except at most for a countable number of i, for which $0 \leq \langle\psi_i|\,\rho\,|\psi_i\rangle \leq 1$. Since

$$\langle\psi_i|\,\rho\,|\psi_i\rangle = \int_{\sigma(\rho)} \lambda d\mu^\rho_{\psi_i}(\lambda),$$

and $\sigma(\rho) \subset \mathbb{R}^+$, this means that $\mu^\rho_{\psi_i}(E) = 0$ for all $E \subset \sigma(\rho)$, except for a countable number of ψ_i. But since $\mu^\rho_{\psi_i}(E) = \langle\psi_i|\,\mu^\rho(E)\,|\psi_i\rangle$, we have that $\{\mu^\rho(E)\}_{E\subset\sigma(\rho)}$ is a countable set of projections. By definition of PVM, the image of each point in $\sigma(\rho)$ has to be an projection orthogonal to the others. Therefore, there is a countable number of points in $\sigma(\rho)$, and the spectral decomposition of ρ becomes

$$\rho = \sum_i \langle\varphi_i|\,\rho\,|\varphi_i\rangle\,|\varphi_i\rangle\langle\varphi_i| \tag{2}$$

for some orthonormal basis $\{|\varphi_i\rangle_i\}$ of $\mathcal{H}$. Let then $p_i := \langle\varphi_i|\,\rho\,|\varphi_i\rangle$. We define

$$|\psi\rangle := \sum_i \sqrt{p_i}\,|\varphi_i\rangle\,|\varphi_i\rangle \in \mathcal{H} \otimes \mathcal{H}.$$

Note that $\mathrm{tr}(\rho) = 1$ implies that $\langle\psi|\psi\rangle = 1$. Finally,

$$\mathrm{tr}_{\mathcal{H}}(|\psi\rangle\langle\psi|) = \sum_j p_i\,|\varphi_i\rangle\langle\varphi_i|\,|\langle\varphi_j|\varphi_i\rangle|^2 = \sum_i \langle\varphi_i|\,\rho\,|\varphi_i\rangle\,|\varphi_i\rangle\langle\varphi_i| = \rho.$$

□

Note that in particular, we have proven that the spectrum of a state is in $[0, 1]$, so every state is a bounded operator. That is, $\mathcal{S}(\mathcal{H}) \subset \mathrm{B}(\mathcal{H})$.

2.3.3. *Positive Operator-Valued Measures.* Analogously, we may want to consider convex combinations of measurement operators. For that, the definition given in Definition 2.3.1 is too restricted. In general, one considers

Definition 2.3.9 (POVM)**.** Let Σ be a σ-algebra on a set Λ. Let $\mathcal{H}$ be a Hilbert space. A Positive Operator-Valued Measure (POVM) on (Λ, Σ) with values in $\mathrm{B}(\mathcal{H})$ is a map $\mu : \Sigma \to \mathrm{B}(\mathcal{H})$ such that

- For all $E \in \Sigma$, $\mu(E)$ is a positive operator.
- $\mu(\emptyset) = 0$, and $\mu(\Sigma) = \mathbb{1}_{\mathcal{H}}$.
- For all $|\psi\rangle \in \mathcal{H}$, and disjoint $E_i \in \Sigma$, $\mu\left(\bigsqcup_{i=1}^{\infty} E_i\right)|\psi\rangle = \sum_{j=1}^{\infty} \mu(E_j)|\psi\rangle$.

For a finite number of outcomes, a POVM is therefore nothing more than a finite set of positive operators that add up to the identity.

In a setting where the Hilbert space is fixed, this more general concept is useful, because the set of PVMs is not convex, but the set of POVMs is. Indeed, consider a set of POVMs $\{\mu_i\}_{i=1}^n$ on (Λ, Σ) with values on $\mathrm{B}(\mathcal{H})$. Then is it straightforward to check that a convex combination

$$\sum_{i=1}^{n} \lambda_i \mu_i, \quad \lambda_i \in \mathbb{R}^+, \ \sum_{i=1}^{n} \lambda_i = 1.$$

is also POVM. However, a convex combination of PVMs is not necessarily a PVM, since the sum of projectors doesn't have to be a projector.

However, in our setting this will never be the case. All of our optimization problems also optimize over Hilbert spaces, and all of our measures are always defined on a nice measurable space (a finite set of points where all subsets are measurable). In this case, we have the following theorem, which we state in only the generality that we need.

Theorem 2.3.10 (Naimark)**.** *Let μ be a POVM over a finite space $(X, \mathcal{B}(X))$ with values in $\mathrm{B}(\mathcal{H})$. There exists a Hilbert space $\mathcal{K}$, a bounded linear map $L : \mathcal{K} \to \mathcal{H}$ and a PVM ν over $(X, \mathcal{B}(X))$ with values in $\mathrm{B}(\mathcal{K})$ such that $L^\dagger L = \mathbb{1}_{\mathcal{H}}$*

$$\mu(E) = L^\dagger \nu(E) L$$

Proof. We construct the PVM explicitely. Let $\mathcal{K} = \mathcal{H} \otimes \mathbb{C}^X$. Note that from the spectral theorem, we define, for $x \in X$,

$$\nu(x) := \mathbb{1}_{\mathcal{H}} \otimes |x\rangle\langle x| \,,$$

where $\{|x\rangle\}_{x \in X}$ is defined to be an orthonormal basis of $\mathbb{C}^X$. Furthermore, let the map L be defined as

$$L|\psi\rangle = \sum_{x \in X} \sqrt{\mu(x)}\,|\psi\rangle \otimes |x\rangle \,,$$

where $\sqrt{\mu(x)}$ is defined via the functional calculus 2.3.4. It is quite easy to check that with these definitions everythings works. Indeed, we have

$$L^\dagger L = \sum_{x \in X} \mu(x) = \mathbb{1}_{\mathcal{H}}, \quad L^\dagger \nu(x) L = \mu(x),$$

so that $L^\dagger \nu(E) L = \mu(E)$ follows by linearity. □

And indeed POVMs are usually regarded as Projective Measurements restricted to a lower dimension. Let μ be a POVM, and ν be a PVM as given by Naimark's theorem 2.3.10. Note that if we have an experiment in which we obtain a certain outcome x with probability $\mathrm{P}(x) = \mathrm{tr}(\mu(x)\rho)$, then from the cyclic property of the trace, we get

$$\mathrm{P}(x) = \mathrm{tr}\Big(\nu(x) L \rho L^\dagger\Big) = \mathrm{tr}(\nu(x)\tilde{\rho}),$$

where $\tilde{\rho} := L\rho L^\dagger \in \mathcal{S}(\mathcal{K})$. Therefore, the same probability distribution is obtained with a PVM on the Hilbert space $\mathcal{K}$, just by changing the state.

2.3.4. *Integrals of states.* Suppose we have a black box generating quantum states $\rho_\lambda \in \mathcal{S}(\mathcal{H})$ according to some probability measure $d\,\mathrm{P}_\Lambda(\lambda)$ on a probability space $(\Lambda, \Sigma, \mathrm{P}_\Lambda)$. This is a quite important situation in Quantum Information Theory that we need to describe in the framework that we have introduced. In particular, the output of said black box must be described by a state in $\mathcal{S}(\mathcal{H})$. In order to get a correct description we expect from the usual laws of probability, for any measurement $\{A_a\}_a$ we perform, to obtain the distribution of outcomes

$$\mathrm{P}(a) = \int_\Lambda \mathrm{P}(a,\lambda) d\,\mathrm{P}_\Lambda(\lambda) = \int_\Lambda \mathrm{tr}(A_a \rho_\lambda) d\,\mathrm{P}_\Lambda(\lambda). \tag{3}$$

If the probability measure P_Λ is discrete and finitely supported, we can further simplify

$$\int_\Lambda \mathrm{tr}(A_a\rho_\lambda) d\,\mathrm{P}_\Lambda(\lambda) = \sum_{\lambda=1}^N \mathrm{tr}(A_a\rho_\lambda) P_\Lambda(\lambda) = \mathrm{tr}\left(A_a \left(\sum_{\lambda=1}^N \rho_\lambda P_\Lambda(\lambda)\right)\right).$$

Therefore, defining the new operator

$$\rho := \sum_{\lambda=1}^N \rho_\lambda \,\mathrm{P}_\Lambda(\lambda),$$

which clearly belongs to $\mathcal{S}(\mathcal{H})$, we would obtain a correct description of the system. Similarly, when P_Λ is continuous, we expect to obtain a state of the form

$$\rho = \int_\Lambda \rho_\lambda d\,\mathrm{P}_\Lambda(\lambda) \tag{4}$$

to describe the black-box system. We can always do this, if the map $\lambda \mapsto \rho_\lambda$ is sufficiently well-behaved. We need some sort of measurability property in order to ensure that the integral of Equation (3) is well-defined no matter what measurement we decide to do. However, it turns out that the weakest notion of measurability $\lambda \mapsto \mathrm{tr}\{A\rho_\lambda\}$ being measurable and the integrals that we get out of that, coincide in the space of trace-class operators over a separable Hilbert space with the notion of strong measurability and Bochner integrability. Therefore, we may think of Equation (4) as a Bochner integral without any loss of generality.

Of particular importance will be the fact that all continuous (with respect to the topology of the trace-class operators) linear operators commute with the Bochner integral. This includes the trace and the partial trace.

2.4. Other formulations of Quantum Mechanics. In Postulate (1) of the Hilbert space formulation of QM, we associate a Hilbert space to each independent physical system. A Hilbert space is usually a complex vector space, and QM makes no exception to this. Furthermore, the evolution equation given in (2) makes explicit the imaginary unit, and thus the use of complex numbers. Used to the role of complex numbers in physics being purely computational, many people tried to formulate QM in a way that avoids their use. This is, of course, possible. After all, one can rewrite complex numbers as a pair of real numbers and call it a day. What is not so easy is, as we will see, to do this while maintaining the locality structure given by Postulate (4).

There are a few other descriptions of quantum mechanics. One particularly helpful one, which we will use in Chapter 3, is that of Wigner quasiprobability distributions. In this approach the role of the usual quantum-mechanical amplitudes $|\psi\rangle$ is played by Wigner functions, which are real-valued functions on the phase space of the system. In this case the composition of several systems also becomes complicated. There exist other quasiprobability distributions, which are less frequently used.

There are other formulations such as Bohmian mechanics or the path-integral formulation, which continue to use complex numbers, and therefore are not relevant for us in this Chapter.

Some other attemps to formulate quantum theory with real numbers are again not compatible with Postulate (4) [Stu60; ABW13].

3. Real Quantum Mechanics

What we call Real Quantum Mechanics (RQM) is not just a rewriting of QM where we hide the complex numbers under the rug at the cost of giving up some physical aspect of the theory, as it happens with the descriptions given in Sections **??** through **??**. Each of those are of independent interest, but the complex numbers mostly remain hidden in their formulations.

For us, RQM consists simply in taking the framework of QM given by Postulates (1)-(4) and imposing that all the mathematical objects are based on the field of real numbers, other than the field of complex numbers.

3.1. The Postulates of Real Quantum Mechanics. The theory we call RQM is therefore one satisfying the following Postulates:

($1_{\mathbb{R}}$) To a physical system A, there corresponds a *real* separable Hilbert space $\mathcal{H}_A$, and a particular state of A is given by a unit vector $\psi \in \mathcal{H}_A$.

($3_{\mathbb{R}}$) To a measurable quantity f of the system A, there corresponds a self-adjoint operator $\hat{f} \in \mathrm{L}(\mathcal{H}_A)$ such that measuring f yields a probability distribution of outcomes with moments $\mathbb{E}_A[f^n] = \langle\psi|(\hat{f})^n|\psi\rangle$.

($4_{\mathbb{R}}$) The Hilbert space of a combined system AB is the tensor product of the Hilbert spaces of system A and system B. That is, $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$.

An independent preparation procedure for systems A and B then yields a tensor product state $\psi \otimes \varphi \in \mathcal{H}_S \otimes \mathcal{H}_B$.

A local operation on system A is given by an operator of the form $\hat{f} \otimes \mathbb{1}_B \in \mathrm{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$.

At first sight it may seem that only Postulate $(1_{\mathbb{R}})$ is different from Postulate (1), but the change of a complex Hilbert space to a real Hilbert space changes all operators to be $\mathbb{R}$-linear (rather than $\mathbb{C}$-linear), and the tensor product to be over the field of real numbers rather than over the field of complex numbers.

The theory of Section 2.3 remains unchanged. From Postulate $(3_{\mathbb{R}})$ we can also get measurement operators, since the spectral theorem 2.3.3 also holds for self-adjoint operators on a real Hilbert space [**Remark 20.18**; MV97]. In particular, we may also formulate Postulate $(3_{\mathbb{R}})$ in the equivalent way

($3'_{\mathbb{R}}$) To a measurable quantity f of the system A with possible measurement outcomes in $\Lambda \subset \mathbb{R}$, corresponds a PVM μ^f on $(X, \mathcal{B}(X))$ with values in $\mathrm{B}(\mathcal{H})$ such that the probability measure of the outcomes of the measurement is given by μ^f_ψ.

Therefore, for any measurement with a finite number of outcomes, we have a corresponding number of projections onto closed subspaces of the real Hilbert space $\mathcal{H}$. One also speaks of POVMs and mixed states. Note that the Purification theorem 2.3.8 holds for real Hilbert spaces, as nothing in the proof requires complex numbers. Naimark's theorem 2.3.10 also holds for real spaces since in the proof one can use as $\mathcal{K} = \mathcal{H} \otimes \mathbb{R}^X$ and everything works the same way. Therefore we can usually restrict ourselves to PVMs as well.

However, the locality structure given by the tensor product over $\mathbb{R}$ is quite different from the one given by the tensor product over $\mathbb{C}$, and is this what results in different predictions. Let us illustrate this with a simple example.

Suppose that there are two parties, Alice and Bob, that share a quantum state. A bipartite state $\rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is called *separable* if it can be written as a convex combination of product states. If a state is not separable, then it is said to be *entangled*. Entangled states are very important for quantum foundations [EPR35], and therefore separable states form an important class of states to study. However, note that the notion of separability depends on whether we consider QM or RQM.

Indeed, consider the eigenvectors $|i\rangle, |-i\rangle$ of σ_y.

Example 3.1.1 ([CFR00])**.** The state

$$\rho := \frac{|i\rangle\langle i| \otimes |i\rangle\langle i| + |-i\rangle\langle -i| \otimes |-i\rangle\langle -i|}{2} \tag{5}$$

is separable as a state in QM but it is not separable as a state in RQM.

Proof. Note that

$$\rho = \frac{1}{8}\begin{pmatrix} 1 & -i & -i & -1 \\ i & 1 & 1 & -i \\ i & 1 & 1 & -i \\ -1 & i & i & 1 \end{pmatrix} + \frac{1}{8}\begin{pmatrix} 1 & i & i & -1 \\ -i & 1 & 1 & i \\ -i & 1 & 1 & i \\ -1 & -i & -i & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \tag{6}$$

so it is both a state in QM as well as in RQM. It is by construction separable. However, it not real-separable. Indeed, suppose we could write it as a convex combination of product states

$$\rho = \sum_{i=1}^{n} \lambda_i \sigma_i \otimes \rho_i, \quad \lambda_i \geq 0, \ \sum_{i=1}^{n} \lambda_i = 1, \ \sigma_i, \rho_i \in \mathcal{S}(\mathbb{R}^2).$$

Then, we would have

$$\rho^{T_A} = \sum_{i=1}^{n} \lambda_i \sigma_i^T \otimes \rho_i = \sum_{i=1}^{n} \lambda_i \sigma_i \otimes \rho_i = \rho.$$

However, since $(|i\rangle\langle i|)^T = |-i\rangle\langle -i|$, we have

$$\rho^{T_A} = \frac{1}{2}(|-i\rangle\langle -i| \otimes |i\rangle\langle i| + |i\rangle\langle i| \otimes |-i\rangle\langle -i|) = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

□

Note that the fact that we had a convex combination (instead of a usual linear combination) did not play any role in the proof. Indeed, with the same argument we can conclude that we cannot decompose ρ as an $\mathbb{R}$-linear combination of real product states. This is usually one of the main arguments previously used against RQM as a physical theory. However, this cannot really be experimentally tested.

3.2. Simulation of QM in RQM in the Bell scenario. We want some argument against RQM which can be detected experimentally. In Quantum Information, the experimentally available data consists on probability distributions of obtaining certain outcomes given a certain measurement. Thus, if we can recover a probability distribution obtained in an experiment using only RQM, we say that we have simulated QM with RQM.

Consider the causal structure depicted in Figure 2.

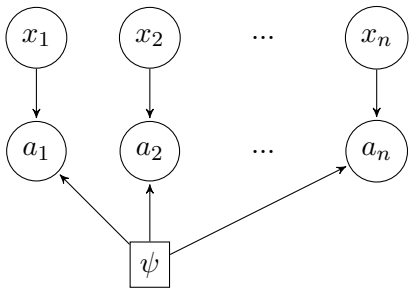

Figure 2. The causal structure of a multipartite Bell experiment. Here, n parties share a common quantum state ψ and are able to perform measurements according to individual classical settings x_i, resulting in classical outcomes a_i. We call this the *Bell scenario*.

It was first proven in [PV08] that one can maximally violate any bipartite Bell inequality in RQM. In [MMG09] the authors further show that this is true of all Bell experiments in the causal structures of Figure 2. We say that RQM can simulate QM in these causal structures.

In order to prove that this is true we need two things. The first one is to give a prescription that sends each party's complex measurement operators to real measurement

operators. The second one is a way to turn complex states into real states. The first one is easy to do, for consider the matrix

$$J := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

This matrix satisfies $J^\dagger = J^T = -J$, $J^2 = -\mathbb{1}_2$ and $JJ^\dagger = \mathbb{1}_2$ and is therefore a good candidate for the imaginary unit i. Indeed, if we substitute all imaginary units with J we get a well-known representation of the complex numbers. For matrices, this takes the following form:

Definition 3.2.1. For all $n \in \mathbb{N}$ and $A \in \mathrm{M}_n(\mathbb{C})$, denote by $\mathrm{Re}[A], \mathrm{Im}[A]$ the entry-wise real and imaginary parts of A, respectively. Note that this depends on the chosen basis. Then, we define the $\mathbb{R}$-linear map

$$\begin{aligned} f^* : \mathrm{M}_n(\mathbb{C}) &\longrightarrow \mathrm{M}_{2n}(\mathbb{R}) \\ A &\longmapsto \mathrm{Re}[A] \otimes \mathbb{1}_2 + \mathrm{Im}[A] \otimes J. \end{aligned}$$

This map satisfies all the properties that we want for a proper simulation:

Lemma 3.2.2. *For all $A, B \in \mathrm{M}_n(\mathbb{C})$, we have*

1. $f^*(A^\dagger) = f^*(A)^T$.
2. $f^*(AB) = f^*(A)f^*(B)$.
3. f^* *is a positive map.*
4. $f^*(\mathrm{U}_n) \subset \mathrm{O}_{2n}$. *The image of a projective measurement is a projective measurement. The image of a POVM is a POVM.*
5. $\mathrm{tr}(f^*(A)f^*(B)) = 2\mathrm{Re}[\mathrm{tr}(AB)]$.

Proof. 1. Indeed, since taking an adjoint is the same a taking a transpose and complex conjugating, we have $\mathrm{Re}[A^\dagger] = \mathrm{Re}[A^T] = \mathrm{Re}[A]^T$ and $\mathrm{Im}[A^\dagger] = -\mathrm{Im}[A^T] = -\mathrm{Im}[A]^T$. Therefore,

$$f^*(A^\dagger) = \mathrm{Re}[A^T] \otimes \mathbb{1}_2 + \mathrm{Im}[A^T] \otimes J^T = f^*(A)^T.$$

2. Since we can write $A = \mathrm{Re}[A] + i\mathrm{Im}[A]$, we have that $\mathrm{Re}[AB] = \mathrm{Re}[A]\mathrm{Re}[B] - \mathrm{Im}[A]\mathrm{Im}[B]$ and $\mathrm{Im}[AB] = \mathrm{Im}[A]\mathrm{Re}[B] + \mathrm{Re}[A]\mathrm{Im}[B]$. Therefore,

$$f^*(A)f^*(B) = (\mathrm{Re}[A]\mathrm{Re}[B] - \mathrm{Im}[A]\mathrm{Im}[B]) \otimes \mathbb{1}_2 + (\mathrm{Im}[A]\mathrm{Re}[B] + \mathrm{Re}[A]\mathrm{Im}[B]) \otimes J = f^*(AB).$$

3. Note that a matrix is positive semidefinite if and only if it can be written as $A = BB^\dagger$ for some matrix B. Therefore, if $A \geq 0$, then $f^*(A) = f^*(B)f^*(B)^T \geq 0$.

4. Since $f^*(\mathbb{1}_n) = \mathbb{1}_{2n}$, every claim follows from 1., 2., 3. and the linearity of f^*.

5. Since $\mathrm{tr}(J) = 0$, we have

$$\mathrm{tr}(f^*(A)f^*(B)) = \mathrm{tr}(f^*(AB)) = 2\,\mathrm{tr}(\mathrm{Re}[AB]) = 2\mathrm{Re}[\mathrm{tr}(AB)].$$

□

So this works exactly as we want to, except for property 4., which introduces an extra factor of 2 in the trace. This is an artifact of having to double the dimension of the system, and is the reason why we need a different map for states. Now, suppose that

we are doing an experiment with a single system in a single laboratory. That is, we are in the trivial causal structure given by Figure 3

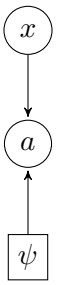

Figure 3. The trivial causal structure of a single party quantum experiment.

In this case, simulating an experiment just means reproducing the statistics $\mathrm{P}(a|x) := \mathrm{tr}(A_a\rho)$, where, in the biggest possible generality, ρ is a density matrix and A_a a POVM element. This is now quite simple to realize in RQM since, by Lemma 3.2.2 we have that $\mathrm{P}(a|x) = \mathrm{tr}(f^*(A_a)f^*(\rho))/2$. Since $f^*(A_a)$ is also a POVM element, our correspondence then works by sending states $\rho \mapsto f^*(\rho)/2$. Indeed, we know that positive operators are sent to positive operators, so we only need to correct the trace for everything to work out.

However, in the usual multipartite Bell scenario this simple strategy doesn't work, because our encoding of the quantum state includes a Hilbert space on which all parties are acting - the one used . This can be corrected by changing our encoding of the state in a way that respects the causality structure. The way to do this is to introduce a new rebit for each party via the following states:

Definition 3.2.3. Let $n \in \mathbb{N}$, $y \equiv y_1...y_n \in \{0,1\}^n$, $h(y) := \sum_{i=1}^n y_i$. We define the following n-qubit states

$$\left|\overline{0}_n\right\rangle := \sqrt{\frac{1}{2^{n-1}}} \sum_{h(y)\text{even}} (-1)^{\frac{h(y)}{2}} |y\rangle\,, \quad \left|\overline{1}_n\right\rangle := \sqrt{\frac{1}{2^{n-1}}} \sum_{h(y)\text{odd}} (-1)^{\frac{h(y)-1}{2}} |y\rangle\,.$$

These states have the nice property that we can act locally with J on any subsystem and the result is independent on which subsystem we have acted on. More precisely, we have that

Lemma 3.2.4. *For all $i \in \{1,...,n\}$, we have*

$$J_i\left|\overline{0}_n\right\rangle = \left|\overline{1}_n\right\rangle, \quad J_i\left|\overline{1}_n\right\rangle = -\left|\overline{0}_n\right\rangle.$$

Proof. Note that $J|0\rangle = |1\rangle$ and $J|1\rangle = -|0\rangle$, and let

$$K := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

For any $y \in \{0,1\}^n$ one has that $J_i|y\rangle = (-1)^{y_i}K_i|y\rangle$. Let $\overline{y} \in \{0,1\}^n$ be such that $|\overline{y}\rangle = K_i|y\rangle$. That is, $\overline{y}$ is obtained from y by negating the i-th component of y. We have that $h(\overline{y}) = h(y) + (-1)^{y_i}$, and each y with an even number of ones corresponds exactly

to one $\overline{y}$ with an odd number of ones. Therefore,

$$\sum_{h(y)\text{even}} (-1)^{\frac{h(y)}{2}} J_i |y\rangle = \sum_{h(y)\text{even}} (-1)^{\frac{h(y)}{2}-y_i} |\overline{y}\rangle = \sum_{h(\overline{y})\text{odd}} (-1)^{\frac{h(\overline{y})-1}{2}} |\overline{y}\rangle ,$$

from which the first claim follows. The second claim is analogous. □

We are now able to define the encoding of our states. It is easier to first define the enconding on vectors, rather than density matrices:

Definition 3.2.5. We define an enconding on states $|\psi\rangle \in \mathbb{C}^m = \mathbb{C}^{m_1} \otimes \cdots \otimes \mathbb{C}^{m_n}$ as

$$\begin{aligned} f_* : \mathbb{C}^m &\longrightarrow \mathbb{R}^{2nm} = (\mathbb{R}^{m_1} \otimes \mathbb{R}^2) \otimes \cdots \otimes (\mathbb{R}^{m_n} \otimes \mathbb{R}^2) \\ |\psi\rangle = \sum_{x=1}^{m} (a_x + ib_x) |x\rangle &\longmapsto S\left[\sum_{x=1}^{m} a_x |x\rangle \left|\overline{0}_n\right\rangle + b_x |x\rangle \left|\overline{1}_n\right\rangle\right] \\ &= S\left[\mathrm{Re}[|\psi\rangle] \otimes \left|\overline{0}_n\right\rangle + \mathrm{Im}[|\psi\rangle] \otimes \left|\overline{1}_n\right\rangle\right] , \end{aligned}$$

where $\{|x\rangle\}_{x=1}^m$ is some previously fixed orthonormal basis of $\mathbb{C}^m$, n is the number of parties in the Bell scenario and S is the linear map performing the system permutation

$$S : \mathbb{R}^{m_1} \otimes \cdots \mathbb{R}^{m_n} \otimes \mathbb{R}^2 \otimes \cdots \otimes \mathbb{R}^2 \to (\mathbb{R}^{m_1} \otimes \mathbb{R}^2) \otimes \cdots \otimes (\mathbb{R}^{m_n} \otimes \mathbb{R}^2).$$

The idea is that we will send each of the qubits forming the ancillary systems of $\left|\overline{0}_n\right\rangle$ and $\left|\overline{1}_n\right\rangle$ to each party. Therefore, each party receives a real state of double the dimension, which corresponds to the dimension of their encoded measurement operators f^*, and thus can act with them. This strategy works as intended, obtaining the correct probabilities:

Proposition 3.2.6. *For all $A \in \mathrm{M}_{m_i}(\mathbb{C})$, $|\psi\rangle , |\varphi\rangle \in \mathbb{C}^{m_1} \otimes \cdots \otimes \mathbb{C}^{m_n}$, with respect to a product basis,*

1. $f_*(A_i |\psi\rangle) = f^*(A)_i f_*(|\psi\rangle)$.
2. $f_*(|\varphi\rangle)^T f_*(|\psi\rangle) = \mathrm{Re}[\langle\varphi|\psi\rangle]$.

Proof. 1. Since we have a product basis, we have that taking real and imaginary parts commutes with extending the operator A from acting on subsystem i to acting on the global system. Furthermore, note that

$$A_i |\psi\rangle = \mathrm{Re}[A]_i \mathrm{Re}[|\psi\rangle] - \mathrm{Im}[A]_i \mathrm{Im}[|\psi\rangle] + i(\mathrm{Im}[A]_i \mathrm{Re}[|\psi\rangle] + \mathrm{Re}[A]_i \mathrm{Im}[|\psi\rangle]),$$

and thus,

$$\begin{aligned} S^{-1}[f_*(A_i |\psi\rangle)] = (\mathrm{Re}[A]_i \mathrm{Re}[|\psi\rangle] - \mathrm{Im}[A]_i \mathrm{Im}[|\psi\rangle]) \otimes \left|\overline{0}_n\right\rangle \\ + (\mathrm{Im}[A]_i \mathrm{Re}[|\psi\rangle] + \mathrm{Re}[A]_i \mathrm{Im}[|\psi\rangle]) \left|\overline{1}_n\right\rangle. \end{aligned}$$

On the other hand, using Lemma 3.2.4, we obtain

$$\begin{aligned} S^{-1}[f^*(A)_i f_*(|\psi\rangle)] &= (\mathrm{Re}[A]_i \otimes (\mathbb{1}_2)_i + \mathrm{Im}[A]_i \otimes J_i)(\mathrm{Re}[|\psi\rangle] \otimes \left|\overline{0}_n\right\rangle + \mathrm{Im}[|\psi\rangle] \otimes \left|\overline{1}_n\right\rangle \\ &= (\mathrm{Re}[A]_i \mathrm{Re}[|\psi\rangle] - \mathrm{Im}[A]_i \mathrm{Im}[|\psi\rangle]) \otimes \left|\overline{0}_n\right\rangle \\ &\quad + (\mathrm{Im}[A]_i \mathrm{Re}[|\psi\rangle] + \mathrm{Re}[A]_i \mathrm{Im}[|\psi\rangle]) \left|\overline{1}_n\right\rangle \\ &= S^{-1}[f_*(A_i |\psi\rangle)]. \end{aligned}$$

2. follows from a similar computation. □

Therefore, if we have a mixed state to begin with, we can encode it by linearity as

$$|\psi\rangle\langle\varphi| \mapsto f_*(|\psi\rangle\langle\varphi|) := f_*(|\psi\rangle) f_*(|\varphi\rangle)^T.$$

Using f^* to encode the measurements and f_* to encode the states we can perfectly simulate any Bell experiment in the causal network given in Figure 2. Indeed, any such experiment has as outcome a probability distribution of the form

$$\mathrm{P}(a_1, ..., a_n|x_1, ..., x_n) = \mathrm{tr}(A_1^{x_1} \otimes \cdots \otimes A_n^{x_n} \rho).$$

On the other hand, from Proposition 3.2.6, we deduce that

$$\mathrm{tr}(f^*(A_1^{x_1}) \otimes \cdots \otimes f^*(A_n^{x_n}) f_*(\rho)) = \mathrm{Re}[\mathrm{tr}(A_1^{x_1} \otimes \cdots \otimes A_n^{x_n} \rho)].$$

Since probabilities are real numbers, both expressions are the same.

Remark 3.2.7. If $n = 1$ we can recover the encoding of states used in the trivial causal structure of Figure 3. Indeed,

$$|\psi\rangle\langle\psi| = \sum_{x,y=1}^{n} \left[(a_x a_y + b_x b_y) + i(b_x a_y - a_x b_y)\right] |x\rangle\langle y| \,,$$

and therefore

$$\frac{1}{2} f^*(|\psi\rangle\langle\psi|) = \frac{1}{2} \sum_{x,y=1}^{n} (a_x a_y + b_x b_y) \, |x\rangle\langle y| \otimes \mathbb{1}_2 + (b_x a_y - a_x b_y) \, |x\rangle\langle y| \otimes J.$$

However, we have

$$f_*(|\psi\rangle\langle\psi|) = \sum_{x,y=1}^{m} |x\rangle\langle y| \otimes (a_x a_y \, |0\rangle\langle 0| + b_x b_y \, |1\rangle\langle 1| + b_x a_y \, |1\rangle\langle 0| + a_x b_y \, |0\rangle\langle 1|).$$

It is, it seems, a different encoding. On the other hand, consider the state $i\,|\psi\rangle = \sum_{x=1}^{m} (-b_x + i a_x) \, |x\rangle$. Since it only differs by a phase from ψ, all the statistics must also coincide, and thus, we may also try simulating this state, as

$$f_*(i\,|\psi\rangle\langle\psi|\,(-i)) = \sum_{x,y=1}^{m} |x\rangle\langle y| \otimes (b_x b_y \, |0\rangle\langle 0| + a_x a_y \, |1\rangle\langle 1| - b_x a_y \, |0\rangle\langle 1| - a_x b_y \, |1\rangle\langle 0|).$$

From here, we obtain

$$\frac{1}{2} f^*(|\psi\rangle\langle\psi|) = \frac{1}{2} f_*(|\psi\rangle\langle\psi|) + \frac{1}{2} f_*(i\,|\psi\rangle\langle\psi|\,(-i)) = f_*((1+i)/\sqrt{2}\,|\psi\rangle\langle\psi|\,(1-i)/\sqrt{2}).$$

That is, there is a phase for which both simulations coincide. Indeed, the map f_* is only defined on density matrices up to a phase

3.3. Simulation of QM in RQM in the PBR scenario. Consider the causal network depicted in Figure 4. In a sense, it is the dual of a Bell scenario. Instead of many parties performing independent measurements on a single quantum state, we have a single party performing a measurement on many independent quantum states.

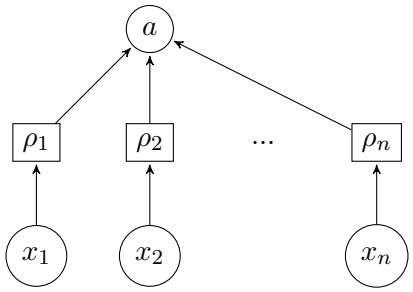

Figure 4. The causal structure dual to that of Figure 2. Here, a single party performs a measurement on n independent quantum states, which are prepared according to some classical settings x_i, obtaining outcome a. We call this the *PBR scenario*.

Although it seems as simple as the trivial causal structure given in Figure 3, the restriction on the state to be a product of many states has some interesting consequences, as shown in [PBR12]. Now we show that all experiments in the PBR scenario are reproducible in RQM, by giving another simulation, as in Section 3.2.

We have that any experiment performed on the PBR scenario outputs a behavior of the form

$$\mathrm{P}(a|x_1,...,x_n) = \mathrm{tr}(A_a(\rho_1^{x_1} \otimes \cdots \otimes \rho_n^{x_n})).$$

In order to simulate this result on real Hilbert spaces, we again need an encoding for the measurement operators, and one for the states. We will now provide a simulation which is drastically different from that of Section 3.2.

Suppose that party i has m_i possible preparation settings $\{1,...,m_i\}$. Then, we associate to $\rho_i^{x_{m_i}}$ an element of an orthonormal basis $\{|x_j\rangle\}_{x_j=1}^{m_i}$ of $\mathbb{R}^{m_i}$. That is, we encode $\rho_i^{x_i}$ as $g_*(\rho_i^{x_i}) := |x_i\rangle\langle x_i|$. These are clearly trace 1 positive operators, and are thus valid states.

On the other hand, we can now encode A_a as

$$A_a \mapsto g^*(A_a) := \sum_{x_1,...,x_n} \mathrm{P}(a|x_1,...,x_n)\,|x_1,...,x_n\rangle\langle x_1,...,x_n|\,.$$

Since this is a diagonal matrix with positive entries, it is positive. On the other hand, since $\mathrm{P}(a|x_1,...,x_n)$ is a conditional probability distribution, we have that

$$\sum_a g^*(A_a) = \mathbb{1}_{\mathbb{R}^{m_1}\otimes\cdots\otimes\mathbb{R}^{m_n}},$$

and thus $\{g^*(A_a)\}$ are a POVM, and therefore a valid measurement.

Finally, by construction we have that

$$\mathrm{P}(a|x_1,...,x_n) = \mathrm{tr}(g^*(A_a)(g_*(\rho_1^{x_1}) \otimes \cdots \otimes g_*(\rho_n^{x_n}))).$$

4. The SWAP scenario

Consider now the causal network depicted in Figure 5. This is the next simplest causal network after the ones studied in Sections 3.2 and 3.3.

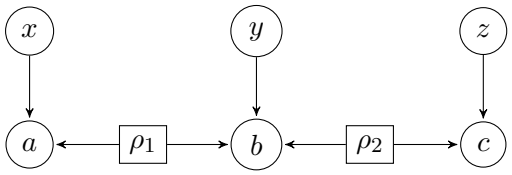

Figure 5. A causal network for a bilocality experiment

Since it is not a Bell scenario, the validity of the simulation strategy we saw in Section 3.2 is not ensured. The main problem is Bob's measurement operators. He is acting on two Hilbert spaces, so according to the prescription given in Section 3.2, he would receive two extra rebits used to encode the real and imaginary degrees of freedom, so we cannot just double the dimension of his measurements via f^*, since then there would be a dimension mismatch.

An obvious candidate would be to decompose each of Bob's measurements $B = \sum_{i=1}^{n} B_i^1 \otimes B_i^2$ and apply $f^* \otimes f^*$. However,

Proposition 4.0.1 (cf. Lemma 3.2.2)**.** *The map f^* is not 2-positive. In particular, $f^* \otimes f^*$ is not positive, and a POVM does not need to map to a POVM under $f^* \otimes f^*$.*

Proof. Indeed, consider the state

$$\left|\psi^+\right\rangle := \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

One has that

$$\left|\psi^+\right\rangle\!\left\langle\psi^+\right| = \frac{1}{4}(\sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z + \mathbb{1}_2 \otimes \mathbb{1}_2),$$

where σ_i are the Pauli matrices, and thus

$$f^* \otimes f^*(\left|\psi^+\right\rangle\!\left\langle\psi^+\right|) = \frac{1}{4}(\sigma_x \otimes \mathbb{1}_2 \otimes \sigma_x \otimes \mathbb{1}_2 - J \otimes J \otimes J \otimes J + \sigma_z \otimes \mathbb{1}_2 \otimes \sigma_z \otimes \mathbb{1}_2 + \mathbb{1}_{16}).$$

But this matrix has eigenvectors with eigenvalue $-1/2$. One of them is $|0010\rangle - |0111\rangle + |1000\rangle + |1101\rangle$. □

For example, then, this map will not properly encode a Bell measurement into RQM. Of course, a priori it's possible that we need to change the encoding again, as we had to do from the trivial causal network to the Bell scenario. Or maybe it could be that we have to combine the simulations provided in Sections 3.2 and 3.3. We prove in the sequel that none of this is possible.

4.1. Imposibility of simulation of QM with RQM. We are going to work in the slightly more general network given by Figure 6

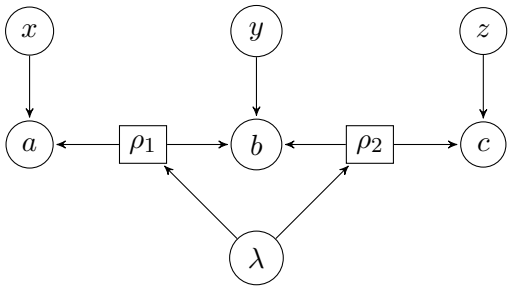

Figure 6. The causal network for the SWAP scenario.

The only difference with the network presented in Figure 5 is that the quantum systems may be correlated with a classical variable. There are many reasons for considering this more general network. First, the set of behaviors that can be obtained in this network is convex, whereas the set obtained from Figure 5 is not. Second, it is more general and our results apply to it. Finally, from an experimental point of view, it is complicated to ensure that there are no classical correlations between the quantum systems.

We will other consider that each party has a finite number of inputs and outputs. We denote the set of inputs as $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ respectively for Alice, Bob and Charlie. Analogously, we denote the sets of outputs by $\mathcal{A}, \mathcal{B}, \mathcal{C}$, respectively. A behavior obtained as a result of an experiment conducted in this causal network has the form

$$\mathrm{P}(a,b,c|x,y,z) = \int_\Lambda \mathrm{tr}\Big((A_a^x \otimes B_b^y \otimes C_c^z)(\sigma_{AB_1}^\lambda \otimes \sigma_{B_2C}^\lambda)\Big) dP_\Lambda(\lambda), \tag{7}$$

where P_Λ is a probability measure on Λ representing the classical correlations, and there are four Hilbert spaces $\mathcal{H}_A, \mathcal{H}_{B_1}, \mathcal{H}_{B_2}, \mathcal{H}_C$ such that $\sigma_{AB_1}^\lambda \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_{B_1})$, $\sigma_{B_2C}^\lambda \in \mathcal{S}(\mathcal{H}_{B_2} \otimes \mathcal{H}_C)$ for all $\lambda \in \Lambda$, and $\{A_a^x\}_{a\in\mathcal{A}}$ is a POVM with values in $\mathrm{B}(\mathcal{H}_A)$ for all $x \in \mathcal{X}$, $\{B_b^y\}_{b\in\mathcal{B}}$ is a POVM with values in $\mathrm{B}(\mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2})$ for all $y \in \mathcal{Y}$, and $\{C_c^z\}_{c\in\mathcal{C}}$ is a POVM with values in $\mathrm{B}(\mathcal{H}_C)$ for all $z \in \mathcal{Z}$.

Depending on if we are in QM or RQM the aforementioned Hilbert spaces will be complex or real, and the tensor product will be over the complex or the real numbers. We will show that there is a QM behavior in this causal network that cannot be reproduced in RQM, by explicitely giving one.

Consider the scenario given by the causal network of Figure 6, with $\mathcal{A} = \mathcal{C} = \{-1, 1\}$, $\mathcal{B} = \{0, 1, 2, 3\}$, $\mathcal{X} = \{1, 2, 3\}$, $\mathcal{Y} = \{1\}$, $\mathcal{Z} = \{1, 2, 3, 4, 5, 6\}$. We define the behavior $\mathrm{P}_\mathbb{C}(a, b, c|x, y, z)$ explicitely through Equation 7. We will set $P_\Lambda(\lambda) = \delta(\lambda)$, so that the behavior is also valid in the causal network of Figure 5. After removing all the superfluous notation, we have

Definition 4.1.1 (Optimal complex behavior)**.** We define the behavior

$$\mathrm{P}_\mathbb{C}(a,b,c|x,z) = \mathrm{tr}\Big((\tilde{A}_a^x \otimes \tilde{B}_b \otimes \tilde{C}_c^z)(\tilde{\sigma}_{AB_1} \otimes \tilde{\sigma}_{B_2C})\Big). \tag{8}$$

explicitely by giving all the operators in the right-hand side. We let both states be the maximally entangled qubit state

$$\tilde{\sigma}_{AB_1} = \tilde{\sigma}_{B_2C} = \Phi^+.$$

Bob measures in the Bell basis. That is,

$$\tilde{B}_0 = \Phi^+,\ \tilde{B}_1 = \Psi^+,\ \tilde{B}_2 = \Phi^-,\ \tilde{B}_3 = \Psi^-.$$

Finally, we define a PVM for Alice and Charlie via the spectral theorem 2.3.3 by giving self-adjoint operators. That is, we define $\tilde{A}^x = \tilde{A}^x_1 - \tilde{A}^x_{-1}$ and $\tilde{C}^z = \tilde{C}^z_1 - \tilde{C}^z_{-1}$ as

$$\tilde{A}^1 = \sigma_z,\ \tilde{A}^2 = \sigma_x,\ \tilde{A}^3 = \sigma_y,$$

and, letting $D_{ij} := (\sigma_i + \sigma_j)/\sqrt{2}$, $E_{ij} := (\sigma_i - \sigma_j)/\sqrt{2}$, we set

$$\tilde{C}^1 := D_{zx},\ \tilde{C}^2 := E_{zx},\ \tilde{C}^3 := D_{zy},\ \tilde{C}^4 := E_{zy},\ \tilde{C}^5 := D_{xy},\ \tilde{C}^6 := E_{xy}.$$

By construction, this behavior admits a decomposition of the form given by Equation 7 in QM. We have used the complex numbers in this decomposition, and as we will see this is unavoidable.

Theorem 4.1.2. [↓] *The behavior* $\mathrm{P}_{\mathbb{C}}(a,b,c|x,z)$ *does not admit a decomposition of the form given by Equation (7) in RQM. That is, if we require the Hilbert spaces and the operators to be real, and the tensor product be over the real numbers.*

In other words, if we are doing an experiment which can be modelled in the causal network given by Figure 6, then it is impossible to get the behavior $\mathrm{P}_{\mathbb{C}}(a,b,c|x,z)$ as an outcome of such an experiment within RQM.

The idea of the proof is as follows. First, we consider a linear functional on the probabilities. We see that $\mathrm{P}_{\mathbb{C}}$ achieves the maximum possible value that this functional can achieve in QM, and we use self-testing techniques to figure out which states we could have started with. Using the self-tested state, we can apply local unitary transformations and arrive at another state which cannot be prepared within this causal network in RQM, thus showing that the initial state needs to live in a complex Hilbert space.

The linear functional is a variation on one introduced in [APVW16]. It is called CHSH_3 because it is the sum of three CHSH functionals:

Definition 4.1.3. Given some behavior $\mathrm{P}(a,b,c|x,z)$, we define

$$S^{xz}_b := \sum_{a,c\in\{-1,1\}} ac\,\mathrm{P}(a,b,c|x,z).$$

We define now the linear functional

$$\begin{aligned}\mathcal{J}_b(\mathrm{P}) := {} & (-1)^{b_2}(S^{11}_b + S^{12}_b) + (-1)^{b_1}(S^{21}_b - S^{22}_b) \\ & + (-1)^{b_2}(S^{13}_b + S^{14}_b) + (-1)^{b_1+b_2+1}(S^{33}_b - S^{34}_b) \\ & + (-1)^{b_1}(S^{25}_b + S^{26}_b) + (-1)^{b_1+b_2+1}(S^{35}_b - S^{36}_b),\end{aligned} \tag{9}$$

where $b \equiv b_1 b_2$ is the binary decomposition of the numbers 0 through 3.

Each line in equation (9) corresponds to some CHSH functional of the correlations of Alice and Charlie conditioned on Bob having observed result b. A quick calculation shows that

Lemma 4.1.4. $\mathcal{J}_b(\mathrm{P}_{\mathbb{C}}) = 6\sqrt{2}\mathrm{P}_{\mathbb{C}}(b)$, *and* $\mathrm{P}_{\mathbb{C}}(b) = 1/4$.

Proof. That the marginal $\mathrm{P}_{\mathbb{C}}(b)$ is well defined and independent of x, z follows from equation (8). Let

$$\rho_b := \mathrm{tr}_{B_1B_2}((\mathbb{1}_A \otimes \tilde{B}_b \otimes \mathbb{1}_C)(\Phi^+ \otimes \Phi^+)).$$

One can compute by hand or remembering that this is precisely how to perform *entanglement swapping* [ZZHE93], that

$$\rho_b = \frac{1}{4}\tilde{B}_b.$$

In particular, $\mathrm{tr}(\rho_b) = \mathrm{P}_{\mathbb{C}}(b) = 1/4$. We may now write

$$\mathrm{P}_{\mathbb{C}}(a,b,c|x,z) = \frac{1}{4}\,\mathrm{tr}\Big((\tilde{A}_a^x \otimes \tilde{C}_c^z)\tilde{B}_b\Big)$$

The signs in equation (9) have been chosen so that, for each b, one gets the maximum violation of the three CHSH inequalities that we are adding up, and therefore, we obtain the result. □

The self-test of the CHSH_3 was performed in [BŠCA18]. It needs the introduction of the following notion:

Definition 4.1.5 (Regularized operators)**.** Given any Hermitian operator $A \in \mathrm{B}(\mathcal{H})$, we define its regularized version via the functional calculus 2.3.4

$$\hat{A} := f(A),$$

where f is the measurable function given by

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$
$$x \longmapsto \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases}$$

From the spectral decomposition, we see that this regularization procedure turns a hermitian operator into a (hermitian) unitary operator with eigenvalues ± 1. Note also that $\hat{A}A = |A|$.

We are now ready to prove the theorem.

Proof of Theorem 4.1.2. Suppose, by contradiction, that there is a probability distribution $P_\Lambda(\lambda)$, real Hilbert spaces $\mathcal{H}_A, \mathcal{H}_{B_1}, \mathcal{H}_{B_2}, \mathcal{H}_C$, measurement operators $A_a^x \in \mathrm{B}(\mathcal{H}_A), B_b \in \mathrm{B}(\mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}), C_c^z \in \mathrm{B}(\mathcal{H}_C)$ and states $\sigma_{AB_1} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_{B_1}), \sigma_{B_2C} \in \mathcal{S}(\mathcal{H}_{B_2} \otimes \mathcal{H}_C)$ such that

$$\mathrm{P}_{\mathbb{C}}(a,b,c|x,z) = \int_\Lambda \mathrm{tr}\Big((A_a^x \otimes B_b \otimes C_c^z)(\sigma_{AB_1}^\lambda \otimes \sigma_{B_2C}^\lambda)\Big) dP_\Lambda(\lambda).$$

Since, as noted in Section 3.1, Theorem 2.3.10 and its consequences also hold in RQM, we may assume that each party's measurement operators form a PVM. Let us call

$$\rho := \int_\Lambda \sigma_{AB_1}^\lambda \otimes \sigma_{B_2C}^\lambda dP_\Lambda(\lambda)$$

We have that, with probability $P(b) = \mathrm{tr}(\rho(\mathbb{1}_A \otimes B_b \otimes \mathbb{1}_C))$, Bob observes outcome b. Conditioned on this result, the joint statistics of Alice and Charlie are obtained from the state

$$\rho_b := \frac{\mathrm{tr}_{B_1B_2}((\mathbb{1}_{\mathcal{H}_A} \otimes B_b \otimes \mathbb{1}_{\mathcal{H}_C})\rho)}{\mathrm{P}(b)}$$

as $\mathrm{P}(a,c|x,z,b) = \mathrm{tr}((A_a^x \otimes C_c^z)\rho_b)$.

Consider now a purification (Theorem 2.3.8) of this state. That is, a state $|\psi_b\rangle \in \mathcal{H}_A \otimes \mathcal{H}_C \otimes \mathcal{H}_P$ such that

$$P(a,c|x,z,b) = \langle\psi_b| A_a^x \otimes C_c^z \otimes \mathbb{1}_{\mathcal{H}_P} |\psi_b\rangle.$$

We will rename $\mathcal{H}_C \equiv \mathcal{H}_C \otimes \mathcal{H}_P$ and $C_c^z \equiv C_c^z \otimes \mathbb{1}_{\mathcal{H}_P}$ to simplify the notation. Let us again call $A^x := A_1^x - A_{-1}^x$ and $C^z := C_1^z - C_{-1}^z$ the self-adjoint operators corresponding to the PVM performed by Alice, and Charlie respectively. We consider the following four variants of the CHSH_3 operator

$$\begin{aligned}\hat{\mathcal{J}}_b := {} & (-1)^{b_2} A^1 (C^1 + C^2) + (-1)^{b_1} A^2 (C^1 - C^2) \\ & + (-1)^{b_2} A^1 (C^3 + C^4) + (-1)^{b_1+b_2+1} A^3 (C^3 - C^4) \\ & + (-1)^{b_1} A^2 (C^5 + C^6) + (-1)^{b_1+b_2+1} A^3 (C^5 - C^6),\end{aligned}$$

where $b = b_1 b_2$ is the binary representation of the numbers 0 through 3, and the tensor product between Alice and Charlie's observables is implied. For all b, this is a combination of 3 CHSH operators, and therefore it is called CHSH_3. In particular, one has that the average value corresponds to the value of the functional defined in Equation (9):

$$\langle\psi_b| \hat{\mathcal{J}}_b |\psi_b\rangle = \frac{\mathcal{J}_b(\mathrm{P}_{\mathbb{C}})}{\mathrm{P}(b)} = 6\sqrt{2}, \tag{10}$$

where the first equality follow from the fact that $\langle\psi_b| A^x \otimes C^z |\psi_b\rangle = S_b^{xz}/P(b)$ from the spectral decomposition of A^x and C^z, and the second equality follows from Lemma 4.1.4. Furthermore, we have the following sum-of-squares (SOS) decomposition:

$$\begin{aligned}\sqrt{2}(6\sqrt{2} - \hat{\mathcal{J}}_b) = {} & \left((-1)^{b_2} A^1 - \frac{C^1 + C^2}{\sqrt{2}}\right)^2 + \left((-1)^{b_1} A^2 - \frac{C^1 - C^2}{\sqrt{2}}\right)^2 \\ & + \left((-1)^{b_2} A^1 - \frac{C^3 + C^4}{\sqrt{2}}\right)^2 + \left((-1)^{b_1+b_2+1} A^3 - \frac{C^3 - C^4}{\sqrt{2}}\right)^2 \\ & + \left((-1)^{b_1} A^2 - \frac{C^5 + C^6}{\sqrt{2}}\right)^2 + \left((-1)^{b_1+b_2+1} A^3 - \frac{C^5 - C^6}{\sqrt{2}}\right)^2\end{aligned} \tag{11}$$

The tensor product with identites is again implied. Note that this is technically a *sum of hermitian squares* (SOHS). That is, a decomposition of the form $A = \sum_i P_i P_i^\dagger$. This is always the kind of decomposition that we want, since we need that each summand is positive. When all the operators involved are hermitian, an SOHS is the same as an SOS. In particular, since $\langle\psi_b| (6\sqrt{2} - \hat{\mathcal{J}}_b) |\psi_b\rangle = 0$, by positivity we must have that $\langle\psi_b| P_i^2 |\psi_b\rangle = 0$ for all i in the SOS $(6\sqrt{2} - \hat{\mathcal{J}}_b) = \sum_i P_i^2$ given by equation (11), and therefore $P_i|\psi_b\rangle = 0$. Grouping terms, we get

$$\begin{aligned}(-1)^{b_2} A^1 \otimes \mathbb{1}_{\mathcal{H}_C} |\psi_b\rangle &= \mathbb{1}_{\mathcal{H}_A} \otimes \frac{C^1 + C^2}{\sqrt{2}} |\psi_b\rangle = \mathbb{1}_{\mathcal{H}_A} \otimes \frac{C^3 + C^4}{\sqrt{2}} |\psi_b\rangle, \\ (-1)^{b_1} A^2 \otimes \mathbb{1}_{\mathcal{H}_C} |\psi_b\rangle &= \mathbb{1}_{\mathcal{H}_A} \otimes \frac{C^1 - C^2}{\sqrt{2}} |\psi_b\rangle = \mathbb{1}_{\mathcal{H}_A} \otimes \frac{C^5 + C^6}{\sqrt{2}} |\psi_b\rangle, \\ (-1)^{b_1+b_2+1} A^3 \otimes \mathbb{1}_{\mathcal{H}_C} |\psi_b\rangle &= \mathbb{1}_{\mathcal{H}_A} \otimes \frac{C^3 - C^4}{\sqrt{2}} |\psi_b\rangle = \mathbb{1}_{\mathcal{H}_A} \otimes \frac{C^5 - C^6}{\sqrt{2}} |\psi_b\rangle.\end{aligned} \tag{12}$$

Let us now call

$$Z^A := A^1, \quad X^A := A^2, \quad Y^A := A^3,$$
$$Z^C := \frac{C^1 + C^2}{\sqrt{2}}, \quad X^C := \frac{C^1 - C^2}{\sqrt{2}}, \quad Y^C := \frac{C^3 - C^4}{\sqrt{2}},$$

in reference to the optimal complex operators. We will also continuously avoid writing the tensor product in the notation. For example, Equation (12) looks like

$$(-1)^{b_2} Z^A |\psi_b\rangle = Z^C |\psi_b\rangle, \quad (-1)^{b_1} X^A |\psi_b\rangle = X^C |\psi_b\rangle,$$
$$(-1)^{b_1+b_2+1} Y^A |\psi_b\rangle = Y^C |\psi_b\rangle. \tag{13}$$

A different SOS decomposition of the CHSH inequality is also useful:

$$\begin{aligned}\sqrt{2}(6\sqrt{2} - \hat{\mathcal{J}}_b) = &\left(C^1 - \frac{(-1)^{b_2} A^1 + (-1)^{b_1} A^2}{\sqrt{2}}\right)^2 + \left(C^2 - \frac{(-1)^{b_2} A^1 - (-1)^{b_1} A^2}{\sqrt{2}}\right)^2 \\ &+ \left(C^3 - \frac{(-1)^{b_2} A^1 - (-1)^{b_1+b_2} A^3}{\sqrt{2}}\right)^2 + \left(C^4 - \frac{(-1)^{b_2} A^1 + (-1)^{b_1+b_2} A^3}{\sqrt{2}}\right)^2 \\ &+ \left(C^5 - \frac{(-1)^{b_1} A^2 - (-1)^{b_1+b_2} A^3}{\sqrt{2}}\right)^2 + \left(C^6 - \frac{(-1)^{b_1} A^2 + (-1)^{b_1+b_2} A^3}{\sqrt{2}}\right)^2\end{aligned} \tag{14}$$

With the usual argument then we have that

$$\left[C^1 - \frac{(-1)^{b_2} A^1 + (-1)^{b_1} A^2}{\sqrt{2}}\right] |\psi_b\rangle = 0.$$

Therefore,

$$\left[C^1 + \frac{(-1)^{b_2} A^1 + (-1)^{b_1} A^2}{\sqrt{2}}\right] \left[C^1 - \frac{(-1)^{b_2} A^1 + (-1)^{b_1} A^2}{\sqrt{2}}\right] |\psi_b\rangle = 0.$$

Since C^1, A^1 and A^2 all square to the identity, the left hand side is $(-1)^{b_1+b_2+1}\{A^1, A^2\}/2$. In particular, we have $\{A^1, A^2\}|\psi_b\rangle = 0$. Similarly, one obtains that all of Alices observables anticommute on the subspace spanned by the $|\psi_b\rangle$:

$$\{Z^A, X^A\}|\psi_b\rangle = 0, \quad \{Z^A, Y^A\}|\psi_b\rangle = 0, \quad \{X^A, Y^A\}|\psi_b\rangle = 0. \tag{15}$$

We now proceed to use these relations to obtain a self-testing result. That is, there is a local unitary operation that Alice and Charlie can perform to bring $|\psi_b\rangle$ to certain fixed states. Such a unitary, which we call $U \otimes V$ is defined in Figure 7. It is adapted from the one used in [BŠCA18] so that it is a real operator.

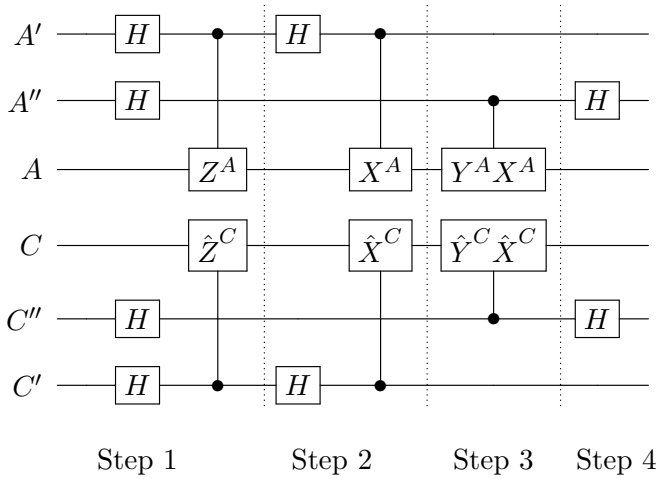

Figure 7. Real local isometry $U \otimes V$ built from each party's untrusted measurement operators, written in circuit form. That is, the depicted maps are applied in from left to right on the corresponding tensor product labelled by the wire. In particular, A denotes the usual Hilbert space of Alice, $\mathcal{H}_A$, and A', A'' denote each a two-dimensional (real) Hilbert space, or *rebit*. Analogously for Charlie. For convenience, the map is divided in four steps. H represents the Hadamard gate. A black dot together with a vertical determines a *controlled gate*. That is, a gate applied coherently when the control qubit is $|1\rangle$.

Note that Z^A, X^A, Y^A are unitary operators, while Z^C, Y^C, X^C might not be. We thus regularize Charlie's operators. The important thing to note is that the regularized versions of Charlie's operators also satisfy the relations (13). Indeed, since $\left|Z^A\right| = \mathbb{1}$ and $A \leq |A|$ for any hermitian operator A,

$$\begin{aligned} \left\|(\hat{Z}^C - Z^C)|\psi_b\rangle\right\| &= \left\|(\mathbb{1} - \hat{Z}^C Z^C)|\psi_b\rangle\right\| = \left\|(\mathbb{1} - \left|Z^C\right|)|\psi_b\rangle\right\| \\ &= \left\|(\mathbb{1} - \left|Z^A\right|\left|Z^C\right|)|\psi_b\rangle\right\| \leq \left\|(\mathbb{1} - Z^A Z^C)|\psi_b\rangle\right\| = 0. \end{aligned} \tag{16}$$

The same is obviously true for the other operators X^C, Y^C. Putting Equations (13) and (16) together, we have that

$$\begin{aligned} (-1)^{b_2} Z^A|\psi_b\rangle = \hat{Z}^C|\psi_b\rangle, \quad (-1)^{b_1} X^A|\psi_b\rangle = \hat{X}^C|\psi_b\rangle, \\ (-1)^{b_1+b_2+1} Y^A|\psi_b\rangle = \hat{Y}^C|\psi_b\rangle. \end{aligned} \tag{17}$$

Let us now see what the action of the map defined on Figure 7 is. We proceed step by step. For clarity of notation, we ignore the order of the tensor products, and instead denote with subindices in which Hilbert spaces each vector lives.

Step 1. In this step we begin by initializing the state in

$$|0000\rangle_{A'A''C'C''} |\psi_b\rangle_{AC}.$$

After application of the unitary maps of Step 1, we obtain

$$\begin{aligned} \frac{1}{2} |{+}{+}\rangle_{A''C''} [(|00\rangle_{A'C'} + (-1)^{b_2} |11\rangle_{A'C'})|\psi_b\rangle_{AC} \\ + ((-1)^{b_2} |01\rangle_{A'C'} + |10\rangle_{A'C'}) Z^A |\psi_b\rangle_{AC}], \end{aligned}$$

where we have used Equation (17) and the fact that Z^A is unitary and hermitian.

Step 2. Recall that

$$H \otimes H|\phi^+\rangle = |\phi^+\rangle, \quad H \otimes H|\phi^-\rangle = |\psi^+\rangle, \tag{18}$$
$$H \otimes H|\psi^+\rangle = |\phi^-\rangle, \quad H \otimes H|\psi^-\rangle = -|\psi^-\rangle.$$

Therefore, after another round of Hadamard gates, we get

$$\begin{cases} \frac{1}{\sqrt{2}} |{+}{+}\rangle_{A''C''} (|\phi^+\rangle_{A'C'}|\psi_b\rangle_{AC} + |\phi^-\rangle_{A'C'} Z^A|\psi_b\rangle_{AC}) & \text{if } b_2 = 0, \\ \frac{1}{\sqrt{2}} |{+}{+}\rangle_{A''C''} (|\psi^+\rangle_{A'C'}|\psi_b\rangle_{AC} - |\psi^-\rangle_{A'C'} Z^A|\psi_b\rangle_{AC}) & \text{if } b_2 = 1. \end{cases}$$

Then, after applying the controlled X gates, using Equations (15) and (17), we get

$$\begin{cases} \frac{1}{2} |{+}{+}\rangle_{A''C''} (|00\rangle_{A'C'} + (-1)^{b_1} |11\rangle_{A'C'})(\mathbb{1} + Z^A)|\psi_b\rangle_{AC} & \text{if } b_2 = 0, \\ \frac{1}{2} |{+}{+}\rangle_{A''C''} (|10\rangle_{A'C'} + (-1)^{b_1} |01\rangle_{A'C'})X^A(\mathbb{1} - Z^A)|\psi_b\rangle_{AC} & \text{if } b_2 = 1. \end{cases}$$

Let us call

$$\mathrm{O}(b) := \begin{cases} \frac{1}{\sqrt{2}}(\mathbb{1} + Z^A) & \text{if } b_2 = 0, \\ \frac{1}{\sqrt{2}} X^A(\mathbb{1} - Z^A) & \text{if } b_2 = 1, \end{cases}$$

and $|\varphi_b\rangle := \mathrm{O}(b)|\psi_b\rangle$. We can write

$$\begin{cases} |{+}{+}\rangle_{A''C''} |\phi^+\rangle_{A'C'}|\varphi_b\rangle_{AC} & \text{if } b = 00, \\ |{+}{+}\rangle_{A''C''} |\psi^+\rangle_{A'C'}|\varphi_b\rangle_{AC} & \text{if } b = 01, \\ |{+}{+}\rangle_{A''C''} |\phi^-\rangle_{A'C'}|\varphi_b\rangle_{AC} & \text{if } b = 10, \\ |{+}{+}\rangle_{A''C''} |\psi^-\rangle_{A'C'}|\varphi_b\rangle_{AC} & \text{if } b = 11, \end{cases}$$

We are going to define

$$|b\rangle := \begin{cases} |\phi^+\rangle & \text{if } b = 00, \\ |\psi^+\rangle & \text{if } b = 01, \\ |\phi^-\rangle & \text{if } b = 10, \\ |\psi^-\rangle & \text{if } b = 11. \end{cases}$$

With this notation, after Step 2 we are left with state $|{+}{+}\rangle_{A''C''} |b\rangle_{A'C'}|\varphi_b\rangle_{AC}$. This was to be expected, as the circuit until now is the famous `SWAP` circuit [**Figure 1**; MYS12].

Step 3. The unitary in Step 3 takes this state to

$$|b\rangle_{A'C'} \frac{1}{2}(|00\rangle_{A''C''} |\varphi_b\rangle_{AC} + |11\rangle_{A''C''} Y^A X^A \hat{Y}^C \hat{X}^C |\varphi_b\rangle_{AC} \tag{19}$$
$$+ |01\rangle_{A''C''} \hat{Y}^C \hat{X}^C |\varphi_b\rangle_{AC} + |10\rangle_{A''C''} Y^A X^A |\varphi_b\rangle_{AC}).$$

However, note that Equations (15) and (17) imply

$$\begin{aligned}
\hat{Y}^C\hat{X}^C|\varphi_b\rangle &= (-1)^{b_1+b_2+1}\hat{X}^C(\mathbb{1}+Z^A)Y^A|\psi_b\rangle \\
&= (-1)^{b_1+b_2}Y^A(\mathbb{1}-Z^A)X^C|\psi_b\rangle \\
&= (-1)^{b_2}Y^A(\mathbb{1}-Z^A)X^A|\psi_b\rangle \\
&= (-1)^{b_2}Y^AX^A|\varphi_b\rangle \qquad \text{if } b_2=0, \\
\hat{Y}^C\hat{X}^C|\varphi_b\rangle &= (-1)^{b_1}\hat{Y}^CX^A(\mathbb{1}-Z^A)X^A|\psi_b\rangle \\
&= (-1)^{b_1}(\mathbb{1}+X^A)\hat{Y}^C|\psi_b\rangle \\
&= (-1)^{b_2+1}(\mathbb{1}+X^A)Y^A|\psi_b\rangle \\
&= (-1)^{b_2+1}Y^AX^A|\varphi_b\rangle \qquad \text{if } b_2=1.
\end{aligned}$$

Both these equations are concisely written as,

$$\hat{Y}^C\hat{X}^C|\varphi_b\rangle = Y^AX^A|\varphi_b\rangle. \tag{20}$$

Symmetrically in Y and X, one obtains

$$\hat{X}^C\hat{Y}^C|\varphi_b\rangle = X^AY^A|\varphi_b\rangle. \tag{21}$$

In particular, one more use of Equation (15) yields

$$Y^AX^A\hat{Y}^C\hat{X}^C|\varphi_b\rangle = -|\varphi_b\rangle.$$

Therefore, Equation (19) becomes

$$|b\rangle_{A'C'}\frac{1}{\sqrt{2}}(|\phi^-\rangle_{A''C''}|\varphi_b\rangle_{AC} + |\psi^+\rangle_{A''C''}Y^AX^A|\varphi_b\rangle_{AC})$$

Step 4. Finally, using equation (18), we obtain the final state

$$|b\rangle_{A'C'}\frac{1}{\sqrt{2}}(|\psi^+\rangle_{A''C''}|\varphi_b\rangle_{AC} + |\phi^-\rangle_{A''C''}Y^AX^A|\varphi_b\rangle_{AC}). \tag{22}$$

Knowing what $U\otimes V\,|0000\rangle_{A'C'A''C''}\,|\psi_b\rangle$ is, we now take the partial trace over systems AC. We have that

$$\begin{aligned}
\mathrm{tr}_{AC}(U\otimes V\,|0000\rangle_{A'C'A''C''}\,|\psi_b\rangle) = \frac{1}{2}\,|b\rangle\langle b|_{A'C'}\otimes[(\Phi^-+\Psi^+)\,\langle\varphi_b|\varphi_b\rangle \\
+ (|\phi^-\rangle\langle\psi^+| + |\psi^+\rangle\langle\phi^-|)\,\langle\varphi_b|\,Y^AX^A\,|\varphi_b\rangle]_{A''C''}.
\end{aligned}$$

However, combining equations (20), (21) and (15), we obtain

$$\begin{aligned}
Y^AX^A|\varphi_b\rangle &= \mathrm{O}(b)\hat{Y}^C\hat{X}^C|\psi_b\rangle \\
&= (-1)^{b_1}\mathrm{O}(b)X^A\hat{Y}^C|\psi_b\rangle \\
&= (-1)^{b_1+b_2}\mathrm{O}(b)X^AY^A|\psi_b\rangle \\
&= (-1)^{b_1+b_2+1}\mathrm{O}(b)Y^AX^A|\psi_b\rangle \\
&= (-1)^{b_2+1}\hat{X}^C\mathrm{O}(b)Y^A|\psi_b\rangle \\
&= -\hat{X}^C\hat{Y}^C|\varphi_b\rangle \\
&= -X^AY^A|\varphi_b\rangle.
\end{aligned}$$

Finally note that, since we are in a real Hilbert space, the scalar product is symmetric. Therefore, since Alice operators are hermitian, we have

$$\langle\varphi_b|\, Y^A X^A\, |\varphi_b\rangle = \langle\varphi_b|\, X^A Y^A\, |\varphi_b\rangle = -\langle\varphi_b|\, Y^A X^A\, |\varphi_b\rangle\,,$$

and therefore $\langle\varphi_b|\, Y^A X^A\, |\varphi_b\rangle = 0$. On the other hand, by unitarity, the output state of Step 2 must have norm 1, and thus $\langle\varphi_b|\varphi_b\rangle = 1$. Therefore, the final state is

$$|b\rangle\langle b|_{A'C'} \otimes \left[\frac{\Psi^+ + \Phi^-}{2}\right]_{A''C''}. \tag{23}$$

Now, if we average over Bob's results, we obtain

$$\sum_{b=0}^{3} \mathrm{P}(b) \left(|b\rangle\langle b|_{A'C'} \otimes \left[\frac{\Psi^+ + \Phi^-}{2}\right]_{A''C''}\right) = \frac{\mathbb{1}_{A'C'}}{4} \otimes \left[\frac{\Psi^+ + \Phi^-}{2}\right]_{A''C''}. \tag{24}$$

One can now appreciate that the state $(\Psi^+ + \Phi^-)/2$ is exactly the same as the state given in equation (5). Indeed, in the basis $\{|00\rangle\,, |01\rangle\,, |10\rangle\,, |11\rangle\}$, and from equation (6), we have

$$\frac{\Psi^+ + \Phi^-}{2} = \frac{1}{4}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} = \frac{|i\rangle\langle i| \otimes |i\rangle\langle i| + |-i\rangle\langle -i| \otimes |-i\rangle\langle -i|}{2}.$$

In particular, it is not equal to its own partial transpose, so that

$$\left(\frac{\mathbb{1}_{A'C'}}{4} \otimes \left[\frac{\Psi^+ + \Phi^-}{2}\right]_{A''C''}\right)^{T_{A'A''}} \neq \frac{\mathbb{1}_{A'C'}}{4} \otimes \left[\frac{\Psi^+ + \Phi^-}{2}\right]_{A''C''}.$$

That is, our final state is not real-separable over the bipartition $A'A''|C'C''$. However, to obtain this state we have performed the operation

$$\begin{aligned}
&\sum_{b=0}^{3} \mathrm{P}(b) \mathrm{tr}_{AC}(U \otimes V |\psi_b\rangle_{AC}\, |0000\rangle_{A'C'A''C''}) \\
&\quad = \mathrm{tr}_{AB_1B_2C}((U \otimes V)\rho_{AB_1B_2C} \otimes |0000\rangle\langle 0000|_{A'C'A''C''}\, (U \otimes V^T) \\
&\quad = \int_\Lambda \mathrm{tr}_{AB_1B_2C}((U \otimes V)\sigma^\lambda_{AB_1} \otimes \sigma^\lambda_{B_2C} \otimes |0000\rangle\langle 0000|_{A'C'A''C''}\, (U \otimes V)^T)\, \mathrm{dP}_\Lambda(\lambda) \\
&\quad = \int_\Lambda \mathrm{tr}_{AB_1}(U(\sigma^\lambda_{AB_1} \otimes |00\rangle\langle 00|_{A'A''})U^T) \otimes \mathrm{tr}_{B_2C}(V(\sigma^\lambda_{B_2C} \otimes |00\rangle\langle 00|_{C'C''})V^T)\, \mathrm{dP}_\Lambda(\lambda),
\end{aligned}$$

which is manifestly real-separable over the bipartition $A'A''|C'C''$, reaching a contradiction. □

Note that it is irrelevant for the proof whether Alice and Charlie can *in practice* actually purify the system, or implement the unitaries given in Figure 7. Even if one modifies RQM to, for example, exclude Hadamard gates from the theory, the proof still reaches a mathematical contradiction: one can via some mathematical procedure obtain a state with a property that it cannot have under our hypotheses.

4.2. Robustness. Theorem 4.1.2 states that there is a behavior that we cannot reproduce exactly in RQM. However, a priori it could be possible to be able to approximate $\mathrm{P}_{\mathbb{C}}(a,b,c|x,z)$ as much as possible. In this section, we refine the proof of Theorem 4.1.2 to show that this is not the case. More precisely, we show

Theorem 4.2.1. [↓] *Let* $\mathrm{P}(a,b,c|x,z)$ *be a behavior with a well-defined marginal* $\mathrm{P}(b)$ *such that*

$$\left|\mathrm{P}(b)-\frac{1}{4}\right|<\varepsilon_c,\quad \mathcal{J}_b(\mathrm{P})>(6\sqrt{2}-\varepsilon_c)\,\mathrm{P}(b)$$

for all b*, and* $\varepsilon_c\approx 7.18\cdot 10^{-5}$*. Then,* $\mathrm{P}(a,b,c|x,z)$ *does not admit a decomposition of the form given by Equation (7) in RQM.*

The proof follows the same steps as the proof ot Theorem 4.1.2, but keeping track of all the small errors that accumulate from the imperfect violation of the inequality. The idea is to keep ε_c small enough so that after applying the unitary from Figure 7 and tracing out systems AC we are still away from the set of states that are invariant under partial transposition of $A'A''$. To make the estimates precise, we therefore need to compute the distance from that set to the output state. There are a variety of possible choices for the distance; we consider the distance given by the trace norm, and we'll need the following property:

Proposition 4.2.2 (The trace-norm is contractive under quantum channels)**.** *Let* $A\in\mathcal{T}(\mathcal{H})$ *be hermitian,* $S:\mathcal{T}(\mathcal{H})\to\mathcal{T}(\mathcal{H})$ *a positive trace-preserving map. We have*

$$\|S(A)\|_1\leq\|A\|_1.$$

In particular, this holds for quantum channels. That is, CPTP maps.

Proof. Indeed, from the spectral theorem 2.3.3 we may decompose $A=A^+-A^-$, where $A^+:=\theta(A), A^-:=\theta(-A)$ are positive operators. We also have $|A|=A^++A^-$. Notice that for any positive operator, the trace norm is just the trace, and in general for any hermitian operator we have $\|A\|_1=\mathrm{tr}(|A|)$. Therefore,

$$\begin{aligned}\|S(A)\|_1=\big\|S(A^+)-S(A^-)\big\|_1\leq\big\|S(A^+)\big\|_1+\big\|S(A^-)\big\|_1&=\mathrm{tr}\big(S(A^+)\big)+\mathrm{tr}\big(S(A^-)\big)\\&=\mathrm{tr}\big(S(A^++A^-)\big)=\mathrm{tr}(|A|)=\|A\|_1\end{aligned}$$

□

Lemma 4.2.3. *Let* $\mathcal{S}^{T_A}$ *be the set of states in* $\mathcal{H}_{A'}\otimes\mathcal{H}_{A''}\otimes\mathcal{H}_{C'}\otimes\mathcal{H}_{C''}$ *which are invariant under transposition on* $\mathcal{H}_{A'}\otimes\mathcal{H}_{A''}$*. Let*

$$\rho^*:=\frac{\mathbb{1}_{A'C'}}{4}\otimes\frac{\Psi^++\Phi^-}{2}=\frac{\mathbb{1}_{A'C'}}{4}\otimes\frac{|i\rangle\langle i|\otimes|i\rangle\langle i|+|-i\rangle\langle -i|\otimes|-i\rangle\langle -i|}{2}$$

be the state that gives the contradiction in the proof of Theorem 4.1.2. Then,

$$\inf_{\tau\in\mathcal{S}^{T_A}}\|\tau-\rho^*\|_1=1$$

Proof. We prove that 1 is both a lower and an upper bound. First, the upper bound. Notice that for $\tau^*:=\mathbb{1}/16$, we have

$$\tau^*-\rho^*=\frac{\mathbb{1}_{A'C'}}{4}\otimes\left(\frac{\mathbb{1}_{A''C''}}{4}-\frac{\Psi^++\Phi^-}{2}\right)=\frac{\mathbb{1}_{A'C'}}{4}\otimes\frac{\Phi^++\Psi^--\Psi^+-\Phi^-}{4}.$$

In particular,

$$\|\tau^* - \rho^*\|_1 = \mathrm{tr}\left(\frac{\mathbb{1}_{A'C'}}{4}\right)\mathrm{tr}\left(\frac{\Phi^+ + \Psi^- + \Psi^+ + \Phi^-}{4}\right) = \mathrm{tr}\left(\frac{\mathbb{1}_{A'A''C'C''}}{16}\right) = 1.$$

For the lower bound, we first compute $\left\|\rho^* - \rho^{*T_{A'A''}}\right\|_1$. Writing

$$\rho^* = \frac{1}{2}\left(\frac{\mathbb{1}_{A'} \otimes |i\rangle\langle i|_{A''}}{2} \otimes \frac{\mathbb{1}_{C'} \otimes |i\rangle\langle i|_{C''}}{2} + \frac{\mathbb{1}_{A'} \otimes |-i\rangle\langle -i|_{A''}}{2} \otimes \frac{\mathbb{1}_{C'} \otimes |-i\rangle\langle -i|_{C''}}{2}\right),$$

and recalling that $|\pm i\rangle\langle \pm i|^T = |\mp i\rangle\langle \mp i|$, and that $|i\rangle\langle i| - |-i\rangle\langle -i| = \sigma_y$, we can see that

$$\begin{aligned}\rho^* - \rho^{*T_{A'A''}} &= \frac{1}{2}\left(\frac{\mathbb{1}_{A'} \otimes |i\rangle\langle i|_{A''}}{2} \otimes \frac{\mathbb{1}_{C'} \otimes (\sigma_y)_{C''}}{2} - \frac{\mathbb{1}_{A'} \otimes |-i\rangle\langle -i|_{A''}}{2} \otimes \frac{\mathbb{1}_{C'} \otimes (\sigma_y)_{C''}}{2}\right) \\ &= \frac{1}{2}\frac{\mathbb{1}_{A'} \otimes (\sigma_y)_{A''}}{2} \otimes \frac{\mathbb{1}_{C'} \otimes \sigma_{y C''}}{2} \\ &= \frac{\mathbb{1}_{A'} \otimes \mathbb{1}_{C'} \otimes \sigma_y \otimes \sigma_y}{8}.\end{aligned}$$

Thus, $\left\|\rho^* - \rho^{*T_{A'A''}}\right\|_1 = 2$. On the other hand, we define the quantum channel

$$\begin{aligned} S : \mathcal{H}_{A'} \otimes \mathcal{H}_{A''} &\longrightarrow \mathcal{H}_{A'} \otimes \mathcal{H}_{A''} \\ \rho &\longmapsto \sum_{\alpha',\alpha'' \in \{i,-i\}} |\alpha'\rangle\langle\alpha'| \otimes |\alpha''\rangle\langle\alpha''| \, \rho \, |\alpha'\rangle\langle\alpha'| \otimes |\alpha''\rangle\langle\alpha''|, \end{aligned}$$

where $|i\rangle, |-i\rangle$ are the eigenvectors of the Pauli y gate σ_y. Given that $|i\rangle, |-i\rangle$ form an orthonormal set, it follows that $S^2 = S$. Writing

$$\rho^* = \frac{1}{2}\left(\frac{\mathbb{1}_{A'} \otimes |i\rangle\langle i|_{A''}}{2} \otimes \frac{\mathbb{1}_{C'} \otimes |i\rangle\langle i|_{C''}}{2} + \frac{\mathbb{1}_{A'} \otimes |-i\rangle\langle -i|_{A''}}{2} \otimes \frac{\mathbb{1}_{C'} \otimes |-i\rangle\langle -i|_{C''}}{2}\right)$$

it also becomes self-evident that $S \otimes \mathbb{1}_{C'C''}(\rho^*) = \rho^*$. From Proposition 4.2.2, we deduce that we only need to consider states of the form $S \otimes \mathbb{1}_{C'C''}(\tau) = \tau$. Any such state is of the form

$$\tau = \sum_{\alpha',\alpha'' \in \{i,-i\}} |\alpha'\rangle\langle\alpha'|_{A'} \otimes |\alpha''\rangle\langle\alpha''|_{A''} \otimes \tau_{C'C''}^{\alpha'\alpha''}$$

for some arbitrary $\tau_{C'C''}^{\alpha'\alpha''}$. On the other hand, since $|\alpha\rangle\langle\alpha|^T = |-\alpha\rangle\langle-\alpha| = \sigma_z |\alpha\rangle\langle\alpha| \sigma_z$, we have that $\tau^{T_{A'A''}} = (\sigma_z{}^{\otimes 2} \otimes \mathbb{1}_{C'C''})\tau(\sigma_z{}^{\otimes 2} \otimes \mathbb{1}_{C'C''})$ for all τ fixed by $S \otimes \mathbb{1}_{C'C''}$, including ρ^*. Therefore, using Proposition 4.2.2, we obtain $\left\|\tau^{T_{A'A''}}\right\|_1 = \|\tau\|_1$ for all such τ. Thus,

$$2 = \left\|\rho^* - \rho^{*T_{A'A''}}\right\|_1 \leq \|\rho^* - \tau\|_1 + \left\|\rho^{*T_{A'A''}} - \tau\right\|_1 = 2\|\rho^* - \tau\|_1,$$

whenever $\tau^{T_{A'A''}} = \tau$, obtaining the promised lower bound. □

We are now ready to proof Theorem 4.2.1. We will follow the steps of the proof of Theorem 4.1.2, while keeping track of all the errors that accumulate.

Proof of Theorem 4.2.1. Suppose that Alice, Bob and Charlie share states and measurements in RQM which satisfy Equation (7), where the behaviour $\mathrm{P}(a, b, c|x, z)$ satisfies that

$$\left|\mathrm{P}(b) - \frac{1}{4}\right| < \varepsilon, \quad \mathcal{J}_b(\mathrm{P}) > (6\sqrt{2} - \varepsilon)\,\mathrm{P}(b)$$

for some $\varepsilon > 0$. Consider, as in the proof of Theorem 4.1.2, the purified states $|\psi_b\rangle$, which satisfy Equation (10) and thus we have

$$\langle\psi_b|\,\hat{\mathcal{J}}_b|\psi_b\rangle = \frac{\mathcal{J}_b(\mathrm{P})}{\mathrm{P}(b)} > (6\sqrt{2}-\varepsilon).$$

Equivalently, we obtain that

$$\langle\psi_b|\,\sqrt{2}(6\sqrt{2}-\hat{\mathcal{J}}_b)|\psi_b\rangle < \sqrt{2}\varepsilon.$$

For small enough ε, we can approximate the self-tested states and measurements of the proof of Theorem 4.1.2 using the same SOS decompositions for $\sqrt{2}(6\sqrt{2}-\hat{\mathcal{J}}_b)$. Indeed, suppose

$$\sqrt{2}(6\sqrt{2}-\hat{\mathcal{J}}_b) = \sum_i P_i^\dagger P_i.$$

Then, it must be by positivity that $\langle\psi_b|\,P_i^\dagger P_i|\psi_b\rangle < \sqrt{2}\varepsilon$ for all i or, equivalently, that

$$\|P_i|\psi_b\rangle\| < \sqrt{\sqrt{2}\varepsilon}.$$

Let us then define $\varepsilon_1 := \sqrt{\sqrt{2}\varepsilon}$. As before, we also denote

$$Z^A := A^1, \quad X^A := A^2, \quad Y^A := A^3,$$
$$Z^C := \frac{C^1+C^2}{\sqrt{2}}, \quad X^C := \frac{C^1-C^2}{\sqrt{2}}, \quad Y^C := \frac{C^3-C^4}{\sqrt{2}}.$$

Note that now Equation (12) holds now only approximately, so we don't have, for example, that

$$\mathbb{1}_{\mathcal{H}_A}\otimes\frac{C^1+C^2}{\sqrt{2}}|\psi_b\rangle = \mathbb{1}_{\mathcal{H}_A}\otimes\frac{C^3+C^4}{\sqrt{2}}|\psi_b\rangle.$$

Therefore, in the definition of Z^C that we are using, we have chosen the operator $(C^1+C^2)/\sqrt{2}$ over $(C^3+C^4)/\sqrt{2}$, which are now different operators, even when restricted to $|\psi_b\rangle$. This choice in practice does not matter, since we are only using the fact that this operator is ε_1-close to Z^A.

Using the SOS decompositions (11) and (14), we find that, analogous to Equations (12, 15), that

$$\left\|((-1)^{b_2}Z^A - Z^C)|\psi_b\rangle\right\|, \quad \left\|((-1)^{b_1}X^A - X^C)|\psi_b\rangle\right\|, \quad \left\|((-1)^{b_1+b_2}Y^A - Y^C)|\psi_b\rangle\right\| \leq \varepsilon_1 \tag{25}$$

and

$$\|\{Z^A, X^A\}|\psi_b\rangle\|, \|\{Z^A, Y^A\}|\psi_b\rangle\|, \|\{Y^A, X^A\}|\psi_b\rangle\| \leq 2(1+\sqrt{2})\varepsilon_1 \tag{26}$$

hold. Finally, Charlie's regularized operators are also approximate equal to the unregularized ones, as in Equation (16):

$$\left\|(\hat{Z}^C - Z^C)|\psi_b\rangle\right\|, \left\|(\hat{Y}^C - Y^C)|\psi_b\rangle\right\|, \left\|(\hat{X}^C - X^C)|\psi_b\rangle\right\| \leq \varepsilon_1 \tag{27}$$

Using all of these relations, we can bound the distance from the state obtained after performing the isometry given in Figure 7 to $|0000\rangle_{A'A''C'C''}|\psi_b\rangle$ and then tracing out systems AC to the state obtained in the ideal case, show in the proof of Theorem 4.1.2. It is convenient to do this in several steps and use the triangle inequality. First, we just

apply the isometry given in Figure 7. In the case where $\varepsilon = 0$, we found out that the output state, given by Equation (22) is

$$|\sigma_b\rangle := |b\rangle_{A'C'} \frac{1}{\sqrt{2}}(|\psi^+\rangle_{A''C''}|\varphi_b\rangle_{AC} + |\phi^-\rangle_{A''C''} Y^A X^A |\varphi_b\rangle_{AC}).$$

Note now that since Alice and Charlie's measurement operators are not ideal anymore, this state is now in principle unnormalized. This is because the relations that we used to arrive at this expression are not satisfied anymore. However, we can bound its norm as follows

Lemma 4.2.4.

$$1 - (3+\sqrt{2})\varepsilon_1 \leq \||\sigma_b\rangle\|^2 \leq 1 + (3+\sqrt{2})\varepsilon_1$$

Proof. Note that

$$\||\sigma_b\rangle\|^2 = \frac{1}{2}(\langle\varphi_b|\varphi_b\rangle + \langle\varphi_b| X^A Y^A Y^A X^A |\varphi_b\rangle) = \langle\varphi_b|\varphi_b\rangle .$$

Recalling the definition of $|\varphi_b\rangle = \mathrm{O}(b)|\psi_b\rangle$, we have

$$\langle\varphi_b|\varphi_b\rangle = \frac{1}{2}\langle\psi_b| (1 \pm Z^A)^2 |\psi_b\rangle = 1 \pm \langle\psi_b| Z^A |\psi_b\rangle,$$

where the sign depends on b. Therefore, $\left|\||\sigma_b\rangle\|^2 - 1\right| = \left|\langle\psi_b| Z^A |\psi_b\rangle\right|$. The following estimates are inspired by [MYS12],

$$\begin{aligned}
\left|\langle\psi_b| Z^A |\psi_b\rangle + \langle\psi_b| Z^A X^A \hat{X}^C |\psi_b\rangle\right| &= \left|\langle\psi_b| Z^A |\psi_b\rangle + \langle\psi_b| \hat{X}^C X^A Z^A |\psi_b\rangle\right| \\
&= \left|\langle\psi_b| (\hat{X}^C Z^A \hat{X}^C + \hat{X}^C X^A Z^A) |\psi_b\rangle\right| \\
&\leq \left\|\hat{X}^C |\psi_b\rangle\right\| \left\|(Z^A \hat{X}^C + X^A Z^A)|\psi_b\rangle\right\| \leq (4 + 2\sqrt{2})\varepsilon_1. \\
\left|\langle\psi_b| Z^A |\psi_b\rangle - \langle\psi_b| Z^A X^A \hat{X}^C |\psi_b\rangle\right| &\leq \left\|(Z^A(\mathbb{1} - X^A \hat{X}^C)|\psi_b\rangle\right\| \\
&\leq \left\|Z^A(\hat{X}^C - X^A)|\psi_b\rangle\right\| \leq 2\varepsilon_1,
\end{aligned}$$

where we used the Cauchy-Schwarz inequality as well as

$$\left\|(Z^A \hat{X}^C + X^A Z^A)|\psi_b\rangle\right\| \leq \left\|(Z^A \hat{X}^C - Z^A X^A)|\psi_b\rangle\right\| + \left\|\{X^A, Z^A\}|\psi_b\rangle\right\| \leq 2\varepsilon_1 + 2(1+\sqrt{2})\varepsilon_1.$$

Therefore, we find that

$$\left|\langle\psi_b| Z^A |\psi_b\rangle\right| \leq (3+\sqrt{2})\varepsilon_1, \tag{28}$$

which implies the statement. □

With the norm controlled, we can now bound the distance between the actual state that we get after applying the isometry defined in Figure 7 and this semi-ideal state. Let us call $|\rho_b^\varepsilon\rangle = U \otimes V(|0000\rangle_{A'A''C'C''} |\psi_b\rangle_{AC})$. We have

Lemma 4.2.5.

$$\||\rho_b^\varepsilon\rangle - |\sigma_b\rangle\| \leq (15 + 13\sqrt{2})\varepsilon_1.$$

Proof. The proof is a repeated application of the triangle inequality using every intermediate step in the application of the unitary of Figure 7. We apply it step by step while keeping track of all the errors that appear whenever we use one of the expressions (25), (26), (27). For example, we first apply the Hadamard gates, and then the controled Z^A and $\hat{Z}^C$ gates, which return a state as written in the first line of "Step 1" below. Since so far we did not use any of the approximate relations, the error incurred to in this step is 0, as written on the left of the state in "Step 1". Next, we use the approximate relation (25) to convert $Z^A\hat{Z}^C|\psi_b\rangle$ to $(-1)^{b_2}Z^AZ^A|\psi_b\rangle = (-1)^{b_2}|\psi_b\rangle$, and similarly $\hat{Z}^C|\psi_b\rangle$ to $\hat{Z}^A|\psi_b\rangle$, incurring into an error $2\varepsilon_1$, written on the left. Continuing this way through the circuit, the intermediate expressions, together with their bounds, are the following:

$$\begin{aligned}
\text{Step 1:}\qquad & 0\varepsilon_1 : \frac{1}{2}\Big[|00\rangle + |11\rangle\, Z^A\hat{Z}^C + |01\rangle\, \hat{Z}^C + |10\rangle\, Z^A\Big]\,|\psi_b\rangle \\
& 2\varepsilon_1 : \frac{1}{2}\Big[(|00\rangle + (-1)^{b_2}\,|11\rangle)\mathbb{1} + ((-1)^{b_2}\,|01\rangle + |10\rangle)Z^A\Big]\,|\psi_b\rangle \\
\text{Step 2:}\qquad & 0\varepsilon_1 : \frac{1}{4}\Big[\,|00\rangle\,(1+(-1)^{b_2})(\mathbb{1}+Z^A) + |11\rangle\,(1+(-1)^{b_2})X^A\hat{X}^C(\mathbb{1}-Z^A) + \\
& \qquad |01\rangle\,(1-(-1)^{b_2})\hat{X}^C(\mathbb{1}+Z^A) + |10\rangle\,(1-(-1)^{b_2})X^A(\mathbb{1}-Z^A)]\Big]\,|\psi_b\rangle \\
& (3+\sqrt{2})\varepsilon_1 : \frac{1}{4}\Big[(1+(-1)^{b_2})(|00\rangle + (-1)^{b_1}\,|11\rangle)(\mathbb{1}+Z^A) + \\
& \qquad (1-(-1)^{b_2})((-1)^{b_1}\,|01\rangle + |10\rangle)X^A(\mathbb{1}-Z^A)]\Big]\,|\psi_b\rangle
\end{aligned}$$

Note that in the last approximation, the two paths $b_2 = 0$ and $b_2 = 1$ have the same upper bound.

The rest of the circuit does not involve $A'C'$ so we can safely ignore them. For the remaining steps we obtain

$$\begin{aligned}
\text{Step 3:}\qquad & 0\varepsilon_1 : \frac{1}{2}\Big[\,|00\rangle\,\mathcal{O}(b) + |11\rangle\,Y^AX^A\hat{Y}^C\hat{X}^C\mathrm{O}(b) + \\
& \qquad |01\rangle\,\hat{Y}^C\hat{X}^C\mathrm{O}(b) + |10\rangle\,Y^AX^A\mathrm{O}(b)\Big]\,|\psi_b\rangle \\
& (8+10\sqrt{2})\varepsilon_1 : \frac{1}{2}\Big[\,|00\rangle\,\mathrm{O}(b) + |11\rangle\,Y^AX^AY^AX^A\mathrm{O}(b) \\
& \qquad + |01\rangle\,Y^AX^A\mathrm{O}(b) + |10\rangle\,Y^AX^A\mathrm{O}(b)\Big]\,|\psi_b\rangle \\
& (2+2\sqrt{2})\varepsilon_1 : \frac{1}{2}\left[(|00\rangle - |11\rangle)\mathrm{O}(b) + (|01\rangle + |10\rangle)Y^AX^A\mathcal{O}(b)\right]|\psi_b\rangle \\
\text{Step 4:}\qquad & 0\varepsilon_1 : \frac{1}{2}\left[(|01\rangle + |10\rangle)\mathrm{O}(b) + (|00\rangle - |11\rangle)Y^AX^A\mathrm{O}(b)\right]|\psi_b\rangle
\end{aligned}$$

where we have taken the larger bound among the two $b_2 = 1$ and $b_2 = 0$ cases to get

$$\left\|(\hat{Y}^C\hat{X}^C\mathrm{O}(b) - Y^AX^A\mathrm{O}(b))|\psi_b\rangle\right\| \leq (20+8\sqrt{2})\varepsilon_1/\sqrt{2} \text{ and} \tag{29}$$

$$\left\|(Y^AX^AY^AX^A\mathrm{O}(b) + \mathrm{O}(b)))|\psi_b\rangle\right\| \leq (8+4\sqrt{2})\varepsilon_1/\sqrt{2}\,. \tag{30}$$

By a series of triangle inequalities going through all the intermediate expressions, we get the Lemma. □

Finally, we consider these states after tracing out systems AC, since these are the relevant states for the proof of Theorem 4.1.2. We will call $\sigma_b := \mathrm{tr}_{AC}(|\sigma_b\rangle\langle\sigma_b|)$, and $\rho_b^\varepsilon := \mathrm{tr}_{AC}(|\rho_b^\varepsilon\rangle\langle\rho_b^\varepsilon|)$. Summing over all of Bob's outcomes, we recover our state of interest

$$\rho^\varepsilon := \sum_b \mathrm{P}(b)\rho_b^\varepsilon.$$

Lemma 4.2.6. *For small enough ε_1,*

$$\left\|\rho^\varepsilon - \sum_b \mathrm{P}(b)\sigma_b\right\|_1 \leq 2\sqrt{\left(\frac{2+(3+\sqrt{2})\varepsilon_1}{2}\right)^2 - \left(1 - \frac{(3+\sqrt{2})\varepsilon_1 + (15+13\sqrt{2})^2{\varepsilon_1}^2}{2}\right)^2}.$$

Proof. We are going to adapt an argument from [**Theorem 3.33**; Wat18] to unnormalized states. Namely, since $|\rho_b^\varepsilon\rangle\langle\rho_b^\varepsilon| - |\sigma_b\rangle\langle\sigma_b|$ has at most rank 2, let λ_1, λ_2 be its two possibly non-zero eigenvalues. Then, we have

$$\begin{aligned}\lambda_1 + \lambda_2 &= \mathrm{tr}(|\rho_b^\varepsilon\rangle\langle\rho_b^\varepsilon| - |\sigma_b\rangle\langle\sigma_b|) &&= 1 - \||\sigma_b\rangle\|^2\\ \lambda_1^2 + \lambda_2^2 &= \mathrm{tr}\big((|\rho_b^\varepsilon\rangle\langle\rho_b^\varepsilon| - |\sigma_b\rangle\langle\sigma_b|)^2\big) &&= 1 + \||\sigma_b\rangle\|^4 - 2|\langle\rho_b^\varepsilon|\sigma_b\rangle|^2,\end{aligned}$$

Solving explicitely for λ_1, λ_2 in this system of equations shows that

$$\||\rho_b^\varepsilon\rangle\langle\rho_b^\varepsilon| - |\sigma_b\rangle\langle\sigma_b|\|_1 = |\lambda_1| + |\lambda_2| = 2\sqrt{\left(\frac{1+\||\sigma_b\rangle\|^2}{2}\right)^2 - \left|\langle\rho_b^\varepsilon|\sigma_b\rangle\right|^2}.$$

On the other hand,

$$\langle\rho_b^\varepsilon|\sigma_b\rangle = \frac{1 + \||\sigma_b\rangle\|^2 - \|\rho_b^\varepsilon - |\sigma_b\rangle\|^2}{2}.$$

The Lemma then follows by application of Lemmas 4.2.4 and 4.2.5. □

We will also now name the state obtained in Equation (23) by

$$\rho_b^* := |b\rangle\langle b|_{A'C'} \otimes \left[\frac{\Psi^+ + \Phi^-}{2}\right]_{A''C''}.$$

Recall also from Equation (24) that the state from Lemma 4.2.3 we want to approximate is $\rho^* = \sum_b \rho_b^*/4$.

Lemma 4.2.7.

$$\|\sigma_b - \rho_b^*\|_1 \leq (7 + 3\sqrt{2})\varepsilon_1.$$

Proof. let us first separate the expression in the norm as

$$\begin{aligned}\|\sigma_b - \rho_b^*\|_1 \leq |\hat{\mu}(b)|\big\||b\rangle\langle b| \otimes \big(|\phi^-\rangle\langle\psi^+| + |\psi^+\rangle\langle\phi^-|\big)\big\|_1\\ + |\mu(b)|\big\||b\rangle\langle b| \otimes \big(|\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+|\big)\big\|_1,\end{aligned}$$

where the coefficients simplify to

$$|\mu(b)| = \frac{1}{2}\left|\langle\psi_b|\, Z^A |\psi_b\rangle\right|$$

and

$$|\hat{\mu}(b)| = \begin{cases}\frac{1}{4}\left|\langle\psi_b|\,(\mathbb{1} + Z^A)X^A Y^A(\mathbb{1} + Z^A)|\psi_b\rangle\right| & b_2 = 0\\ \frac{1}{4}\left|\langle\psi_b|\,(\mathbb{1} - Z^A)X^A Y^A(\mathbb{1} - Z^A)|\psi_b\rangle\right| & b_2 = 1.\end{cases}$$

To bound $|\hat{\mu}(b)|$, let us consider

$$\begin{aligned}
&\frac{1}{4}\Big|\langle\psi_b|\,(\mathbb{1}+(-1)^{b_2}Z^A)\{X^A,Y^A\}(\mathbb{1}+(-1)^{b_2}Z^A)|\psi_b\rangle\Big| \\
&\quad\le \frac{1}{4}\Big\|(\mathbb{1}+(-1)^{b_2}Z^A)|\psi_b\rangle\Big\|\Big\|\{X^A,Y^A\}(\mathbb{1}+(-1)^{b_2}Z^A)|\psi_b\rangle\Big\| \\
&\quad\le \frac{1}{2}\Big\|\{X^A,Y^A\}(\mathbb{1}+(-1)^{b_2}Z^A)|\psi_b\rangle\Big\| \\
&\quad\le \frac{1}{2}\left(\|\{X^A,Y^A\}|\psi_b\rangle\| + \Big\|\hat{Z}^C\{X^A,Y^A\}|\psi_b\rangle\Big\| + \Big\|\{X^A,Y^A\}((-1)^{b_2}Z^A-\hat{Z}^C)|\psi_b\rangle\Big\|\right) \\
&\quad\le (4+2\sqrt{2})\varepsilon_1.
\end{aligned}$$

This, together with Equation (28), produces the desired bound as

$$\|\sigma_b-\rho_b^*\|_1 \le 2(|\hat{\mu}(b)|+|\mu(b)|) \le (4+2\sqrt{2})\varepsilon_1+(3+\sqrt{2})\varepsilon_1 = (7+3\sqrt{2})\varepsilon_1.$$

□

The theorem now follows from Lemmas 4.2.6, 4.2.7 and 4.2.3 as follows. We have, using Lemmas 4.2.6, 4.2.7 and a small computation that, for $\varepsilon \le \varepsilon_c$ and $\mathrm{P}(a,b,c|x,z)$ satisfying the hypotheses of the Theorem,

$$\|\rho^*-\rho^\varepsilon\|_1 \le \left\|\rho^\varepsilon-\sum_b \mathrm{P}(b)\sigma_b\right\|_1 + \left\|\sum_b \mathrm{P}(b)\sigma_b-\rho^*\right\|_1 < 1.$$

Therefore, by Lemma 4.2.3 we know that ρ^ε is not invariant under transposition on systems $A'A''$. However, just like in Theorem 4.1.2, the way we have computed this state would ensure that it would be invariant. Therefore, the original decomposition cannot exist. □

4.3. Numerics. We have seen that there is a gap between the maximum value we can achieve with the CHSH_3 inequality in QM and RQM. However, this gap is, so far, too small to be detectable in an actual experiment. In this Section, using standard numerical tools, we prove the existence of a bigger gap.

4.3.1. *Semidefinite Programming hierarchies.* We have an optimization problem of the form

$$\begin{aligned}
&\sup_{\mathcal{H},\rho,\{A^i\}_{i\in\mathcal{X}},\{B_j\}_{j\in\mathcal{B}},\{C^k\}_{k\in\mathcal{Z}}} && \mathrm{tr}\Big(Q(A^i,B_j,C^k)\rho\Big) \\
&\text{s.t.} && \rho \text{ is a state and } A^i,B_j,C^k \text{ are measurement operators}
\end{aligned} \tag{31}$$

where Q is a given non-commutative tensor polynomial, such as CHSH_3. In practice, one usually forgets about the tensor product structure and includes the commutation of variables corresponding to different parties as an extra constraint. Technically, this changes the problem [JNV+20], but in practice this is good enough. That is, we usually

have a problem of the form

$$\begin{aligned} \sup_{\{A_i\}_{i=1}^n,\rho,\mathcal{H}} \quad & \mathrm{tr}(p(A_1,...,A_n)\rho) \\ \text{s.t.} \quad & \mathcal{H} \text{ is a Hilbert space} \\ & A_i \in \mathrm{B}(\mathcal{H}) \text{ self-adjoint} \\ & \rho \in \mathcal{S}(\mathcal{H}) \\ & h_i(A_1,...,A_n) = 0,\ i = 1,...,k \end{aligned} \tag{32}$$

where h_i are some non-commutative polynomials, such as $[A_i, A_j]$, A_i^2, or $A_i^2 - A - i$. This kind of problems are hard since in particular they contain all polynomial optimization problems, which are NP-hard [Lau09]. Our problem is particularly untractable, since we are also optimizing over the class of all (real) Hilbert spaces. However, it is possible to compute upper bounds with a hierarchy of Semidefinite Programs (SDPs), a collection of a kind of problems which lie in P [VB96], each one giving a better upper bound than the previous one.

The Navascués-Pironio-Acín (NPA) hierarchy was introduced in [NPA07; NPA08] precisely to do this. The idea is the following: suppose that there is a Hilbert space $\mathcal{H}$, a state $\psi \in \mathcal{H}$ and self-adjoint operators $\{A^x, C^z\}_{i\in\mathcal{X},k\in\mathcal{Z}}$ such that a certain conditional probability distribution comes out as the result of a quantum experiment $\mathrm{P}(a,c|x,z) = \langle\psi| A_a^x C_c^z |\psi\rangle$. We can assume that we have projective measurements and pure states, since we are not restricting the Hilbert space dimension. If this is the case, then one can also consider terms of the form $\langle\psi| w |\psi\rangle$, where w is a word on the letters $\Sigma := \{A_a^x, C_c^z\}_{a\in\mathcal{A},c\in\mathcal{C},x\in\mathcal{X},z\in\mathcal{Z}}$ of length $n \in \mathbb{N}$. We always consider reduced words under the equivalence relation given by

$$\begin{aligned} w_1 E E w_2 &\sim w_1 E w_2 \quad \text{for all } E \in \Sigma, w_1, w_2 \in \Sigma^* \\ w_1 A_a^x C_c^z w_2 &\sim w_1 C_c^z A_a^x w_2 \quad \text{for all } x \in \mathcal{X}, z \in \mathcal{Z}, a \in \mathcal{A}, b \in \mathcal{B}, w_1, w_2 \in \Sigma^* \end{aligned}$$

Consider now, for each $n \in \mathbb{N}$ the matrix

$$\Gamma^n := \left(\langle\psi| w_i^\dagger w_j |\psi\rangle\right)_{w_i,w_j\in\Sigma^{\leq n}}, \tag{33}$$

This is the Gram matrix of the vectors $\{w|\psi\rangle\}_{w\in\Sigma^{\leq n}}$, and is therefore positive semidefinite. It is called a *moment matrix*. On the other hand, some of its entries are determined by the fact that $\mathrm{P}(a,c|x,z) = \langle\psi| A_a^x C_c^z |\psi\rangle$. So if we are maximizing a linear functional on $\mathrm{P}(a,c|x,z)$ with the constraints that this behavior has to come from a quantum state and commuting measurements, we can instead consider the problem of maximizing the same functional with the constraint that the matrix Γ^n is positive semidefinite. For each $n \in \mathbb{N}$, this is now an SDP which provides an upper bound to the original problem, which in the limit of $n \to \infty$ converges to the actual solution [NPA08]. Such a relaxation of an optimization problem is called an SDP *hierarchy*.

Using words with letters in different alphabets, one can obtain variations of the NPA hierarchy. Let us illustrate how this works via the standard example:

Example 4.3.1 (CHSH)**.** We are trying to compute the optimal value of the CHSH inequality with these techniques. That is, we want to solve the problem

$$
\begin{aligned}
(34) \qquad & \sup_{\mathcal{H},\psi,\{A^x,C^z\}_{x,z=0,1}} && \langle\psi|\left(A^0C^0 + A^1C^0 + A^0C^1 - A^1C^1\right)|\psi\rangle \\
& \text{s.t.} && \mathcal{H} \text{ is a Hilbert space} \\
& && |\psi\rangle \in \mathcal{H}, A^x, C^z \in \mathrm{B}(\mathcal{H}) \\
& && [A^x, C^z] = 0, \\
& && A^{x\dagger} = A^x, C^{z\dagger} = C^z, (A^x)^2 = \mathbb{1}, (C^z)^2 = \mathbb{1}.
\end{aligned}
$$

The last constraint states that we have a dichotomic observable. We can therefore construct the NPA hierarchy in the alphabet $\Sigma := \{A^x, C^z\}_{x,z=0,1}$, with the equivalence relations given by the constraints of the problem:

$$
\begin{aligned}
w_1 E E w_2 &\sim w_1 w_2 \quad \text{for all } E \in \Sigma, w_1, w_2 \in \Sigma^* \\
w_1 A^x C^z w_2 &\sim w_1 C^z A^x w_2 \quad \text{for all } x, z \in \{0,1\}, w_1, w_2 \in \Sigma^*.
\end{aligned}
$$

If we go to level 1 of the NPA hierarchy, we obtain the moment matrix

$$
\Gamma^1 = \begin{pmatrix}
1 & \mathrm{tr}\left(A^0\rho\right) & \mathrm{tr}\left(A^1\rho\right) & \mathrm{tr}\left(C^0\rho\right) & \mathrm{tr}\left(C^1\rho\right) \\
\cdot & 1 & \mathrm{tr}\left(A^0A^1\rho\right) & \mathrm{tr}\left(A^0C^0\rho\right) & \mathrm{tr}\left(A^0C^1\rho\right) \\
\cdot & \cdot & 1 & \mathrm{tr}\left(A^1C^0\rho\right) & \mathrm{tr}\left(A^1C^1\rho\right) \\
\cdot & \cdot & \cdot & 1 & \mathrm{tr}\left(C^0C^1\rho\right) \\
\cdot & \cdot & \cdot & \cdot & 1
\end{pmatrix}.
$$

Introducing a variable $\{d_{w_i^\dagger w_j}\}_{w_i,w_j\in\Sigma^{\leq 1}}$ for each entry of the matrix, and taking into account all the relations coming from the constraints of the original problem, we can relax the original problem to the following

$$
\begin{aligned}
& \sup_{d_w} && d_{A^0B^0} + d_{A^1B^0} + d_{A^0B^1} - d_{A^1B^1} \\
& \text{s.t.} && \Gamma^1 \geq 0, \\
(35) \qquad & && d_{A^x}, d_{B^y} \text{ behave as } \mathrm{P}(0|x) - \mathrm{P}(1|x), \mathrm{P}(0|y) - \mathrm{P}(1|y), \text{respectively} \\
& && d_{A^xB^y} \text{ behave as } \mathrm{P}(0,0|x,y) + \mathrm{P}(1,1|x,y) - \mathrm{P}(0,1|x,y) - \mathrm{P}(1,0|x,y).
\end{aligned}
$$

where

$$
\Gamma^1 := \begin{pmatrix}
1 & d_{A^0} & d_{A^1} & d_{B^0} & d_{B^1} \\
\cdot & 1 & d_{A^0A^1} & d_{A^0B^0} & d_{A^0B^1} \\
\cdot & \cdot & 1 & d_{A^1B^0} & d_{A^1B^1} \\
\cdot & \cdot & \cdot & 1 & d_{B^0B^1} \\
\cdot & \cdot & \cdot & \cdot & 1
\end{pmatrix}.
$$

4.3.2. *The* CHSH_3 *scenario.* With these tools, we take a look at the inequality used to prove Theorems 4.1.2 and 4.2.1

In this case, we consider the operator

$$
\begin{aligned}
\hat{\mathcal{J}} := &(B_{00} + B_{10} - B_{01} - B_{11})A_1(C_1 + C_2 + C_3 + C_4) \\
&+ (B_{00} - B_{10} + B_{01} - B_{11})A_2(C_1 - C_2 + C_5 + C_6) \\
&+ (-B_{00} + B_{10} + B_{01} - B_{11})A_3(C_3 - C_4 + C_5 - C_6).
\end{aligned}
$$

This corresponds to the sum of the operators corresponding to CHSH_3 after Bob has obtained a certain measurement outcome b. That is,

$$\hat{\mathcal{J}} = \sum_{b=0}^{3} \hat{\mathcal{J}}_b.$$

Therefore, we have the following optimization problem

$$\begin{aligned}
\sup_{\mathcal{H},\rho,\{A^i\}_{i\in\mathcal{X}},\{B_j\}_{j\in\mathcal{B}},\{C^k\}_{k\in\mathcal{Z}}} \quad & \mathrm{tr}\Big(\hat{\mathcal{J}}\rho\Big) \\
\text{s.t.} \quad & \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2} \otimes \mathcal{H}_C \\
& \mathcal{H}_A, \mathcal{H}_{B_1}, \mathcal{H}_{B_2}, \mathcal{H}_C \text{ are real Hilbert spaces} \\
& \rho = \int_\Lambda \rho_1^\lambda \otimes \rho_2^\lambda \, \mathrm{dP}_\Lambda(\lambda) \\
& \rho_1^\lambda \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_{B_1}), \rho_2^\lambda \in \mathcal{S}(\mathcal{H}_{B_2} \otimes \mathcal{H}_C) \\
& A^x \in \mathrm{B}(\mathcal{H}_A), B_b \in \mathrm{B}(\mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}), C^z \in \mathrm{B}(\mathcal{H}_C) \\
& (A^x)^\dagger = A^x, (C^z)^\dagger = C^z, (A^x)^2 = \mathbb{1}, (C^z)^2 = \mathbb{1}, \\
& B_b \geq 0, \quad \sum_{b=0}^{3} B_b = \mathbb{1}_{\mathcal{H}}.
\end{aligned} \tag{36}$$

We have to modify the NPA hierarchy to account for the novel constraints that appear in our problem. In particular, the fact that we have real rather than complex Hilbert spaces, and also the locality structure given by Postulate $(4_{\mathbb{R}})$ and the causal network of Figure 6.

In order to do this, we introduce two changes to the usual NPA hierarchy. First, we are going to consider, as in Theorems 4.1.2, 4.2.1 moment matrices for the behaviour $\mathrm{P}(a,c|x,z,b)$ conditioned on Bob obtaining outcome b. Given a feasible point of Problem (36), it must be

$$\mathrm{P}(a,c|x,z,b) = \mathrm{tr}((A_a^x \otimes C_c^z)\omega_b),$$

where

$$\omega_b := \mathrm{tr}_{B_1B_2}((\mathbb{1}_A \otimes B_b \otimes \mathbb{1}_C)\rho)$$

is an unnormalized state. Now, instead of the usual moment matrix for $\mathrm{P}(a,c|x,z,b)$ we introduced a sort of bipartite moment matrix, first studied in [MBL+13], as follows.

Consider two alphabets, one for Alice $\Sigma_A := \{A_1^x\}_{x\in\mathcal{X}}$ and one for Charlie $\Sigma_C := \{C_1^z\}_{z\in\mathcal{Z}}$. Since all measurement operators for Alice and Charlie are dichotomic, we only need to consider one of the outcomes. Let us for the moment focus on Alice, although we consider the following procedure for Charlie as well. We will in general consider reduced words under the equivalence relations given by

$$w_1 E E w_2 \sim w_1 E w_2 \quad \text{for all } E \in \Sigma_A, w_1, w_2 \in \Sigma_A^*.$$

Usually, if Alice had state ρ we would consider the moment matrix given by

$$\Gamma^n = \Big(\mathrm{tr}\Big(\alpha_j \rho \alpha_i^\dagger\Big)\Big)_{\alpha_i,\alpha_j\in\Sigma_A^{\leq n}},$$

as defined in Equation (33) for pure states. Equivalently, we consider the following map

$$\Omega_A^n(\rho) := \sum_{\alpha_i,\alpha_j \in \Sigma_A^{\leq n}} \mathrm{tr}\left(\alpha_j \rho \alpha_i^\dagger\right) |\alpha_i\rangle\langle\alpha_j| \,.$$

We already know that Ω_A^n is positive, since for every state we get a moment matrix, which is positive. But we can say a bit more.

Lemma 4.3.2. *The map $\Omega_A^n : \mathcal{T}(\mathcal{H}_A) \to \mathrm{B}(span(|\alpha_i\rangle_{\alpha_i \in \mathcal{A}}))$ is completely positive.*

Proof. The map Ω_A^n has a decomposition in Kraus operators $\Omega_A^n(\rho) = \sum_n K_n \rho K_n^\dagger$ with

$$K_n = \sum_{\alpha_j \in \Sigma_A^{\leq n_A}} |\alpha_j\rangle\langle n|\, \alpha_j,$$

where $\{|n\rangle\}_{n \in |\mathcal{H}_A|}$ is an orthonormal basis of $\mathcal{H}_A$. This proves that the map is completely positive. □

In particular, the map $\Omega_A^{n_A} \otimes \Omega_C^{n_C}$ is also positive. We define our moment matrix as

$$\Gamma^b := \Omega_A^{n_A} \otimes \Omega_C^{n_C}(\omega_b),$$

where ω_b is the shared state between Alice and Charlie that reproduces the conditional statistics $\mathrm{P}(a, c|x, z, b)$. This moment matrix depends now on two parameters (n_A, n_C) instead of one, but explicitely contains a bipartite structure which makes easy to impose the extra constraints that we need in the formulation of our problem. In particular, Γ_b are all positive by virtue of Lemma 4.3.2. Furthermore, notice that

$$\sum_b \Gamma^b = \Omega_A^{n_A} \otimes \Omega_C^{n_C} \left(\sum_b \omega_b\right) = \int_\Lambda \Omega_A^{n_A}(\mathrm{tr}_{B_1}(\rho_1^\lambda)) \otimes \Omega_C^{n_C}(\mathrm{tr}_{B_2}(\rho_2^\lambda)) d\,\mathrm{P}_\Lambda(\lambda).$$

is real-separable, and therefore

$$\left(\sum_b \Gamma^b\right)^{T_A} = \sum_b \Gamma^b.$$

Finally, we notice that some of the entries of this matrix are related to each other. Because of the ciclicity of the trace, whenever $\alpha_i^\dagger \alpha_j = \alpha_k^\dagger \alpha_l \equiv \alpha$, the coefficient of $|\alpha_i\rangle\langle\alpha_j|$ is the same as that of $|\alpha_k\rangle\langle\alpha_l|$ in $\Omega_A^n(\rho)$. Therefore, we may write in general

$$\Gamma^b = \sum_{\alpha \in \Sigma_A^{\leq n_A} \cdot \Sigma_A^{\leq n_A}, \gamma \in \Sigma_C^{\leq n_C} \cdot \Sigma_C^{\leq n_C}} d^b_{\alpha,\gamma} M^\alpha \otimes N^\gamma,$$

for some real coefficients $d^b_{\alpha,\gamma}$ and matrices $(M^\alpha)_{a,a'} = \delta_{\alpha,a'a^\dagger}$, $(N^\gamma)_{c,c'} = \delta_{\gamma,c'c^\dagger}$.

As before, we now forget that these moment matrices come from a feasible point of Problem (36) and we simply impose this extra constraint in the NPA hierarchy that we obtain with these moment matrices. This turns out to be enough to witness a separation between RQM and QM. The optimization problem that we solve is therefore formulated as

$$
\begin{aligned}
\sup_{d,P} \quad & \mathcal{J}(\mathrm{P}), \\
\text{s.t.} \quad & \Gamma_b \geq 0, \\
& \left(\sum_b \Gamma_b\right)^{T_A} = \sum_b \Gamma_b, \\
& d^b_{\mathbb{1},\mathbb{1}} = \mathrm{P}(b), d^b_{A^x_1,\mathbb{1}} = \mathrm{P}_{AB}(1,b|x), d^b_{\mathbb{1},C^z_1} = \mathrm{P}_{BC}(b,1|z), d^b_{A^x_1,C^z_1} = \mathrm{P}(1,b,1|x,z), \\
& \mathrm{P}(a,b,c|x,z) \geq 0, \sum_a \mathrm{P}(a,b,c|x,z) = \mathrm{P}_{BC}(b,c|z), \sum_c \mathrm{P}(a,b,c|x,z) = \mathrm{P}_{AB}(a,b|x), \\
& \sum_a \mathrm{P}_{AB}(a,b|x) = \sum_c \mathrm{P}_{BC}(b,c|z) = P(b), \sum_b \mathrm{P}(b) = 1,
\end{aligned}
$$

where the conditions on $\mathrm{P}(a,b,c|x,z)$ in the last two lines enforce that P corresponds to a non-signalling, normalized tripartite distribution. We are able to solve this SDP at level 2 for both Alice and Charlie, obtaining an upper bound for Problem 36 of 7.66, which is manifestly much smaller than the $6\sqrt{2} \approx 8.48$.

5. Discussion

In this chapter we have put the axioms of quantum theory to the test in the setting of general causal networks, a setting in which the notion of time it simplified to its bare bones, as we are only interested on possible causal influences between parties, and nothing more. We have seen that, in the standard Bell scenario and in the less frequently used PBR scenario, it doesn't matter whether we use real or complex Hilbert spaces. In these cases, all experimental results which are obtainable in QM are also obtainable in RQM. However, in the `SWAP` scenario there is quite a considerable gap between the predictions of both theories. It is so big, that it has already been detected experimentally [CWL+22; LMW+22; WJG+22], although in [LMW+22] we use a slightly different inequality in order to avoid having to do a Bell state measurement. It turns out that many functionals can witness this separation between QM and RQM. For example, a systematic -but not complete- numerical study was carried in [BB22]. We have also now found some simpler functionals which witness this gap. However, many questions still remain open. For example, what is the simplest scenario in which this gap appears? Having Charlie perform so many measurements imposes a big constraint on how far up the NPA hierarchy we can compute, so simpler scenarios are desirable to compute real values. Another important question is to determine the actual value of these Bell inequalities in RQM. In this Chapter we have presented upper bounds. We also have lower bounds determined from seesaw optimization methods. However, the two types of bounds do not match.

Future work will address some of these and other questions. In the meantime, we conclude that the correct field to consider for quantum mechanics is that of the complex numbers.

Chapter 3

Quantum projectiles

This Chapter is currently under peer review in npj Quantum Information as

- D. Trillo, T. P. Le, M. Navascués.
  "Quantum supremacy in mechanical tasks: projectiles, rockets and quantum backflow"
  arXiv:2209.00725,

which is Reference [TLN22], and to which all authors contributed equally.

The paper has been adapted for inclusion in this thesis, as follows. Section 2 introduces all the necessary mathematical formalism with no original work, and can be seen as an extension of Appendix A of [TLN22]. The rest of the paper has been reordered, and some Sections have additional explanations, but are otherwise mostly unchanged.

1. Introduction

While current research in quantum theory focuses on the exploitation of quantum effects in communication and computation scenarios, quantum systems are also known to be advantageous for some mechanical tasks. A paradigmatic example is the tunneling effect, by which a quantum particle can be detected in regions of space that are classically forbidden by energy considerations. Another example is quantum backflow, a phenomenon in which a free quantum particle with positive momentum can be observed to propagate backwards. This effect was first identified by Allcock in the context of the time-of-arrival problem [All69], and later isolated by Bracken and Melloy [BM94]. More recent examples of quantum supremacy in mechanical systems can be found in [Tsi06] and [ZALS22].

However, the advantages that quantum mechanical systems might offer for transportation, understood as the quick dispatch of massive particles through free space, are unexplored. Admittedly, there has been a consistent effort to investigate the properties of a hypothetical quantum time-of-arrival operator [MME07]. Perhaps due to its foundational character, though, this research program has not produced so far any concrete task where quantum mechanical systems have the upper hand.

In this Chapter, we prove the supremacy of quantum mechanical systems over their classical counterparts in a practical transportation task, which we call the *projectile scenario*. We consider a situation where a non-relativistic one-dimensional quantum particle (a projectile) of mass $M > 0$ is prepared in some bounded region of space B and left to propagate freely. After some time ΔT, we measure if the particle is in some distant target region R. For a fixed initial quantum state ρ with spatial support in B, we compare the probability of detection in R with that of a classical particle, initially prepared in B with the same momentum distribution as ρ.

We find that there exist what one might call *ultra-fast states* (*ultra-slow states*), whose probability of detection in R at time ΔT is strictly greater (smaller) than that of any classical particle. Measuring quantum supremacy in the ultrafast (ultraslow) regime through the difference between the quantum and the maximum classical (the minimum classical and the quantum) probabilities of arrival, we find that the maximum quantum advantage in either case does not depend on the distance between the preparation and target regions, but only on the parameter $\alpha := M|B|^2/\Delta T$. For finite values of α, the maximum quantum-classical gap can be computed up to precision δ by diagonalizing an $N \times N$ matrix, with $N = O(\log(1/\delta))$.

We prove that the maximum quantum advantage, achieved in the limit $\alpha \to \infty$, equals the Bracken-Melloy constant [BM94], which bounds the strength of quantum backflow, and which was numerically estimated to have the value $c_{bm} \approx 0.038452$ [PGKW05; EFV05]. This conjectured value was, however, not computed with any rigorous error bounds. In fact, until now there seemed to be no reason to believe that c_{bm} was smaller than 1. In this regard, we argue that $0.0315 \leq c_{bm} \leq 0.0725$, hence providing the first non-trivial upper bound on c_{bm}.

As we show, the appearance of c_{bm} is not a coincidence: through simple metaplectic transformations we connect the quantum projectile problem with a variety of scenarios related to quantum backflow, such as quantum backflow itself, diffraction in time [Mos52] and quantum backflow in the presence of a constant force [MB98]. All such effects are therefore manifestations of the same mathematical phenomenon, seen through different coordinate systems. In light of the recent interest in experimentally demonstrating quantum backflow [PTMM13; EZB20; BG21; MYDP21], we argue that projectile scenarios are more experimentally friendly and operationally interesting.

To arrive at a transportation task with a quantum advantage beyond the Bracken-Melloy constant, we consider a scenario in which several projectiles are sequentially released, namely, a quantum rocket. However, it turns out that c_{bm} also limits the advantage of a quantum rocket over a classical analog with the same lift-off zone, combustion chamber size and rocket and fuel momentum distributions.

Nevertheless, we show that a superior quantum advantage can actually be attained in a variant of the projectile scenario where the quantum projectile is compared with a classical particle having the same position and momentum distributions.

The Chapter is structured as follows: in Section 2 we introduce the phase-space formulation of quantum mechanics, along with all the mathematical background required for it. Then, in Section 3 we study the projectile scenario, proving that it is equivalent to quantum backflow and showing lower and upper bounds for c_{bm}. Finally, in Section

4, we study some generalized effects that can achieve a better quantum advantage than the projectiles.

2. Phase space QM

2.1. Formally solving the Schrödinger equation. Since we want to measure a certain operator on our system after some time has elapsed, we are now in the situation where we have to solve Schrödinger's equation to determine the new state of the system. Remember that Postulate (2) states that if $|\psi(t)\rangle$ is the state of the system at time t, then

$$i\partial_t |\psi(t)\rangle = H |\psi(t)\rangle$$

for some self-adjoint operator H. Formally, a solution is then given by

$$|\psi(t)\rangle = e^{-iHt} |\psi(0)\rangle .$$

We know that the operator e^{-iHt} can be defined via the functional calculus 2.3.4. It also turns out to have particularly nice properties.

Definition 2.1.1. A strongly continuous one-parameter unitary group on $\mathcal{H}$ is a set of unitary operators $\{U(t)\}_{t\in\mathbb{R}}$ such that, for all $s,t \in \mathbb{R}$ and $|\psi\rangle \in \mathcal{H}$,

$$U(0) = \mathbb{1}_{\mathcal{H}}, \quad U(s+t) = U(s)U(t), \quad \lim_{s\to t} \| U(t) |\psi\rangle - U(s) |\psi\rangle \| = 0.$$

In general, a map $U : X \to \mathrm{B}(\mathcal{H})$, where X is a topological space, is said to be strongly continuous if for all $|\psi\rangle \in \mathcal{H}$ the map $U_\psi : X \to \mathcal{H}$ defined as $U_\psi(x) := U(x) |\psi\rangle$ is continuous. This is an important property that is needed for all the uniqueness theorems of this Section.

Proposition 2.1.2. *Let H be a self-adjoint operator on $\mathcal{H}$, and $U(t) := e^{-itH}$. Then $U(t)$ is a strongly continuous one-parameter unitary group, and for all $|\psi\rangle \in Dom(H)$, we have that $|\psi(t)\rangle := U(t) |\psi(0)\rangle$ is also in the domain of H and solves the Schrödinger equation.*

Proof. [**Proposition 10.14 and Lemma 10.17**; Hal13] □

On the other hand, a famous theorem of Stone states that the converse also holds. A strongly continuous one-parameter unitary group $U(t)$ always has an *infinitesimal generator* H such that $U(t) = e^{-iHt}$.

Theorem 2.1.3 (Stone)**.** *Let $U(t)$ be a strongly continuous one-parameter unitary group on $\mathcal{H}$. Then the operator*

$$H |\psi\rangle := i \lim_{t\to 0} \frac{U(t) |\psi\rangle - |\psi\rangle}{t},$$

where the limit is in the norm of $\mathcal{H}$, is densely defined and self-adjoint. Furthermore, $U(t) = e^{-iHt}$ for all $t \in \mathbb{R}$.

Proof. [**Section 10.2**; Hal13] □

In particular, postulating the Schrödinger equation is then the same as postulating a strongly continuous unitary evolution.

In this Chapter, unlike in the previous one, we constantly deal with unbounded self-adjoint operators. This is so because we are considering a particle in a line, and therefore we have an operator As anticipated in Chapter 2, we will deal with a different formulation of QM than the one given in Chapter 2.

2.2. The position and momentum operators. The physical system we are dealing with in this Chapter is that of a non-relativistic particle of mass M on an infinite line. We assume that it has no other degrees of freedom, so that it is free, or at most seeing the action of a position-dependent potential $V(x)$. Classicaly, such a system is described by Hamilton's canonical equations of motion on its phase space, which is the symplectic space $(\mathbb{R}^2, \mathrm{J})$. In canonical coordinates (q, p), the symplectic form J is represented by the matrix

$$\mathrm{J} := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{37}$$

so that $J(\xi, \xi') := \xi^T \mathrm{J} \xi'$ for all $\xi, \xi' \in \mathbb{R}^2$, abusing the notation. Hamilton's equations, which are given by

$$\frac{d\xi_i}{dt} = \sum_j \mathrm{J}_{ij} \frac{\partial H}{\partial \xi_j}$$

for $\xi \in \mathbb{R}^2$, can thus be written as

$$\frac{dq}{dt} = \partial_p H, \quad \frac{dp}{dt} = -\partial_q H.$$

Here, $H \equiv H(q,p) = p^2/2m + V(q)$ is the Hamiltonian of the system.

We need to first figure out a description of this system in QM. Since we will only deal with a single system, we only need to look at Postulates (1)-(3). However, the Postulates like this don't mean much. We also need a rule that tells us what Hilbert space to use, how to assign an self-adjoint operator to an observable, and so on. In particular, we need an operator that corresponds to position. Since, the possible outcomes of a position measurement are all of $\mathbb{R}$, we now from the spectral theorem 2.3.3 that such an operator must be of the form

$$X := \int_{\mathbb{R}} x d\mu(x),$$

for some PVM μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with values on $\mathrm{B}(\mathcal{H})$. In particular, the existence of such a PVM implies that the Hilbert space on which it takes values must be infinite dimensional. But it is possible to say a lot more about the structure of the Hilbert space. In particular, another formulation of the spectral theorem 2.3.3 states that it has to be a direct integral of "generalized eigenspaces", a notion which we introduce now for completeness.

In order to define an object which we can in good conscience call $\int_\Lambda \mathcal{H}_\lambda d\mu(\lambda)$, we somehow need to able to talk about measurability of a map $\lambda \mapsto \mathcal{H}_\lambda$. If all the Hilbert spaces that we intend to integrate $\mathcal{H}_\lambda$ are separable, a simple way of doing this is to choose an orthonormal basis $\{e_j(\lambda)\}_{i=1}^\infty \subset \mathcal{H}_\lambda$ such that

$$\langle e_i(\lambda) | e_j(\lambda) \rangle_\lambda$$

is a measurable function of λ. Note that this is possible only if the map $\lambda \mapsto \dim \mathcal{H}_\lambda$ is measurable. Also, we need to extend the definition of orthonormal basis to include the zero vector, in order to account for different dimensionality of the Hilbert spaces. Having made such a choice, we can define measurability by appealing to the measurability of the coordinate functions in such a basis.

Definition 2.2.1. Let (Λ, Σ, μ) be a measure space. We say that a collection $\{\mathcal{H}_\lambda\}_{\lambda\in\Lambda}$ of separable Hilbert spaces is *measurable* if it comes equipped with maps $\{e_i\}_{i=1}^\infty$ such that for all i

$$e_i : \Lambda \to \bigcup \mathcal{H}_\lambda, \quad e_i(\lambda) \subset \mathcal{H}_\lambda, \quad \langle e_i(\lambda)|e_j(\lambda)\rangle_\lambda \text{ is a measurable map for all } i,j$$

and for all $\lambda \in \Lambda$, the set $\{e_i(\lambda) \mid e_i(\lambda) \neq 0\}_{i=1}^\infty$ is an orthonormal basis of $\mathcal{H}_\lambda$.

Given a map $s : \Lambda \to \bigcup_{\lambda\in\Lambda} \mathcal{H}_\lambda$ such that $s(\lambda) \in \mathcal{H}_\lambda$, we say that it is measurable if

$$\lambda \mapsto \langle e_i(\lambda)|s(\lambda)\rangle_\lambda$$

is measurable for all $i \in \mathbb{N}$.

Note that with these definitions of measurability, the maps that we would like to integrate are also measurable. For example

$$\big\langle s(\lambda)\big|s'(\lambda)\big\rangle_\lambda = \sum_{i=1}^\infty \overline{\langle e_i(\lambda)|s(\lambda)\rangle_\lambda}\,\big\langle e_i(\lambda)\big|s'(\lambda)\big\rangle_\lambda$$

is now measurable.

Definition 2.2.2 (Direct integrals)**.** Let $\mathcal{H}_\lambda$ be a measurable collection of separable Hilbert spaces, and (Λ, Σ, μ) a σ-finite measure space. We define the *direct integral* of $\mathcal{H}_\lambda$ with respect to μ as the space of equivalence classes of almost-everywhere-equal measurable sections $s : \Lambda \to \bigcup_{\lambda\in\Lambda} \mathcal{H}_\lambda$ such that

$$\|s\|^2 := \int_\Lambda \langle s(\lambda)|s(\lambda)\rangle_\lambda \, d\mu(\lambda) < \infty,$$

with the scalar product given by

$$\big\langle s\big|s'\big\rangle := \int_\Lambda \big\langle s(\lambda)\big|s'(\lambda)\big\rangle_\lambda \, d\mu(\lambda).$$

We denote such as space by

$$\int_\Lambda^\oplus \mathcal{H}_\lambda d\mu(\lambda).$$

Example 2.2.3. Consider the constant map $x \mapsto \mathbb{C}_x \equiv \mathbb{C}$ for all $x \in \mathbb{R}$, and choose the standard orthonormal basis of $\mathbb{C}$ (that is, the number 1) for all $x \in \mathbb{R}$.

$$L^2(\mathbb{R}) := \int_\mathbb{R}^\oplus \mathbb{C} dx,$$

where dx is the usual Lebesgue measure on $\mathbb{R}$. Indeed, under our definition, elements of the direct integral are equivalence classes of measurable maps $s : \mathbb{R} \to \mathbb{C}$ which are square-integrable.

Theorem 2.2.4 (Spectral theorem, second form)**.** *Let A be a self-adjoint operator acting on a Hilbert space $\mathcal{H}$. There exists a σ-finite measure μ on $\sigma(A)$, a direct integral of Hilbert spaces*

$$\mathcal{K} := \int_{\sigma(A)}^{\oplus} \mathcal{H}_\lambda d\mu(\lambda)$$

and a unitary map $U : \mathcal{K} \to \mathcal{H}$ satisfying

$$U(\mathrm{Dom}(A)) = \left\{ s \in \int_{\sigma(A)}^{\oplus} \mathcal{H}_\lambda d\mu(\lambda) \;\middle|\; \int_{\sigma(A)} \|\lambda s(\lambda)\|_\lambda^2 d\mu(\lambda) < \infty \right\}$$

and

$$[UAU^\dagger(s)](\lambda) = \lambda s(\lambda)$$

for all $s \in U(\mathrm{Dom}(A))$.

Proof. [**Theorem 10.9**; Hal13] □

In some sense, the spaces $\mathcal{H}_\lambda$ are a sort of generalized eigenspaces for the operator A. The spectral theorem 2.2.4 shows that the space on which the position operator acts must be isometric to

$$\int_{\mathbb{R}}^{\oplus} \mathcal{H}_x d\mu(x) \tag{38}$$

for some measurable collection of Hilbert spaces $\mathcal{H}_x$, and measure $d\mu(x)$. Not much else can be said from general principles, and in fact the Hilbert spaces $\mathcal{H}_x$ depend on the internal structure of the system that we are considering. For example, a particle with spin might have a two-dimensional Hilbert space as its generalized eigenspaces. In this Chapter, we consider the simplest Hilbert spaces given by Equation (38). That is, the (equivalence classes of) square-integrable complex-valued functions $L^2(\mathbb{R}, dx)$. These choices correspond to a particle with no internal structure whatsoever.

The Hilbert space $L^2(\mathbb{R}, dx)$ in which X acts as the multiplication operator, as in Theorem 2.2.4, is refered to as the *position representation*. Consider here the original version of the spectral theorem (Theorem 2.3.3). We may write

$$X = \int_{\mathbb{R}} x d\mu_X.$$

It is customary to write the PVM as $d\mu_X \equiv |x\rangle\langle x|\, dx$. This is quite intuitive, as if we denote an abstract element of $L^2(\mathbb{R})$ as $|\psi\rangle$ and then define $\langle x|\psi\rangle = \psi(x)$, we may write

$$\int_{\mathbb{R}} \psi(x)\, |x\rangle\, dx := |\psi\rangle\,,$$

as $\langle\varphi| \int_{\mathbb{R}} \psi(x)\, |x\rangle\, dx = \int_{\mathbb{R}} \psi(x)\, \langle\varphi|x\rangle\, dx = \int_{\mathbb{R}} \overline{\varphi(x)}\psi(x) dx = \langle\varphi|\psi\rangle$. Similarly, for any nice enough kernel $K(x, y)$, we may define the operator

$$\left(\int_{\mathbb{R}^2} K(x, y)\, |x\rangle\langle y|\, dxdy \right) |\psi\rangle = \int_{\mathbb{R}^2} K(x, y)\psi(y)\, |x\rangle\, dxdy$$

Any such operator, if it is of trace-class and has enough regularity properties (see Section 2.3), has trace

$$\begin{aligned}\sum_{n=1}^{\infty}\int_{\mathbb{R}^2} K(x,y)\overline{\psi_n(x)}\psi_n(y)dxdy &= \sum_{n=1}^{\infty}\int_{\mathbb{R}} dy\psi_n(y)\int_{\mathbb{R}} dxK(x,y)\overline{\psi_n(x)}\\ &= \int_{\mathbb{R}} dy\sum_{n=1}^{\infty}\psi_n(y)\int_{\mathbb{R}} dxK(x,y)\overline{\psi_n(x)}\\ &= \int_{\mathbb{R}} dy\,\langle y|\sum_{n=1}^{\infty}\langle\psi_n|K_y\rangle\,|\psi_n\rangle\\ &= \int_{\mathbb{R}} dy\,\langle y|K_y\rangle\\ &= \int_{\mathbb{R}} dyK(y,y).\end{aligned}$$

for any orthonormal basis $\{\psi_n\}_{n=1}^{\infty}$ of $L^2(\mathbb{R})$, and with $\langle x|K_y\rangle$ denoting the map $K(x,y)$ for a fixed y. Thus, for such an operator K it is legitimate to write

$$\mathrm{tr}(K) = \int_{\mathbb{R}} \langle x|\,K\,|x\rangle\,dx.$$

In this notation then, the PVM that we have is the map $\mu_X : \mathcal{B}(\mathbb{R}) \to \mathrm{B}(L^2(\mathbb{R}))$ which assigns to every measurable subset E of $\mathbb{R}$ the projection

$$\Pi_E\psi(x) := \psi(x)\chi_E(x) = \left(\int_{\mathbb{R}} \chi_E(x)\,|x\rangle\langle x|\,dx\right)|\psi\rangle\,,$$

where χ_E is the characteristic function of E.

On the other hand, Stone's theorem 2.1.3 provides a way to define a momentum operator. It is easy to check that the operators $U(t)\psi(x) := \psi(x-t)$ form a strongly continuous one-parameter unitary group on $L^2(\mathbb{R})$. Since momentum is the infinitesimal generator of translations, we define by analogy

$$P\psi(x) := i\lim_{t\to 0}\frac{\psi(x-t)-\psi(x)}{t} = -i\partial_x\psi(x).$$

Similarly, there is a decomposition for the momentum operator

$$P = \int_{\mathbb{R}} pd\mu_P = \int_{\mathbb{R}} p\,|p\rangle\langle p|\,dp.$$

and a Hilbert space $L^2(\mathbb{R}, dp)$ where P acts as the multiplication operator. Mutatis mutandi, we can repeat for P all the notational definitions we made for X. This is called the *momentum representation*, and the unitary operator guaranteed by Theorem 2.2.4 that relates the position and momentum representations is the Fourier transform.

2.3. The Fourier transform. All missing proofs of this section can be found in any standard reference for Fourier analysis.

Definition 2.3.1. Let $f(x) \in L^1(\mathbb{R})$. The Fourier transform of f is defined as

$$\mathcal{F}[f](p) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)e^{-ixp}dx.$$

The "inverse" Fourier transform of f is defined as

$$\mathcal{F}^{-1}[f](p) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{ixp} dx.$$

The Fourier transform and its "inverse", acting on the same function, have the same properties, and so we write them just for the Fourier transform. Note that the integral is well-defined precisely because f is integrable. Inverse is in quotation marks, because when these operators are inverses of each other depends on the function space we are working with, as it will become clear in the following. For example, we have the following

Lemma 2.3.2 (Riemann-Lebesgue)**.** *The Fourier transform is a linear map $\mathcal{F} : L^1(\mathbb{R}) \to \mathcal{C}_0(\mathbb{R}) \subset L^\infty(\mathbb{R})$. That is, it takes integrable functions to continuous functions vanishing at infinity.*

However, it can happen that the Fourier transform of an integrable function is not integrable.

Example 2.3.3. The characteristic function of $[-1, 1]$

$$\chi_{[-1,1]}(x) := \begin{cases} 1 & \text{if } x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

has a non-integrable Fourier transform. Indeed, let $g(p) := \mathcal{F}\chi_{[-1,1]}(p)$. Then

$$g(p) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-ixp} dx = \frac{1}{2\pi} \frac{e^{ip} - e^{-ip}}{ip} = \sqrt{\frac{2}{\pi}} \frac{\sin p}{p}.$$

The function $\mathrm{sinc}(x) := \sin x/x$ is not integrable, since on any interval $[(n-1)\pi, n\pi]$ with $n \in \mathbb{N}$ we have

$$|\mathrm{sinc}(x)| \geq \frac{\sin x}{n\pi}.$$

On the other hand $|\sin x|$ is $\pi-$periodic. Let us call $C := \int_0^\pi |\sin x|$. Then

$$\int_0^\infty |\mathrm{sinc}(x)| \geq \frac{C}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

Therefore, a priori the inverse Fourier transform is not defined on the image of the Fourier transform of $L^1(\mathbb{R})$. We can control the Fourier transform in this other way:

Theorem 2.3.4 (Plancherel)**.** *Let $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then $\mathcal{F}[f] \in L^2(\mathbb{R})$. Furthermore,*

$$\int_{\mathbb{R}} \overline{\mathcal{F}[f](p)} g(p) dp = \int_{\mathbb{R}} \overline{f(x)} \mathcal{F}^{-1}[g](x).$$

In particular, $\|\mathcal{F}[f]\|_2 = \|f\|_2$.

Since $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is dense in $L^2(\mathbb{R})$, Plancherel's theorem 2.3.4 allows us to extend Definition 2.3.1 to $L^2(\mathbb{R})$ by continuity (with respect to the L^2-norm). On $L^2(\mathbb{R})$, the Fourier transform is a unitary operator with inverse $\mathcal{F}^{-1}$.

By constraining even more the initial space we recover even better integrability properties of the Fourier transform. Recall the space of rapidly-decaying functions:

Definition 2.3.5 (Schwarz space)**.** We define the space of rapidly-decaying functions (also known as Schwarz space, and its elements as test functions) the set

$$\mathcal{S}(\mathbb{R}) := \left\{ \varphi \in \mathcal{C}^\infty(\mathbb{R}) \;\middle|\; \left\| x^\alpha \partial_x^\beta \varphi(x) \right\|_\infty \leq \infty \quad \forall \alpha, \beta \in \mathbb{N}_0 \right\}.$$

A similar definition holds of the domain space $\mathbb{R}^n$ by changing α, β to be multiindices in $\mathbb{N}_0^n$. The correct topology in this space is defined with respect to the family of norms

$$\|\varphi\|_k := \max_{\alpha+\beta \leq k} \left\| x^\alpha \partial_x^\beta \varphi(x) \right\|_\infty,$$

for $k \in \mathbb{N}$. The test functions with this topology form a complete metrizable topological vector space.

Note that, as sets, $\mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}) \subset L^2(\mathbb{R})$, as for all $\varphi \in \mathcal{S}(\mathbb{R})$ one has

$$\int_{\mathbb{R}} |\varphi(x)| dx = \int_{\mathbb{R}} \frac{1}{1+x^2}(1+x^2)|\varphi(x)| dx \leq \left\|(1+x^2)^{-1}\right\|_{L^1} \left\|(1+x^2)\varphi(x)\right\|_{L^\infty} < \infty.$$

Furthermore, since all smooth functions of compact support are in $\mathcal{S}(\mathbb{R})$, we have in particular that $\mathcal{S}(\mathbb{R})$ is dense in $L^1(\mathbb{R})$ (in the L^1-norm) and also dense in $L^2(\mathbb{R})$ (in the L^2-norm). Therefore, their Fourier transform is defined, and we also have

Proposition 2.3.6. *If $\varphi \in \mathcal{S}(\mathbb{R})$, then $\mathcal{F}[\varphi] \in \mathcal{S}(\mathbb{R})$. Furthermore, $\mathcal{F} : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ is an isomorphism of topological vector spaces, with inverse $\mathcal{F}^{-1}$.*

We are interested in the dual of this space:

Definition 2.3.7 (Tempered distributions)**.** A tempered distribution T is an element of the (topological) dual $\mathcal{S}'(\mathbb{R})$ of the Schwarz space $\mathcal{S}(\mathbb{R})$.

Each test function ϕ (in fact, every function in $L^1(\mathbb{R}) \cup L^\infty(\mathbb{R})$) defines a tempered distribution T_φ as

$$T_\varphi(\psi) := \int_{\mathbb{R}} \varphi(x)\psi(x) dx.$$

It turns out that the inclusion $\mathcal{S}(\mathbb{R}) \subset \mathcal{S}'(\mathbb{R})$ is dense in the weak topology. That is, for all $T \in \mathcal{S}'(\mathbb{R})$ there is a sequence $\{\varphi_n\}_{n=1}^\infty \subset \mathcal{S}(\mathbb{R})$ such that

$$T(\psi) = \lim_{n\to\infty} \int_{\mathbb{R}} \varphi_n(x)\psi(x) dx \quad \text{for all } \psi \in \mathcal{S}(\mathbb{R}).$$

This density and , we may extend the definition of the Fourier transform to $\mathcal{S}'(\mathbb{R})$, as $\mathcal{F}[T](\varphi) := T(\mathcal{F}^{-1}[\varphi])$. Indeed, this definition is consistent with the injection $\mathcal{S}(\mathbb{R}) \subset \mathcal{S}'(\mathbb{R})$ by Plancherel's theorem 2.3.4:

$$\mathcal{F}[T_\varphi](\psi) = \int_{\mathbb{R}} \varphi(x)\mathcal{F}^{-1}[\psi](x) dx = \int_{\mathbb{R}} \mathcal{F}[\varphi](p)\psi(p) dp = T_{\mathcal{F}[\varphi]}(\psi).$$

Finally, from Proposition 2.3.6, and Plancherel's theorem 2.3.4 we obtain that $\mathcal{F} : \mathcal{S}'(\mathbb{R}) \to \mathcal{S}'(\mathbb{R})$ is also an isomorphism.

Example 2.3.8. Consider the function $\operatorname{sgn}(x) \in L^\infty(\mathbb{R})$. Since it is bounded, it defines a tempered distribution $T_{\operatorname{sgn}(x)}$ but, as usual, we will abuse notation and simply refer to

the tempered distribution also as $\operatorname{sgn}(x)$. On the other hand, the function $1/x$ defines a tempered distribution by taking a Cauchy principal value on the integral. That is,

$$\text{P.V.}\left(\frac{1}{x}\right)(\varphi(x)) := \lim_{\varepsilon\to 0}\int_{\mathbb{R}\setminus[-\varepsilon,\epsilon]}\frac{1}{x}\varphi(x)dx$$

is a continuous functional on $\mathcal{S}(\mathbb{R})$.

These two distributions are related by a Fourier transform:

$$\mathcal{F}[\operatorname{sgn}](p) = \frac{1}{i}\frac{2}{\pi}\text{P.V.}\frac{1}{p}.$$

Other important properties of Fourier transforms are the following

Proposition 2.3.9. *Let $\overline{f(x)} = f(-x)$ for some $f \in L^2(\mathbb{R})$. Then, $\mathcal{F}[f](p) \in \mathbb{R}$ for all $p \in \mathbb{R}$.*

Theorem 2.3.10 (Convolution theorem)**.** *Let $\psi \in \mathcal{S}(\mathbb{R})$, $K \in \mathcal{S}'(\mathbb{R})$. Then, for almost all $p \in \mathbb{R}$,*

$$\mathcal{F}\left[\int_{\mathbb{R}} d\lambda\psi(\lambda)K(x-\lambda)\right](p) = \sqrt{2\pi}\mathcal{F}[\psi](p)\mathcal{F}[K](p).$$

Finally, note that integration by parts gives for all $\psi \in \mathcal{S}(\mathbb{R})$:

$$p\mathcal{F}[\psi](p) = \mathcal{F}[-i\partial_x\psi](p), \quad -i\partial_p\mathcal{F}[\psi] = \mathcal{F}[x\psi(x)]$$

These identities can also be extended to $\psi \in \mathcal{S}'(\mathbb{R})$ with a suitable definition of ∂_p. In particular, this shows that indeed the position and momentum representation are related via the Fourier transform, as stated in Section 2.2. From this, we can prove the following widely used identity relating the position and momentum representations:

$$d\mu_X(x)d\mu_P(p)\text{“} = |x\rangle\langle p|\,\langle x|p\rangle\,dxdp\text{”} = \frac{e^{ixp}}{\sqrt{2\pi}}\,|x\rangle\langle p|\,dxdp. \tag{39}$$

2.4. The Wigner function. As we have seen, the Fourier transform relates the position and momentum representations. It is also quite useful to do a partial Fourier transform and work in phase space, where both position and momentum have equal footing.

The Wigner function of a quantum state $\rho \in \mathcal{S}(L^2(\mathbb{R}))$ is usually defined as

$$\mathrm{W}_\rho(x,p) := \frac{1}{2\pi}\int_{-\infty}^{\infty} dy\left\langle x-\frac{y}{2}\right|\rho\left|x+\frac{y}{2}\right\rangle e^{ipy}, \tag{40}$$

where we are working in the position representation. In order to make sense of this integral, remember that a state $\rho \in \mathcal{S}(L^2(\mathbb{R}))$ can be represented as an infinite convex combination of pure states (cf. Equation (2))

$$\rho = \sum_{i=1}^{\infty}\lambda_i\,|\psi_i\rangle\langle\psi_i|\,, \quad |\psi_i\rangle \in L^2(\mathbb{R}).$$

Then, equation (40) becomes

$$\mathrm{W}_\rho(x,p) := \frac{1}{2\pi}\int_{-\infty}^{\infty} dy\sum_{i=1}^{\infty}\psi_i\left(x-\frac{y}{2}\right)\overline{\psi_i\left(x+\frac{y}{2}\right)}e^{ipy}$$

Let us now call

$$\mathrm{D}_\rho(x_1, x_2) := \sum_{i=1}^{\infty} \lambda_i \psi_i(x_1)\overline{\psi_i(x_2)},$$

where $\{\psi_i\}_{i=1}^{\infty}$ is an orthonormal basis of $L^2(\mathbb{R})$, and $\lambda_i \geq 0$, $\sum_{i=1}^{\infty} \lambda_i = 1$. Then we have

$$\mathrm{W}_\rho(x, p) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dy \mathrm{D}_\rho \left(x - \frac{y}{2}, x + \frac{y}{2}\right) e^{ipy}.$$

This is a partial (inverse) Fourier transform. That is, a Fourier transform on only one variable of a multivariable function. The properties of the Wigner function are intimately tied to the properties of the Fourier transform.

In order to get integrable Wigner functions we define by analogy the following set of states.

Definition 2.4.1 (Tempered state)**.** We say that $\rho \in \mathcal{S}(\mathcal{H})$ is *tempered* if the defining kernel $\mathrm{D}_\rho(x_1, x_2) := \langle x_1 | \rho | x_2 \rangle \in \mathcal{S}(\mathbb{R}^2)$ is a test function.

We want to focus our analysis on such states, since they have very nice regularity properties which are important in order to work with Wigner functions. Thankfully, we don't lose any generality because of the following

Lemma 2.4.2. *The tempered states are dense in $\mathcal{S}(\mathcal{H})$ with the operator topology. Furthermore, the span of pure tempered states is dense in $\mathcal{S}(\mathcal{H})$.*

Proof. Indeed, let $\varepsilon > 0$, and take any state

$$\rho = \sum_{i=1}^{\infty} \lambda_i |\psi_i\rangle\langle\psi_i|, \quad |\psi_i\rangle \in L^2(\mathbb{R}).$$

By definition there is an $N \in \mathbb{N}$ such that

$$\rho^N := \frac{\sum_{i=1}^{N} \lambda_i |\psi_i\rangle\langle\psi_i|}{\sum_{i=1}^{N} \lambda_i}$$

is $\varepsilon/2$-close to ρ in the operator topology. By taking a bigger N, we may also consider $\sum_{i=1}^{N} \lambda_i \geq 1 - \varepsilon > 0$. On the other hand, $\mathcal{S}(\mathbb{R})$ is dense in $L^2(\mathbb{R})$, so let $|\varphi_i\rangle \in \mathcal{S}(\mathbb{R})$ be $\varepsilon/4$-close to $|\psi_i\rangle$. Without loss of generality, we may consider $\langle\varphi_i|\varphi_i\rangle = 1$. Then, the state

$$\sigma := \frac{\sum_{i=1}^{N} \lambda_i |\varphi_i\rangle\langle\varphi_i|}{\sum_{i=1}^{N} \lambda_i}$$

is tempered, and for all $|\psi\rangle \in L^2(\mathbb{R}^2)$

$$\begin{aligned}\left\|(\sigma-\rho^N)\,|\psi\rangle\right\|_{L^2(\mathbb{R})} &= \frac{1}{\sum_{i=1}^N \lambda_i}\left\|\sum_{i=1}^N \lambda_i\left(\langle\varphi_i|\psi\rangle\,|\varphi_i\rangle - \langle\psi_i|\psi\rangle\,|\psi_i\rangle\right)\right\|_{L^2(\mathbb{R})}\\ &\leq \sum_{i=1}^N \frac{\lambda_i}{\sum_{i=1}^N \lambda_i}\big(\|\langle\varphi_i|\psi\rangle\,|\varphi_i\rangle - \langle\varphi_i|\psi\rangle\,|\psi_i\rangle\|_{L^2(\mathbb{R})}\\ &\qquad + \|\langle\varphi_i|\psi\rangle\,|\psi_i\rangle - \langle\psi_i|\psi\rangle\,|\psi_i\rangle\|_{L^2(\mathbb{R})}\big)\\ &\leq \sum_{i=1}^N \frac{\lambda_i}{\sum_{i=1}^N \lambda_i} 2\||\psi\rangle\|_{L^2(\mathbb{R})}\||\varphi_i\rangle - |\psi_i\rangle\|_{L^2(\mathbb{R})}\\ &\leq \frac{\varepsilon}{2}\||\psi\rangle\|_{L^2(\mathbb{R})}.\end{aligned}$$

Therefore, σ is $\varepsilon/2$-close to ρ^N, which is $\varepsilon/2$-close to ρ, so σ is ε-close to ρ.

The second part follows. □

In general then, we will be able to determine the properties of any linear bounded operator acting on $\mathcal{S}(\mathcal{H})$, by its action on pure states which are tempered. In particular, we can restrict our study of Wigner functions to tempered states, for which all the properties are easier to prove, although we continue to state all the theorems in generality. In particular we have

Proposition 2.4.3 (Properties of the Wigner function)**.** *For all $\rho \in \mathcal{S}(L^2(\mathbb{R}))$, the Wigner function W_ρ is well defined and satisfies the following:*

(1) $\mathrm{W}_\rho \in \mathcal{C}_0(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$.

(2) W_ρ *is in general not in* $L^1(\mathbb{R}^2)$*, but* $\mathrm{W}_\rho \in L^1(\mathbb{R}^2)$ *if* ρ *is tempered.*

(3) W_ρ *takes values in* $\mathbb{R}$.

(4) $\mathrm{tr}\{A\rho\} = 2\pi\int_{\mathbb{R}^2} \mathrm{W}_\rho(x,p)\mathrm{W}_A(x,p)dxdp$ *for* A *and* ρ *tempered.*

(5) If $\psi(x) = 0$ *for all* $x \in R$ *for some convex region* $R \in \mathbb{R}$*, then* $\mathrm{W}_{|\psi\rangle\langle\psi|}(x,p) = 0$ *for all* $x \in R$.

Proof. If ρ is a tempered state, this follows easily from the properties of the Fourier transform given in Section 2.3, and the definition of the Wigner function.

□

Finally, we end this section by computing an important class of Wigner functions that we need for our numerical investigations later.

Example 2.4.4 (The Wigner function of number states)**.** Consider the operators $|m\rangle\langle n|$, where $\{|n\rangle\}_{n=1}^\infty$ represents the number basis of $L^2(\mathbb{R})$; that is,

$$\langle x|n\rangle = \sqrt{\pi}H_n(x)e^{-\frac{x^2}{2}},$$

where H_n is the n-th Hermite polynomial. First note that, by linearity of the definition of the Wigner function, for any state $\rho = \sum_{m,n=1}^N \rho_{mn}\,|m\rangle\langle n|$, we have that

$$W_\rho(x,p) = \sum \rho_{mn} W_{|m\rangle\langle n|}(x,p), \quad \rho_{mn} = \langle m|\,\rho\,|n\rangle\,. \tag{41}$$

Now, take ρ to be a coherent state, i.e., $\rho = |\alpha\rangle\langle\alpha|$, with

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{k=0}^{\infty} \frac{\alpha^k}{\sqrt{k!}} |k\rangle \,, \tag{42}$$

for $|k\rangle$ an element of the number basis. It follows that

$$\rho_{mn} = e^{-|\alpha|^2} \frac{\alpha^m \bar{\alpha}^n}{\sqrt{m!n!}}. \tag{43}$$

On the other hand, the Wigner function of a coherent state is known to be [Bra21b]

$$W_\rho(x,p) = \frac{1}{\pi} e^{-r^2 - 2|\alpha|^2 + \sqrt{2}(\alpha(x-ip) + \bar{\alpha}(x+ip))}, \tag{44}$$

with $r := \sqrt{x^2+p^2}$. Cancelling the factor $e^{-|\alpha|^2}$ in both sides of Equation (41) and expanding the remaining exponential in Equation (44) as a power series in $\alpha, \bar{\alpha}$, we can compare the coefficients multiplying $\alpha^m \bar{\alpha}^n$ on both sides of the resulting equation, thus obtaining

$$W_{|m\rangle\langle n|}(x,p) = \frac{\sqrt{m!n!}}{\pi} e^{-r^2} \sum_{k=0}^{\min(m,n)} \frac{(-1)^k}{k!} \frac{(\sqrt{2}r)^{m+n-2k}}{(m-k)!(n-k)!} e^{i\theta(n-m)}, \tag{45}$$

where $\theta = \arg(x+ip)$.

Note that in [Tsi06] Tsirelson provides the complex conjugated formula for the same quantity. This mistake does not, however, invalidate the main result of [Tsi06], namely, the computation of the spectrum of a given linear operator. This is so because the spectra of a self-adjoint operator and its complex conjugate in a given basis coincide.

Next, we invoke Equation (45) to derive the matrix elements

$$\mathcal{O}_{nm}(\phi) := \langle n| \, \Theta(\cos(\phi)X + \sin(\phi)P) \, |m\rangle$$

and show that

$$\mathcal{O}_{nm}(\phi) = \frac{\sqrt{m!n!}}{\pi} \frac{e^{i\phi(n-m)}(i^{n-m} - i^{m-n})}{i(n-m)} \sum_{k=\max(m,n)}^{m+n} \frac{(-1)^{m+n-k} 2^{k-\frac{m+n}{2}-1} \Gamma\left(k - \frac{m+n}{2} + 1\right)}{(m+n-k)!(k-n)!(k-m)!}. \tag{46}$$

We will use this expression to lower bound the maximum quantum advantage in the standard and restricted projectile scenarios. To begin, using Proposition 2.6.2 we obtain

$$\mathcal{O}_{nm}(\phi) = \int \mathrm{d}x\mathrm{d}p W_{|m\rangle\langle n|}(x,p) \Theta(x\cos\phi + p\sin\phi). \tag{47}$$

We can evaluate the right-hand side of the above equation by changing to polar coordinates. The result is

$$\mathcal{O}_{nm}(\phi) = \frac{\sqrt{m!n!}}{\pi} w_{nm} \frac{e^{i\phi(n-m)}(i^{n-m} - i^{m-n})}{i(n-m)}, \tag{48}$$

with

$$w_{nm} := \sum_{k=0}^{\min(m,n)} \frac{(-1)^k 2^{\frac{m+n}{2}-k-1}\Gamma\left(\frac{m+n}{2}-k+1\right)}{k!(m-k)!(n-k)!}$$
$$= \sum_{k=\max(m,n)}^{m+n} \frac{(-1)^{m+n-k} 2^{k-\frac{m+n}{2}-1}\Gamma\left(k-\frac{m+n}{2}+1\right)}{(m+n-k)!(k-n)!(k-m)!}, \tag{49}$$

where, in the last step, we changed the sum variable $k \to m+n-k$ so that a comparison with Equation (1.5) in [Tsi06] can be made.

As it turns out, the final expression for w_{nm} can be written in terms of the generalized hypergeometric function ${}_pF_q$. Thanks to such an identity, it is possible to compute w_{nm} accurately for large values of m, n, as we do in Section 3.1

2.5. The Weyl operators. We have argued that we should be considering the Hilbert space $L^2(\mathbb{R}, dx)$, in which the position operator consists on multiplication by x, and the momentum operator is $-i\partial_x$. It is immediate to check that in this setting, these two operators satisfy the famous canonical commutation relations (CCR)

$$[X, P] = i\mathbb{1}_{L^2(\mathbb{R})}. \tag{50}$$

However, the astute reader will notice that expression (50) is troublesome. On the left hand side we have an operator which is not defined on all of $L^2(\mathbb{R})$, and so one has to carefully consider domain issues in order to use the CCR to prove anything. This makes the CCR ill-suited as a characterization of the operators X and P. It is better to instead consider the strongly continuous one-parameter unitary groups that they generate. That is, the operators e^{-iaX}, e^{-ibP}.

Proposition 2.5.1 (The exponentiated CCR)**.** *For the usual position and momentum operators X, P acting on $L^2(\mathbb{R})$, it holds that*

$$e^{-iaX}e^{-ibP} = e^{-iab}e^{-ibP}e^{-iaX}. \tag{51}$$

This is now a statement about bounded operators, so with it we never have to worry about domain issues. If X, P were bounded operators, then Equation (51) would follow from the usual Baker-Campbell-Hausdorff (BCH) formula

$$e^A e^B = e^{A+B}e^{\frac{1}{2}[A,B]}, \quad \text{for all } A, B \text{ such that } [A,[A,B]] = [B,[A,B]] = 0.$$

However, not all (unbounded) operators that satisfy the CCR satisfy the exponentiated CCR (51), as a result of domain technicalities on the CCR (50) [**Example 14.5**; Hal13].

Proof. Let $\psi \in L^2(\mathbb{R}, dx)$. Since we defined P as the infinitesimal generator of translations, we have that $(e^{-ibP}\psi)(x) = \psi(x-b)$. On the other hand, using the spectral theorem 2.3.3, we have that $(e^{-iaX}\psi)(x) = e^{-iax}\psi(x)$. Therefore, $(e^{-iaX}e^{-ibP})\psi(x) = e^{-iax}\psi(x-b)$. On the other hand, $(e^{-iab}e^{-ibP}e^{-iaX}\psi)(x) = e^{-iab}e^{-ia(x-b)}\psi(x-b) = e^{-iax}\psi(x-b)$. □

The power of using the exponentiated CCR instead of the CCR is manifest in the celebrated theorem by Stone and von Neumann

Theorem 2.5.2 (Stone-von Neumann)**.** *Let A, B be self-adjoint operators acting irreducibly on a hilbert space $\mathcal{H}$. If A, B satisfy the exponentiated CCR (51), then there exists a unitary $U : \mathcal{H} \to L^2(\mathbb{R})$ such that*

$$Ue^{-itA}U^\dagger = e^{-itX}, \quad Ue^{-itB}U^\dagger = e^{-itP}.$$

If A, B do not act irreducibly on $\mathcal{H}$, we can decompose $\mathcal{H}$ into a direct sum of Hilbert spaces on which they do.

Proof. [vNeu31] □

One would also like to have something like a BCH formula for the operators X and P. For this, we need to consider for $a, b \in \mathbb{R}$ the sum $aX + bP$. Again we immediately have to deal with domain issues, which can be worked around. However, it is easier to define the operator $aX + bP$ using Stone's theorem 2.1.3 on the appropiate strongly continuous one-parameter unitary group. If the BCH formula held, we would have

$$e^{-it(aX+bP)} = e^{i\frac{t^2ab}{2}}e^{-iatX}e^{-ibtP},$$

so in order to seamlessly define $aX + bP$ we can look at the right hand side.

Proposition 2.5.3. *Let $U_{a,b}(t)$ be operators on $L^2(\mathbb{R}, dx)$ defined as*

$$(U_{a,b}(t)\psi)(x) := (e^{i\frac{t^2ab}{2}}e^{-itaX}e^{-itbP}\psi)(x) = e^{i\frac{t^2ab}{2}}e^{-itax}\psi(x - tb).$$

for $\psi \in L^2(\mathbb{R}, dx)$. These operators form a strongly continuous one-parameter unitary group.

Proof. Proposition 13.5 Hall. □

We can now define $aX+bP$ as the infinitesimal generator of $U_{a,b}(t)$, and automatically we get that its domain is dense on $L^2(\mathbb{R})$ and it is self-adjoint. It turns out that $aX+bP$, as one would naïvely define it, does satisfy the BCH formula in some dense domain, and both definitions coincide [**Section 14.2**; Hal13]. Now that we have the generator, we can forget about the one-parameter group. The following operators play an important role in Weyl quantization.

Definition 2.5.4 (Weyl operators)**.** For every $a, b \in \mathbb{R}$, we define the following operator on $L^2(\mathbb{R})$:

$$\mathbb{W}(a, b) := e^{i(bX-aP)} = U_{b,-a}(-1) = e^{-\frac{1}{2}iab}e^{ibX}e^{-iaP}.$$

They satisfy the following relations

Proposition 2.5.5 (The Weyl relations)**.**

$$\mathbb{W}(a, b)\mathbb{W}(a', b') = e^{i\frac{ab'-a'b}{2}}\mathbb{W}(a + a', b + b')$$

Proof. Indeed, using the exponentiated CCR (51), we obtain

$$\begin{aligned}\mathbb{W}(a,b)\mathbb{W}(a',b') &= e^{-i\frac{ab+a'b'}{2}} e^{ibX} e^{-iaP} e^{ib'X} e^{-ia'P} \\ &= e^{-i\frac{ab+a'b'}{2}} e^{-iab'} e^{ibX} e^{ib'X} e^{-iaP} e^{-ia'P} \\ &= e^{-i\frac{ab+a'b'+2a'b}{2}} e^{i(b+b')X} e^{-i(a+a')P} \\ &= e^{-i\frac{a'b-ab'}{2}} \mathbb{W}(a+a',b+b').\end{aligned}$$

□

Note that the exponential contains the symplectic form $ab' - ba' = \mathrm{J}((a,b),(a',b'))$. This hints already that the Weyl operators play a crucial role in the phase space formulation of quantum mechanics. Another important property of the Weyl operators is that they are strongly continuous as a map $\mathbb{W} : \mathbb{R}^2 \to \mathrm{B}(L^2(\mathbb{R}))$. This, together with the Weyl relations, also determine the Weyl operators up to a unitary transformation, in a more general Stone-von Neumann theorem.

Theorem 2.5.6. *Let $U : \mathbb{R}^2 \to \mathrm{B}(\mathcal{H})$ be a strongly continuous map such that for all $(a,b) \in \mathbb{R}^2$, the operator $U(a,b)$ is a unitary which acts irreducibly on a Hilbert space $\mathcal{H}$. If $U(a,b)$ satisfies the Weyl relations; that is,*

$$U(a,b)U(a',b') = e^{i\frac{ab'-a'b}{2}} U(a+a',b+b'),$$

then there is a unitary $V : \mathcal{H} \to L^2(\mathbb{R}^2)$ such that for all $(a,b) \in \mathbb{R}^2$,

$$VU(a,b)V^\dagger = \mathbb{W}(a,b).$$

The unitary V is unique up to a phase.

Proof. This is actually an intermediate step that von Neumann uses in [vNeu31] to prove Theorem 2.5.2. □

2.6. Symmetries of phase space. As in every physical theory, the symmetries of our objects of study will play an important role. The group of symmetries of a symplectic space is called the symplectic group. Given a symplectic space (X,J), we have

$$Sp(X,J) := \{A \in \mathrm{GL}(X) \mid J(Ax,Ay) = J(x,y) \text{ for all } x,y \in X\}.$$

In our case of interest, the symplectic space is $(\mathbb{R}^2,\mathrm{J})$, and we consider always canonical coordinates (q,p), so that J is given by Equation (37). In this case, $\mathrm{Sp}(\mathbb{R}^2,J)$ is identified with a subgroup of $\mathrm{GL}_2(\mathbb{R})$, which we denote by $\mathrm{Sp}_2(\mathbb{R})$. We have

$$\mathrm{Sp}_2(\mathbb{R}) = \{A \in \mathrm{GL}_2(\mathbb{R}) \mid A^T J A = J\}.$$

Indeed, $\mathrm{J}(Ax,Ay) = (Ax)^T \mathrm{J} Ay = x^T A^T \mathrm{J} Ay = x^T \mathrm{J} y = \mathrm{J}(x,y)$ for all $x,y \in \mathbb{R}^2$ if and only if $A^T \mathrm{J} A = \mathrm{J}$. Note now that

$$A^T J A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & \det(A) \\ -\det(A) & 0 \end{pmatrix},$$

and therefore $\mathrm{Sp}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})$.

Using the Weyl operators, we can get a unitary representation of $\mathrm{Sp}_2(\mathbb{R})$ on $L^2(\mathbb{R})$. Indeed, consider $\sigma \in \mathrm{Sp}_2(\mathbb{R})$, and let us abuse the notation again to denote by $\sigma(x,p)$

Name	Group element σ	Metaplectic group element U_σ
Translation	$\tau_{a,b}$	$\mathbb{W}(a,b)$
Shear	$S_\nu = \begin{pmatrix} 1 & \nu \\ 0 & 1 \end{pmatrix}$	$e^{-i\nu\frac{P^2}{2}}$
Squeezing	$D_\mu = \begin{pmatrix} \mu & 0 \\ 0 & 1/\mu \end{pmatrix}$	$e^{-i\log\mu\frac{XP+PX}{2}}$
Rotation	$R_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$	$e^{-i\alpha\frac{X^2+P^2}{2}}$
Reflection	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\rho \longmapsto \overline{\rho}$

Table 1. Image of the metaplectic representation $\sigma \mapsto U_\sigma$ on the elements of the affine symplectic group. Since every element σ of $\mathrm{Sp}_2\mathbb{R} = \mathrm{SL}_2(\mathbb{R})$ can be decomposed as a product $\sigma = R_\alpha D_\mu S_\nu$ with $\alpha, \nu, \mu \in \mathbb{R}$, and $\mu > 0$, this is enough to recover all unitaries. As proven in Corollary 2.6.1.1, the Wigner function is covariant with all these transformations, including the reflection, albeit via an antiunitary map.

the action of σ on the vector with coordinates $(x,p)^T$. It follows from the fact that $\mathrm{J}(\sigma(x,p),\sigma(x',p')) = \mathrm{J}((x,p),(x',p'))$ that

$$\mathbb{W}(\sigma(x,p))\mathbb{W}(\sigma(x',p')) = e^{i\frac{xp'-x'p}{2}}\mathbb{W}(\sigma(x+x',p+p')). \tag{52}$$

It follows from von Neumann's uniqueness theorem 2.5.6 that there exists a unitary $U_\sigma \in \mathrm{B}(L^2(\mathbb{R}))$ such that

$$\mathbb{W}(\sigma(x,p)) = U_\sigma \mathbb{W}(x,p)U_\sigma^\dagger. \tag{53}$$

In particular, this correspondence is a projective representation, since the unitary U_σ is unique up to a phase, and

$$\mathbb{W}(\sigma_1(\sigma_2(x,p))) = U_{\sigma_1}\mathbb{W}(\sigma_2(x,p))U_{\sigma_1}^\dagger = U_{\sigma_1}U_{\sigma_2}\mathbb{W}(x,p)U_{\sigma_2}^\dagger U_{\sigma_1}^\dagger = U_{\sigma_1\sigma_2}\mathbb{W}(x,p)U_{\sigma_1\sigma_2}^\dagger.$$

The image of $\mathrm{Sp}_2(\mathbb{R})$ under the homomorphism $\sigma \mapsto U_\sigma$, together with the Weyl operators is called the metaplectic representation. It is a projective representation of the affine symplectic group; that is, $\mathrm{Sp}_2(\mathbb{R})$ together with all phase space translations $\tau_{a,b} : (x,p) \mapsto (x+b, p+a)$, operations which also preserve the symplectic form J. The unitary maps corresponding in the metaplectic representation to generators of the affine symplectic group are easily determined to be those in Table 1.

Proof of Table 1. First, we note that the translations $\tau_{a,b}$ indeed satisfy the relations

$$\tau_{a,b}\tau_{a',b'} = \tau_{a+a',b+b'}, \quad \sigma\tau_{a,b}\sigma^{-1} = \tau_{\sigma(a,b)},$$

where σ is a shear, squeezing or rotation. Therefore, since the Weyl operators satisfy relations (52, 53), we indeed have a projective representation of the affine symplectic group.

Now, we check that the U_σ for $\sigma \in \mathrm{Sp}_2(\mathbb{R})$ satisfy Equation (53). It turns out that any at-most-quadratic combination of X and P which is symmetric is a self-adjoint operator on a suitable domain which includes the tempered states. Since it is all we need, we assume that we are acting on a tempered pure state. Then, let us check for example Equation (53) for the shears. We work for simplicity in momentum representation:

$$\begin{aligned} e^{-i\nu\frac{P^2}{2}}\mathbb{W}(x,p)e^{i\nu\frac{P^2}{2}}\hat{\psi}(p') &= e^{i\frac{xp}{2}}e^{-i\nu\frac{P^2}{2}}e^{-ixP}e^{ipX}e^{i\nu\frac{p'^2}{2}}\hat{\psi}(p') \\ &= e^{i\frac{xp}{2}}e^{-i\nu\frac{p'^2}{2}}e^{-ixp'}e^{\nu\frac{(p'-p)^2}{2}}\hat{\psi}(p'-p) \\ &= e^{i\frac{xp+\nu p^2}{2}}e^{-ixp'}e^{-ipp'\nu}\hat{\psi}(p'-p) \\ &= e^{i\frac{(x+\nu p)p}{2}}e^{-i(x+\nu p)P}e^{ipX}\hat{\psi}(p') \\ &= \mathbb{W}(x+\nu p,p)\hat{\psi}(p'). \end{aligned}$$

Similar but lengthier computations show the same for the other generators of $\mathrm{Sp}_2(\mathbb{R})$. □

The action of the metaplectic group induces an action on the Wigner functions. The best way to see it is through the following alternative definition of the Wigner functions, introduced in [Gro76]

Theorem 2.6.1. *For all $\rho \in \mathcal{S}(L^2(\mathbb{R}))$ tempered, we have*

$$\mathrm{W}_\rho(x,p) = \frac{1}{\pi}\,\mathrm{tr}\Big\{\rho\mathbb{W}(x,p)\Pi\mathbb{W}(x,p)^\dagger\Big\},$$

where Π is the parity operator, defined as $\Pi\psi(x) := \psi(-x)$.

Proof. First note that

$$\begin{aligned} \mathbb{W}(x,p)\Pi\mathbb{W}(x,p)^\dagger\psi(z) &= e^{ipX}e^{-ixP}\Pi e^{ixP}e^{-ipX}\psi(z) \\ &= e^{ipX}e^{-ixP}\Pi(e^{i(-pz-xp)}\psi(z+x)) \\ &= e^{ipX}e^{-ixP}(e^{i(pz-xp)}\psi(-z+x)) \\ &= e^{2ip(z-x)}\psi(-z+2x) \end{aligned}$$

Let $\rho = |\psi\rangle\langle\psi|$ be tempered. Then the trace of $\rho\mathbb{W}(x,p)\Pi\mathbb{W}(x,p)^\dagger$ is well defined, and equal to

$$\begin{aligned} \mathrm{tr}\Big\{\rho\mathbb{W}(x,p)\Pi\mathbb{W}(x,p)^\dagger\Big\} &= \int_{\mathbb{R}} dz e^{2ip(z-x)}\psi(-z+2x)\overline{\psi(z)} \\ &= \frac{1}{2}\int_{\mathbb{R}} dy e^{ipy}\psi\left(x-\frac{y}{2}\right)\overline{\psi\left(x+\frac{y}{2}\right)} \\ &= \pi\mathrm{W}_\rho(x,p). \end{aligned}$$

The claim follows from linearity and continuity. □

As a corollary, we can see how the metaplectic group transforms the Wigner function:

Corollary 2.6.1.1. *Let σ be an element of the affine symplectic group, and U_σ its metaplectic representation. Then,*

$$\mathrm{W}_\rho(\sigma^{-1}(x,p)) = \mathrm{W}_{U_\sigma \rho U_\sigma^\dagger}(x,p).$$

Furthermore, as noted in Table 1, if σ is the reflection, then

$$\mathrm{W}_\rho(x,-p) = \mathrm{W}_{\overline{\rho}}(x,p).$$

Proof. First, suppose that σ is a shear, a squeezing or a rotation. Then, using Equation (53), one gets

$$\begin{aligned}\mathrm{W}_\rho(\sigma^{-1}(x,p)) &= \frac{1}{\pi}\,\mathrm{tr}\Big\{\rho U_\sigma^\dagger \mathbb{W}(x,p) U_\sigma \Pi U_\sigma^\dagger \mathbb{W}(x,p)^\dagger U_\sigma\Big\} \\ &= \frac{1}{\pi}\,\mathrm{tr}\Big\{\left(U_\sigma \rho U_\sigma^\dagger\right)\mathbb{W}(x,p)\left(U_\sigma \Pi U_\sigma^\dagger\right)\mathbb{W}(x,p)^\dagger\Big\}.\end{aligned}$$

On the other hand, $U_\sigma \Pi U_\sigma^\dagger = \Pi$. This can be determined by realizing that the operators $P^2, X^2+P^2, XP+PX$ all preserve parity. Indeed, we can decompose every $L^2(\mathbb{R})$ function as a sum of an even and an odd part

$$\psi(x) = \frac{\psi(x)+\psi(-x)}{2} + \frac{\psi(x)-\psi(-x)}{2} =: \psi_{\mathrm{even}}(x) + \psi_{\mathrm{odd}}(x),$$

and $\Pi\psi_{\mathrm{even}} = \psi_{\mathrm{even}}$, $\Pi\psi_{\mathrm{odd}} = -\psi_{\mathrm{odd}}$. We obtain a decomposition of $L^2(\mathbb{R}) = {L^2}_{\mathrm{even}}(\mathbb{R}) \oplus {L^2}_{\mathrm{odd}}(\mathbb{R})$. The aforementioned operators all respect this decomposition, as can be easily checked on tempered functions, and so they commute with the parity operator.

Now suppose that σ is a translation. Then, we have from the Weyl relations that

$$\begin{aligned}\mathrm{W}_\rho(\tau_{-a,-b}(x,p)) &= \mathrm{W}_\rho(x-a,p-b) \\ &= \frac{1}{\pi}\,\mathrm{tr}\Big\{\left(\mathbb{W}(a,b)\rho\mathbb{W}(a,b)^\dagger\right)\mathbb{W}(x,p)\Pi\mathbb{W}(x,p)^\dagger\Big\}\end{aligned}$$

Finally, for the reflection we have,

$$\begin{aligned}\mathrm{W}_\rho(x,-p) &= \overline{\mathrm{W}_\rho(x,-p)} \\ &= \frac{1}{2\pi}\overline{\int_{\mathbb{R}} dy e^{-ipy}\left\langle x-\frac{y}{2}\right|\rho\left|x+\frac{y}{2}\right\rangle} \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} dy e^{ipy}\left\langle x-\frac{y}{2}\right|\overline{\rho}\left|x+\frac{y}{2}\right\rangle \\ &= \mathrm{W}_{\overline{\rho}}(x,p).\end{aligned}$$

□

Consider now an affine-linear transformation of phase space σ, such that $\mathrm{J}(\sigma(x,p)) = \pm\mathrm{J}(x,p)$. It follows that σ can be decomposed as a product of the elements given by Table 1, and thus

$$\mathrm{W}_\rho(\sigma^{-1}(x,p)) = \mathrm{W}_{U_\sigma \rho U_\sigma^\dagger}(x,p), \tag{54}$$

where by abuse of notation we define $U_\sigma \rho U_\sigma^\dagger \equiv \overline{\rho}$.

We end this Section by proving one last property of Wigner functions which we will make continuous use of.

Proposition 2.6.2. *Let $a, b, c \in \mathbb{R}$, $f \in L^\infty(\mathbb{R})$ and ρ a tempered state.*

$$\mathrm{tr}(f(a\mathbb{1} + bX + cP)\rho) = \int_{\mathbb{R}^2} f(a + bx + cp)\mathrm{W}_\rho(x, p)dxdp.$$

Proof. Note first that $A := a\mathbb{1} + bX + cP$ is a self-adjoint operator with continuous spectrum $\mathbb{R}$. This follows, for example, by noticing that the spectrum of the unitary $e^{-i(a\mathbb{1}+bX+cP)} := e^{-ia}e^{-i(bX+cP)}$ is all the complex numbers of modulus 1. Indeed, $e^{iaX}e^{icP}$ has the same spectrum as e^{iaX}, as can be seen by conjugating with the unitary $e^{-icP/2}$. In particular, we have that for all ψ, the measure $\mu_\psi^{a\mathbb{1}+bX+cP}$ is absolutely continuous with respect to the Lebesgue measure. Therefore, using the Radon-Nykodym theorem, we obtain a measurable function $g(z)$ such that $d\mu_\psi^A = g(z)dz$. Furthermore, since $\mu_\psi^A(\mathbb{R}) = 1$, the map g is in $L^1(\mathbb{R})$.

On the other hand, f is the Fourier transform of some other tempered distribution, and we may write for all $\psi \in \mathcal{S}(\mathbb{R})$

$$\int_{\mathbb{R}} f(x)\psi(x)dx =: T_f(\psi) = \lim_{n\to\infty} \int_{\mathbb{R}} \mathcal{F}\psi_n(x)\psi(x)dx$$

for some sequence of test functions ψ_n. Furthermore, from the functional calculus 2.3.4, we have

$$\begin{aligned}
\mathrm{tr}(f(a\mathbb{1} + bX + cP)\rho) &= \int_{\mathbb{R}} f(z)d\mu_\psi^A \\
&= \int_{\mathbb{R}} f(z)g(z)dz \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} \mathcal{F}[\psi_n](z)g(z)dz \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} \mathcal{F}[\psi_n](z)d\mu_\psi^A \\
&= \lim_{n\to\infty} \mathrm{tr}(\mathcal{F}[\psi_n](a\mathbb{1} + bX + cP)\rho).
\end{aligned}$$

This allows us to extend the functional calculus of $a\mathbb{1}+bX+cP$ to tempered distributions in a weak sense. We consider now the test functions ψ_n. With the same arguments we can see that

$$\mathcal{F}[\psi_n](a\mathbb{1} + bX + cP) = \int_{\mathbb{R}} \psi_n(t)e^{-it(a\mathbb{1}+bX+cP)}dt.$$

Using the exponentiated CCR and the properties of Wigner functions 2.4.3 we arrive at

$$\mathrm{tr}(\mathcal{F}[\psi_n](aX + bX + cP)\rho) = \int_{\mathbb{R}^2} \psi_n(a + bx + cp)\mathrm{W}_\rho(x, p)dxdp,$$

from which the result follows. □

Conversely, we will see that for tempered states, for which the Wigner function is integrable, integrating the Wigner function over certain regions of phase-space corresponds to computing the expectation value of some self-adjoint operator, as

$$\int_{\mathbb{R}^2} \chi_E(x, p)\mathrm{W}_\rho(x, p)dp = \mathrm{tr}(A_E\rho). \tag{55}$$

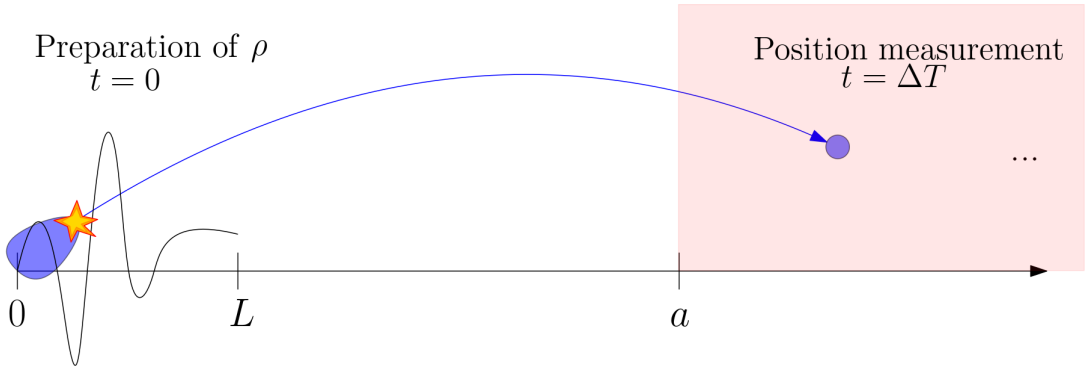

Figure 1. Projectile scenario. A projectile is prepared at time $t = 0$ in $[0, L]$ and, at time $t = \Delta T$, we verify that it has reached region $[a, \infty)$. Maximum quantum advantage in probability of arrival as compared to a classical particle is found to be the Bracken-Melloy constant, $0.0315 \leq c_{bm} \approx 0.038452 \leq 0.0725$.

3. Quantum Projectiles

We are now ready to introduce the task which is the topic of this Chapter. Our starting point is a classical projectile of mass M, prepared at time $t = 0$ in the region $[0, L]$. At time $t = \Delta T > 0$, we observe whether the projectile has reached region $[a, \infty)$, with $a > L$ as in Figure 1). If we ignore where exactly in $[0, L]$ the projectile was prepared, then the probability of finding it in $[a, \infty)$ at time ΔT is, at most, $\text{Prob}\,(p \geq M(a - L)/\Delta T)$, where p denotes the projectile's linear momentum. This corresponds to a configuration where the projectile was prepared at $x = L$ at time $t = 0$. Similarly, the probability to find the projectile in $[a, \infty)$ at time ΔT is, at least, $\text{Prob}\,(p \geq Ma/\Delta T)$, which corresponds to an initial preparation at $x = 0$.

Now, let us assume that the projectile is, in fact, a quantum mechanical system. Let $\mathcal{S}(R)$ denote the set of quantum states with spatial support in $R \subset \mathbb{R}$. We will omit the parentheses whenever R is an interval, and thus denote by $\rho \in \mathcal{S}[0, L]$ the initial quantum state of the projectile. Analogously, we also denote by $\mathcal{P}(R)$ the set of quantum states whose momentum has support in $R \subset \mathbb{R}$. While the projectile is freely propagating, its dynamics are governed by the kinetic Hamiltonian $H = P^2/2M$, where P denotes the projectile's linear momentum operator. The probability to find the quantum projectile in region $[a, \infty)$ after time ΔT can be found by simple application of the Born rule: it is $\text{tr}\big(U\rho U^\dagger \Theta(X - a)\big)$, where $U := e^{-iH\Delta T}$ and Θ is the Heaviside step function. Furthermore, we can write the classical probability of success as

$$\text{Prob}\left(p \geq \frac{M(a-L)}{\Delta T}\right) = \text{tr}\left\{\Theta\left(P - \frac{M(a-L)}{\Delta T}\right)\rho\right\}. \tag{56}$$

If, after time ΔT, the quantum projectile is found in $[a, \infty)$ with probability greater than the classical maximum, we say that the projectile is *ultra-fast*. If, on the contrary, the projectile is detected with probability lower than the classical minimum, we say that the projectile is *ultra-slow*. To gauge how ultra-fast or ultra-slow a quantum projectile in state ρ is, we consider the difference between the quantum and optimal classical probabilities of arrival.

Let us deal with the ultrafast case first. As we saw, the probability of success for this task is given by $\mathrm{tr}\{U\rho U^\dagger \Theta(X-a)\}$. Furthermore, the unitary U corresponds to a metaplectic representation of a shear, as proven in Table 1. Using Propositions 2.6.2, 2.4.3 and Equation (54), we find that

$$\begin{aligned}\mathrm{tr}\Big(U\rho U^\dagger\Theta(X-a)\Big) &= \int_{\mathbb{R}^2}\Theta(x-a)\mathrm{W}_{U\rho U^\dagger}(x,p)dxdp\\ &= \int_{\mathbb{R}^2}\Theta(x-a)\mathrm{W}_{\rho}(x-\Delta Tp/M,p)dxdp\\ &= \int_{\mathbb{R}^2}\Theta(x+\Delta Tp/M-a)\mathrm{W}_{\rho}(x,p)dxdp\\ &= \mathrm{tr}\{\Theta(X+\Delta TP/M-a)\rho\}.\end{aligned}$$

Thus the quantum advantage, if it exists, is given by $\mathrm{tr}(\rho\Omega_F(M,a,\Delta T))$, with

$$\Omega_F(M,a,\Delta T) := \Theta\left(X+\frac{\Delta T}{M}P-a\right)-\Theta\left(\frac{\Delta T}{M}P-a+L\right),$$

We wish to find the largest advantage achievable with a quantum projectile. That is, we are interested in the quantity

$$\varphi_F(M,L,a,\Delta T) := \sup_{\rho\in\mathcal{S}[0,L]}\mathrm{tr}(\rho\Omega_F(M,a,\Delta T)).$$

Now, given a set of states S and an operator A, we have, for any unitary U, that

$$\sup_{\rho\in S}\mathrm{tr}(\rho A) = \sup_{\rho\in USU^\dagger}\mathrm{tr}\Big(\rho UAU^\dagger\Big) = \sup_{\rho\in USU^\dagger}\mathrm{tr}\Big(U^\dagger\rho UA\Big).$$

We next exploit this observation to prove that φ_F is just a function of $\alpha := ML^2/\Delta T$. In particular, φ_F does not depend on a, the location of the target region: remarkably, quantum projectiles are equally advantageous no matter how large the flight distance.

Let $\sigma : \mathbb{R}^2 \to \mathbb{R}^2$ be an affine-linear transformation such that the linear part is symplectic. Equivalently, we have that σ acting formally on the pair of operators (X,P) formally satisfies $[\sigma(X,P)_1,\sigma(X,P)_2] = [X,P] = i$. Let us call such maps *metaplectic*. Then, as we have shown in Section 2.6, there exists a unitary U_σ such that

$$\mathrm{tr}\Big(U_\sigma^\dagger\rho U_\sigma f(a\mathbb{1}+bX+cP)\Big) = \mathrm{tr}(\rho f(a\mathbb{1}+b\sigma(X)+c\sigma(P))), \tag{57}$$

where we denote by $\sigma(X)\equiv\sigma(X,P)_1$ and $\sigma(P)\equiv\sigma(X,P)_2$.

Now, consider the unitary V associated to the metaplectic map

$$\begin{aligned}x &\longmapsto \sqrt{\frac{M}{\Delta T}}(x-L),\\ p &\longmapsto \sqrt{\frac{\Delta T}{M}}p-\sqrt{\frac{M}{\Delta T}}(a-L).\end{aligned} \tag{58}$$

For $\alpha = ML^2/\Delta T$, it follows that

$$\varphi_F(M,L,a,\Delta T) = \varphi(\alpha) := \sup_{\rho\in\mathcal{S}[-\sqrt{\alpha},0]}\mathrm{tr}(\rho\Omega), \tag{59}$$

where $\Omega := \Theta(X+P)-\Theta(P)$. Hence, φ_F is just a function of α. We call the right-hand side of the above equation the *standard problem*. Note that the standard problem

corresponds to determining the maximum quantum advantage of an ultrafast projectile of mass $M = 1$, prepared in the region $[-\sqrt{\alpha}, 0]$, to be found in region $[0, \infty)$ after time $\Delta T = 1$.

So far we have only considered ultrafast projectiles. For the ultraslow case, the story is pretty much the same: namely, the minimum probability that a classical projectile, prepared at time $t = 0$ in $[0, L]$ with the same momentum distribution as the quantum state ρ, reaches the target region at time $t = \Delta T$ is given by

$$\text{Prob}\left(p \geq \frac{Ma}{\Delta T}\right) = \text{tr}\left[\rho\Theta\left(\frac{\Delta T}{M}P - a\right)\right],$$

and so the quantum advantage, if it exists, corresponds to $\text{tr}(\rho\Omega_S(M, a, \Delta T))$, with

$$\Omega_S(M, a, \Delta T) := \Theta\left(\frac{\Delta T}{M}P - a\right) - \Theta\left(X + \frac{\Delta T}{M}P - a\right).$$

The maximum quantum advantage is thus

$$\varphi_S(M, L, a, \Delta T) := \sup_{\rho \in \mathcal{S}[0,L]} \text{tr}(\rho\Omega_S(M, a, \Delta T)).$$

As it turns out, $\varphi_S = \varphi$, and so the functions φ_F, φ_S are identical. Indeed, consider the transformation

$$\sigma(x, p) = \left(-\sqrt{\frac{M}{\Delta T}}x, \sqrt{\frac{\Delta T}{M}}p + \sqrt{\frac{M}{\Delta T}}(x - a)\right). \tag{60}$$

Since $[\sigma(X, P)_1, \sigma(X, P)_2] = -i$, this map does not define a unitary transformation over the set of quantum states. Rather, it defines an *anti-unitary* transformation U_σ, as explained in Section 2.6. Now, the argument above relating linear optimizations over subsets of quantum states also extends to anti-unitary transformations. One can verify that, applying U_σ to the standard problem with $\alpha = ML^2/\Delta T$, one ends up with the definition of φ_S, and, therefore, $\varphi_S(M, L, a, \Delta T) = \varphi\left(ML^2/\Delta T\right)$.

We finish this section by introducing yet another projectile scenario. As before, we wish the quantum projectile to have a larger probability of arrival, but this time we award some advantage to the classical projectile: namely, we compare the probability to detect the quantum projectile in the region $[a, \infty)$ with the maximum probability of detecting the classical one in the larger region $[a - b, \infty)$, with $b > 0$. This problem can be reduced, via the transformation (58), to an optimization of $\langle\Theta(X + P) - \Theta(P + \beta)\rangle_\rho$ over $\rho \in \mathcal{S}[-\sqrt{\alpha}, 0]$, with $\alpha = ML^2/\Delta T$, $\beta = b\sqrt{M/\Delta T}$. We denote this problem the *extended standard problem*, with solution $\varphi(\alpha, \beta)$. Clearly, $\varphi(\alpha, \beta)$ is non-increasing in β and $\varphi(\alpha, 0) = \varphi(\alpha)$. Obviously, $\lim_{\beta\to\infty}\varphi(\alpha, \beta) = 0$, and so one cannot reduce the extended standard problem to the standard problem.

3.1. Solving the standard problem: lower bounds. From the formulation of the standard problem (59), one can immediately deduce that φ is a non-decreasing function of $\alpha \in [0, \infty)$, with $\varphi(0) = 0$ and $\varphi(\alpha) \leq 1$. It remains to see that $\varphi(\alpha) \neq 0$ for some α. To do this, we need to study the spectrum of $\Omega := \Theta(X + P) - \Theta(P)$ restricted to the space $\mathcal{S}[-\sqrt{\alpha}, 0] \subset \mathcal{S}(\mathbb{R})$. In position representation, we have the following integral representation for Ω.

Proposition 3.1.1.

$$\Omega\big|_{\mathcal{S}[-\sqrt{\alpha},0]} = \frac{1}{2\pi}\int_{[-\sqrt{\alpha},0]^2} dxdy \frac{e^{\frac{i}{2}(y^2-x^2)}-1}{i(y-x)}\,|x\rangle\langle y| \tag{61}$$

Proof. We begin with the distributional identity

$$\operatorname{sgn} z = \frac{1}{i\pi}\text{P.V.}\int_{\mathbb{R}}\frac{dt}{t}e^{itz}$$

given by Example 2.3.8. The operator Ω is defined via the functional calculus as

$$\Theta(X+P)-\Theta(P) = \frac{\operatorname{sgn}(X+P)+\mathbb{1}}{2} - \frac{\operatorname{sgn}(P)+\mathbb{1}}{2} = \frac{1}{2}(\operatorname{sgn}(X+P)-\operatorname{sgn}(P)).$$

We are now going to apply $\operatorname{sgn}(X+P)$ and $\operatorname{sgn}(P)$ to a tempered pure state. In this case, we never encounter a problem exchanging integrals.

$$\begin{aligned}
\operatorname{sgn}(P) &= \int_{\mathbb{R}} \operatorname{sgn} p\,|p\rangle\langle p|\,dp \\
&= \frac{\sqrt{2\pi}}{i\pi}\text{P.V.}\int_{\mathbb{R}}\frac{dt}{t}\int_{\mathbb{R}} dp e^{itp}\,|p\rangle\langle p| \\
&= \frac{1}{i\pi}\frac{1}{\sqrt{2\pi}}\text{P.V.}\int_{\mathbb{R}}\frac{dt}{t}\int_{\mathbb{R}^2} dxdy\int_{\mathbb{R}} dp e^{itp}e^{ip(x-y)}\,|x\rangle\langle y| \\
&= \frac{1}{i\pi}\int_{\mathbb{R}^2} dxdy\text{P.V.}\int_{\mathbb{R}}\frac{dt}{t}\delta(t+x-y)\,|x\rangle\langle y| \\
&= \frac{1}{i\pi}\text{P.V.}\int_{\mathbb{R}^2} dxdy\frac{1}{y-x}\,|x\rangle\langle y|
\end{aligned}$$

Similarly,

$$\operatorname{sgn}(X+P) = \frac{1}{i\pi}\text{P.V.}\int_{\mathbb{R}^2}\frac{e^{\frac{i}{2}(y^2-x^2)}}{y-x}\,|x\rangle\langle y|\,.$$

□

Let $K(x,y)$ be the kernel of this integral operator. If $\alpha > 0$, then we can choose $z \in (-\sqrt{\alpha},0)$ such that $K(0,z) = K(z,0)^* \neq 0$. Since $K(0,0) = 0$, by the determinant criterion it follows that the 2×2 matrix $\{K(x,y)\}_{x,y=0,z}$ is not negative semidefinite. In particular, it has a positive eigenvalue λ, with eigenvector $(c_0,c_z)^T$. Now, consider the ket

$$|\psi_\varepsilon\rangle = \frac{1}{\sqrt{\varepsilon}}\int_{[-\sqrt{\alpha},0]} dx(c_0\chi_{[-\varepsilon,0]}(x) + c_z\chi_{[z-\varepsilon,z]}(x))\,|x\rangle\,,$$

where χ_C denotes the characteristic function of $C\subset\mathbb{R}$. For small enough ε, $|\psi_\epsilon\rangle\langle\psi_\epsilon| \in \mathcal{S}[-\sqrt{\alpha},0]$ and $\langle\psi_\varepsilon|\,\Omega\,|\psi_\varepsilon\rangle \approx \varepsilon\lambda > 0$. We conclude that $\varphi(\alpha) > 0$ for all $\alpha > 0$, so ultrafast and ultraslow states exist in all projectile scenarios.

The problem of computing $\varphi(\alpha)$ for different values of α is more convoluted. Note that the kernel $K(x,y)$ is analytic in x,y; hence, for $x,y\in[-\sqrt{\alpha},0]$, we can approximate it up to arbitrary precision by a polynomial on x and y of sufficiently high degree. When we replace $K(x,y)$ by its N^{th} order Taylor expansion, we arrive at a new operator Ω_N, which can be shown to be close in operator norm to Ω, restricted to the subspace of wave functions defined in $[-\sqrt{\alpha},0]$. In turn, Ω_N only has support on the finite-dimensional subspace spanned by vectors of the form $\int_{[-\sqrt{\alpha},0]} dxx^k\,|x\rangle$, where k runs from 0 to the

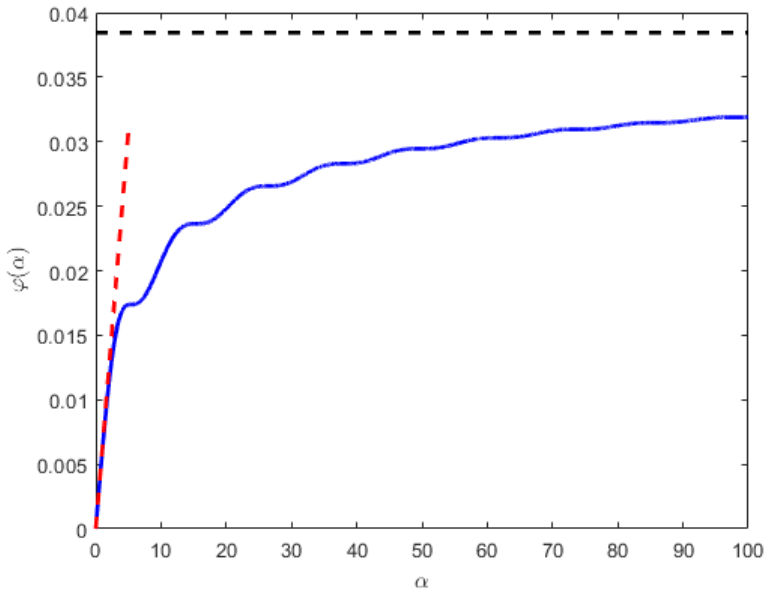

Figure 2. Solid blue: plot of $\varphi(\alpha)$ for $\alpha \in [0, 100]$, computed with precision $\delta = 10^{-4}$. Dashed red: linear upper bound $(2\sqrt{3}-3)\alpha/24\pi$. Dashed black: the conjectured value of the Bracken-Melloy constant c_{bm}

degree in x of the kernel of Ω_N. Hence Ω_N can be exactly diagonalized. The function $\varphi(\alpha)$ thus computed is plotted for $\alpha \in [0, 100]$ in Figure 2. As it can be appreciated, $\varphi(\alpha)$ roughly looks like a concave function, but not quite: at regular intervals, the slope of the function becomes very small. Such 'steps' seem to decrease in amplitude as α grows, and, actually, for $\alpha \gg 1$, the function appears to be well approximated by the Ansatz $r + s\alpha^{-1/2}$.

In addition, via variational methods, we show that $\varphi(\infty) \geq 0.0315$. Recall that $\varphi(\infty)$ is the result of maximizing $\mathrm{tr}(\Omega\rho)$ over all quantum states $\rho \in \mathcal{S}(-\infty, 0]$. Hence, any quantum state satisfying this constraint gives a lower bound on $\varphi(\infty)$. For any ρ, we can enforce this constraint by projection:

$$\hat{\rho} := \frac{\Theta(-X)\rho\Theta(-X)}{1-\epsilon} \in \mathcal{S}(-\infty, 0] \tag{62}$$

with $\epsilon = 1 - \mathrm{tr}(\Theta(-X)\rho)$. Using $\|\Omega\|_\infty \leq 1$, it is easy to prove that, for $\rho = |\psi\rangle\langle\psi|$,

$$\mathrm{tr}(\hat{\rho}\Omega) \geq \frac{\mathrm{tr}(\rho\Omega)}{1-\epsilon} - 2\sqrt{\frac{\epsilon}{1-\epsilon}} - \frac{\epsilon}{1-\epsilon}. \tag{63}$$

which provides a way to lower bound $\varphi(\infty)$, given an arbitrary quantum state not necessarily in $\mathcal{S}(-\infty, 0]$.

Consider, thus, a state ρ with support in $\mathcal{H}_N := \mathrm{span}\{|n\rangle\}_{n=0}^{N}$. The restrictions of the operators $\Omega = \Theta(X+P) - \Theta(P)$ and $\Theta(-X)$ to $\mathcal{H}_N$ can be computed through Equation (46). Taking $N = 1000$, we find, via matrix diagonalization, the pure state $|\psi\rangle \in \mathcal{H}_N$ maximizing the overlap

$$\langle\psi| \left(\Omega_N + \lambda\Theta(-X)_N\right) |\psi\rangle\,, \tag{64}$$

with $\lambda = 2500$. Defining $\rho^\star := |\psi\rangle\langle\psi|$, we compute the averages $\mathrm{tr}\left(\rho\hat{\Omega}\right)$, $\mathrm{tr}(\rho\Theta(-X))$ and, applying Equation (63), we find that $\varphi(\infty) \geq 0.0315$.A plot of the Wigner function of a quantum state approximately in $\mathcal{S}(-\infty, 0]$ and approximately achieving this value can be found in Figure 4 (left).

3.2. Solving the standard problem: upper bounds. To grasp the maximum quantum advantage, we need to study the limiting case $\alpha = \infty$. The problem thus consists in determining the spectrum of $\Omega := \Theta(X + P) - \Theta(P)$, restricted to the space $L^2(-\infty, 0]$. To study this case, it is convenient to switch to the Wigner function representation.

Recall now that, for any tempered state ρ, we have, by Proposition 2.6.2, that

$$\mathrm{tr}(\rho\Omega) = \int_{\mathbb{R}^2} dxdpW_\rho(x,p)(\Theta\left(x+p\right) - \Theta\left(p\right)).$$

The last factor on the integrand will vanish everywhere, except in the regions $\Lambda^+ = \{x + p \geq 0, p \leq 0\}$, where it equals 1, and $\Lambda^- = \{x + p \leq 0, p \geq 0\}$, where it equals -1. However, if $\rho \in \mathcal{S}(-\infty, 0]$, then $W_\rho(x, p) = 0$, for $x > 0$. Since $(x, p) \in \Lambda^+$ implies $x \geq 0$, it follows that the first region does not contribute to the integration above. Hence,

$$\varphi(\infty) = \sup_{\rho\in\mathcal{S}[-\infty,0]} -\int_{\Lambda^-} dxdpW_\rho(x,p).$$

The problem of integrating Wigner functions over wedges (without any further constraints) was studied by Werner [Wer88] in the context of time-of-arrival operators. The idea is that all wedges can be taken to each other via a metaplectic transformation, and therefore it suffices to study the wedge $[0, \infty) \times [0, \infty)$. Under this transformation, $\varphi(\infty)$ becomes

$$\sup_{\rho:\mathrm{tr}(\rho\Theta(X+P))=1} -\int_{[0,\infty)^2} dxdpW_\rho(x,p),$$

where we have used that $\mathcal{S}(-\infty, 0]$ is the space of states that satisfy the condition $\mathrm{tr}(\rho\Theta(-X)) = 1$. Werner considers the operator corresponding to integrating Wigner functions over the quadrant $x, p \geq 0$, and numerically determines its spectrum to be $[-0.155940, 1.007678]$. Therefore, $\varphi(\infty) \leq 0.155940$. This bound, however, does not take into consideration the constraint $\mathrm{tr}(\rho\Theta(X + P)) = 1$. To account for it, we add to Werner's operator a linear combination of operators corresponding to integrating Wigner functions over hyperbolic regions in the quadrant $x, p \leq 0$. Since our Wigner functions vanish in that quadrant, the infimum of the spectrum of the new operator (which can also be determined with the techniques in [Wer88]) also provides an upper bound for $\varphi(\infty)$. We numerically find the bound $\varphi(\infty) \leq 0.0725$.

The problem of upper bounding $\varphi(\infty)$ is equivalent to that of lower bounding the bottom of the spectrum of the operator A defined through $\mathrm{tr}(\rho A) = \int_{\mathbb{R}^2} \theta(x)\theta(p)W_\rho(x, p)$, constrained to the space $\mathcal{Q}$ of wave-functions $|\psi\rangle$ satisfying $\Theta(X + P)\,|\psi\rangle = |\psi\rangle$.

Restricted to this space, $A = A + B$, for any operator B that integrates a Wigner function on some region $R \subset \{(x, p) \in \mathbb{R}^2 : x + p \leq 0\}$. Therefore,

$$\sup_B \inf\{\lambda : \lambda \in \sigma(A + B)\} \leq \inf\{\lambda : \lambda \in \sigma(A|_{\mathcal{Q}})\}.$$

Unfortunately, computing integrals of Wigner functions on arbitrary regions of phase space is arbitrarily complicated, so we must restrict ourselves to tractable regions.

Define $R_k := \{(x,p) \in \mathbb{R}^2 : xp \geq k, x \leq 0, p \leq 0\}$. Such hyperbolic regions are invariant under the action of the squeezing group $e^{-it(XP+PX)}$, and it turns out that the operator B_k representing integration over R_k can be block-diagonalized in a basis $\{|\eta\rangle_+, |\eta\rangle_-\}_\eta$ of squeezing generalized eigenvectors, exactly like Werner does for B_0 in [Wer88]. Hyperbolic regions were also independently considered in full generality in [WB05], where the spectrum is also numerically computed. The result can only be expressed as follows in terms of integrals which do not have an analytical expression, as far as we are aware:

Theorem 3.2.1. *For all $k \geq 0$, we have*

$$B_k = \int_{-\infty}^{\infty} d\eta \sum_{\sigma_1\sigma_2=+,-} K^k_{\sigma_1\sigma_2}(\eta)\, |\eta\rangle_{\sigma_1} \langle\eta|_{\sigma_2}\,,$$

where

$$\begin{aligned}
K^k_{--}(\eta) &:= 0,\\
K^k_{+-}(\eta) &:= \frac{1}{2\pi i}\int_0^{\infty} dx e^{i\eta x}\frac{e^{-2kiCoth(x)}}{Cosh(x)},\\
K^k_{-+}(\eta) &:= \overline{K^k_{+-}(\eta)},\\
K^k_{++}(\eta) &:= \lim_{\varepsilon\to 0}\int_{-\infty}^{\infty} dx e^{i\eta x}\frac{e^{-2kiTanh(x)}}{\varepsilon Cosh(x) + 2iSinh(x)}
\end{aligned}$$

Proof. For $k \geq 0, \varepsilon > 0$

$$\begin{aligned}
I^{k,\varepsilon} &:= \frac{1}{2\pi}\int_{\mathbb{R}^2} dpdq\theta(-p)\theta(qp-k)\theta(-q)e^{-\varepsilon pq}\int_{\mathbb{R}} dp' e^{-ip'q}\overline{\psi\left(p+\frac{p'}{2}\right)}\psi\left(p-\frac{p'}{2}\right)\\
&= \frac{1}{2\pi}\int_{\mathbb{R}^2} dpdp'\theta(-p)\overline{\psi\left(p+\frac{p'}{2}\right)}\psi\left(p-\frac{p'}{2}\right)\left(\int_{\mathbb{R}} dq\theta(-q)\theta(qp-k)e^{(-ip'-\varepsilon p)q}\right).
\end{aligned}$$

The last integral is easily computed for all $p < 0$ to be

$$\int_{\mathbb{R}} dq\theta(-q)\theta(qp-k)e^{(-ip'-\varepsilon p)q} = \int_{-\infty}^{-\frac{k}{p}} e^{(-ip'-\varepsilon p)q} = \frac{\left[e^{(-ip'-\varepsilon p)q}\right]_{-\infty}^{-\frac{k}{p}}}{-ip'-\varepsilon p} = -\frac{e^{\frac{(ip'+\varepsilon p)k}{p}}}{ip'+\varepsilon p}.$$

Substituting this back into the original integral, and performing the change of variables

$$\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} := \begin{pmatrix} 1 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} p \\ p' \end{pmatrix}, \quad \begin{pmatrix} p \\ p' \end{pmatrix} := \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & -1 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$$

we obtain

$$\begin{aligned}
I^{k,\varepsilon} &= \frac{1}{2\pi}\int_{\mathbb{R}^2} dp_1dp_2\theta(-p_1-p_2)\overline{\psi(p_1)}\psi(p_2)\frac{-e^{\frac{2k\left(i(p_1-p_2)+\varepsilon\frac{p_1+p_2}{2}\right)}{p_1+p_2}}}{i(p_1-p_2)+\varepsilon\frac{p_1+p_2}{2}}\\
&= \frac{-e^{k\varepsilon}}{2\pi}\int_{\mathbb{R}^2} dp_1dp_2\overline{\psi(p_1)}\psi(p_2)\frac{\theta(-(p_1+p_2))e^{2ik\frac{p_1-p_2}{p_1+p_2}}}{\frac{1}{2}\varepsilon(p_1+p_2)+i(p_1-p_2)}
\end{aligned}$$

We are now going to divide this integral in four pieces

$$I^{k,\varepsilon} = \sum_{\sigma_1,\sigma_2=+,-} I^{k,\varepsilon}_{\sigma_1,\sigma_2}$$

corresponding to integrating over each of the four quadrants $(\operatorname{sgn} p_1, \operatorname{sgn} p_2) = (\sigma_1, \sigma_2)$. Then, for each integral $I^{k,\varepsilon}_{\sigma_1\sigma_2}$ we perform the change of variables $p_i = \sigma_i e^{2\lambda_i}$, so that

$$I^{k,\varepsilon}_{\sigma_1,\sigma_2} := \frac{-e^{k\varepsilon}}{2\pi}\int_{R^2} d\lambda_1 d\lambda_2 4e^{2\lambda_1+2\lambda_2}\overline{\psi(\sigma_1 e^{2\lambda_1})}\psi(\sigma_2 e^{2\lambda_2})\kappa^{k,\varepsilon}_{\sigma_1,\sigma_2}(\lambda_1,\lambda_2) \tag{65}$$

where

$$\begin{aligned}
\kappa^{k,\varepsilon}_{\sigma_1,\sigma_2}(\lambda_1,\lambda_2) &:= \frac{\theta(-\sigma_1 e^{2\lambda_1}-\sigma_2 e^{2\lambda_2})e^{2ik\frac{\sigma_1 e^{2\lambda_1}-\sigma_2 e^{2\lambda_2}}{\sigma_1 e^{2\lambda_1}+\sigma_2 e^{2\lambda_2}}}}{\frac{1}{2}\varepsilon(\sigma_1 e^{2\lambda_1}+\sigma_2 e^{2\lambda_2})+i(\sigma_1 e^{2\lambda_1}-\sigma_2 e^{2\lambda_2})} \\
&= \frac{e^{-\lambda_1}e^{-\lambda_2}\theta(e^{-\lambda_1}e^{-\lambda_2}(-\sigma_1 e^{2\lambda_1}-\sigma_2 e^{2\lambda_2}))e^{2ik\frac{e^{-\lambda_1}e^{-\lambda_2}(\sigma_1 e^{2\lambda_1}-\sigma_2 e^{2\lambda_2})}{e^{-\lambda_1}e^{-\lambda_2}(\sigma_1 e^{2\lambda_1}+\sigma_2 e^{2\lambda_2})}}}{e^{-\lambda_1}e^{-\lambda_2}(\frac{1}{2}\varepsilon(\sigma_1 e^{2\lambda_1}+\sigma_2 e^{2\lambda_2})+i(\sigma_1 e^{2\lambda_1}-\sigma_2 e^{2\lambda_2}))} \\
&= e^{-\lambda_1}e^{-\lambda_2}\frac{\theta(-\sigma_1 e^{\lambda_1-\lambda_2}-\sigma_2 e^{\lambda_2-\lambda_1})e^{2ik\frac{\sigma_1 e^{\lambda_1-\lambda_2}-\sigma_2 e^{\lambda_2-\lambda_1}}{\sigma_1 e^{\lambda_1-\lambda_2}+\sigma_2 e^{\lambda_2-\lambda_1}}}}{\frac{1}{2}\varepsilon(\sigma_1 e^{\lambda_1-\lambda_2}+\sigma_2 e^{\lambda_2-\lambda_1})+i(\sigma_1 e^{\lambda_1-\lambda_2}-\sigma_2 e^{\lambda_2-\lambda_1})} \\
&= \frac{e^{-\lambda_1}e^{-\lambda_2}}{2}\hat{K}^{k,\varepsilon}_{\sigma_1,\sigma_2}(\lambda_1-\lambda_2),
\end{aligned}$$

with

$$\hat{K}^{k,\varepsilon}_{\sigma_1,\sigma_2}(\lambda) := 2\frac{\theta(-\sigma_1 e^{\lambda}-\sigma_2 e^{-\lambda})e^{2ik\frac{\sigma_1 e^{\lambda}-\sigma_2 e^{-\lambda}}{\sigma_1 e^{\lambda}+\sigma_2 e^{-\lambda}}}}{\frac{1}{2}\varepsilon(\sigma_1 e^{\lambda}+\sigma_2 e^{-\lambda})+i(\sigma_1 e^{\lambda}-\sigma_2 e^{-\lambda})}.$$

Therefore,

$$\begin{aligned}
\hat{K}^{k,\varepsilon}_{++}(\lambda) &:= 0, \\
\hat{K}^{k,\varepsilon}_{+-}(\lambda) &:= \frac{2\theta(-\lambda)e^{2ik\mathrm{Coth}(\lambda)}}{\varepsilon\mathrm{Sinh}(\lambda)+2i\mathrm{Cosh}(\lambda)}, \\
\hat{K}^{k,\varepsilon}_{-+}(\lambda) &:= \overline{\hat{K}^{k,\varepsilon}_{+-}(-\lambda)}, \\
K^{k,\varepsilon}_{--}(\lambda) &:= -\frac{2e^{2ik\mathrm{Tanh}(\lambda)}}{\varepsilon\mathrm{Cosh}(\lambda)+2i\mathrm{Sinh}(\lambda)}.
\end{aligned}$$

Substituting back into Equation (65), we obtain

$$\begin{aligned}
I^{k,\varepsilon}_{\sigma_1,\sigma_2} &= \frac{-e^{\varepsilon k}}{2\pi}\int_{\mathbb{R}^2} d\lambda_1 d\lambda_2 \overline{\sqrt{2}e^{\lambda_1}\psi(\sigma_1 e^{2\lambda_1})}\sqrt{2}e^{\lambda_2}\psi(\sigma_2 e^{2\lambda_2})\hat{K}^{k,\varepsilon}_{\sigma_1,\sigma_2}(\lambda_1-\lambda_2) \\
&= \frac{-e^{\varepsilon k}}{2\pi}\int_{\mathbb{R}^2} d\lambda_1 d\lambda_2 \overline{\varphi_{\sigma_1}(\lambda_1)}\varphi_{\sigma_2}(\lambda_2)\hat{K}^{k,\varepsilon}_{\sigma_1,\sigma_2}(\lambda_1-\lambda_2),
\end{aligned}$$

where $\varphi_{\sigma_i}(\lambda) := \sqrt{2}e^{\lambda}\psi(\sigma_i e^{2\lambda})$. We now see that we have a convolution, so we can use the Convolution Theorem 2.3.10 to obtain

$$I^{k,\varepsilon}_{\sigma_1,\sigma_2} = -e^{\varepsilon k}\int_{\mathbb{R}} d\eta \overline{\varphi_{\sigma_1}(\eta)}\varphi_{\sigma_2}(\eta)K^{k,\varepsilon}_{\sigma_1\sigma_2}(\eta),$$

where

$$K^{k,\varepsilon}_{\sigma_1\sigma_2}(\eta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\eta\lambda}\hat{K}^{k,\varepsilon}_{\sigma_1\sigma_2}(\lambda).$$

Taking the limit $\varepsilon \to 0$, we obtain the desired result. □

As proven in [Wer88], the operator A is block-diagonalized by the same unitary transformation. Setting $B = \sum_k b_k B_k$, we thus have that the bottom of the spectrum of $\tilde{A} := A + B$ equals $\inf_\eta\{\lambda_{\min}(\tilde{A}(\eta))\}$, where $\lambda_{\min}(Z)$ denotes the minimum eigenvalue of the matrix Z and $\{\tilde{A}(\eta) : \eta \in \mathbb{R}\}$ is a one-parameter family of 2×2 matrices.

In this regard, the best combination we could find before the integrals defining the entries of $\tilde{A}(\eta)$ became too numerically unstable to be reliable is

$$\tilde{A} := A + 0.7673B_0 - 0.8767B_{0.1} + 0.09895B_{0.5},$$

whose spectrum as a function of η is shown in Figure 3.2.

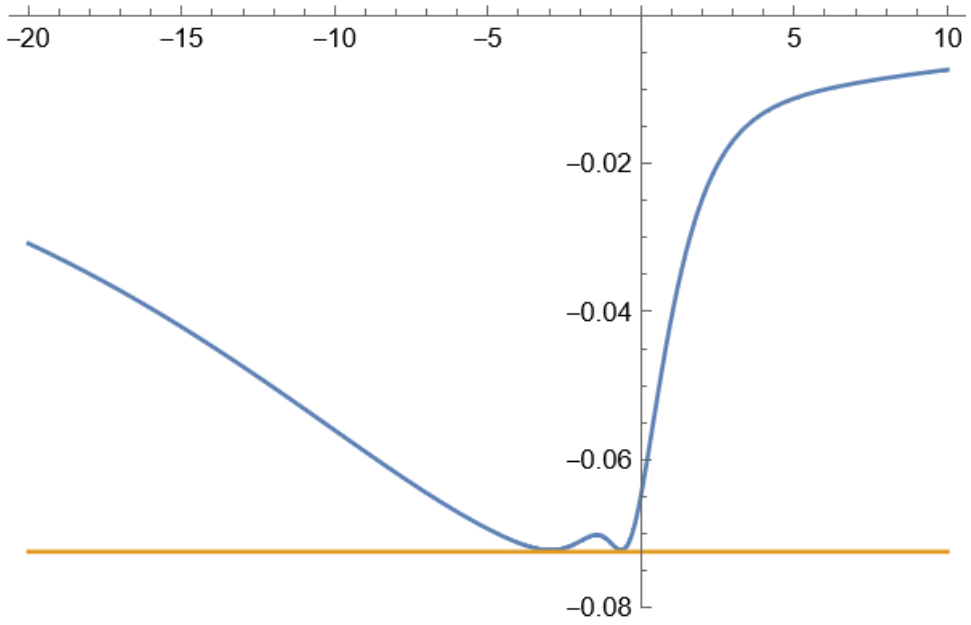

Figure 3. The bottom of the spectrum of the operator $\tilde{A}$, restricted to each two-dimensional subspace $\text{span}\{|\eta\rangle_+, |\eta\rangle_-\}$. The horizontal line is -0.0725. This spectrum was computed by numerically integrating with Mathematica, taking $\varepsilon = 0.001$ rather than a limit.

In the next section, we will show that $\varphi(\infty) = c_{bm}$, the Bracken-Melloy constant [BM94], which is conjectured to have the value 0.038452 [EFV05]. Our bounds $0.0315 \leq c_{bm} \leq 0.0725$ therefore support this widespread belief.

3.3. Quantum Backflow. As we have seen, the ultrafast (ultraslow) projectile problem is equivalent to the standard problem, since a unitary (anti-unitary) transformation takes us from the latter to the former. We next see that the standard projectile problem is similarly connected to the most extreme manifestation of other quantum mechanical effects. The exact correspondences are summarized in Table 2.

Let us start with the phenomenon of quantum backflow [All69; BM94; AGP16; EFV05]. Consider a pure state that only has positive momentum and that is evolving freely. In position representation, we can write it as

$$\psi(x,t) = \frac{1}{2\pi}\int_0^\infty dp e^{ipx} e^{-ip^2 t/2M}\phi(p).$$

Scenario	Operator	Set of states	$\sigma(x)$	$\sigma(p)$	α
Standard problem	$\Theta(P+X)-\Theta(P)$	$\mathcal{S}[-\sqrt{\alpha},0]$	x	p	α
Ultrafast projectile	$\Theta\left(X+\frac{\Delta T}{M}P-a\right)-\Theta\left(\frac{\Delta T}{M}P-(a-L)\right)$	$\mathcal{S}[0,L]$	$\sqrt{\frac{M}{\Delta T}}(x-L)$	$\sqrt{\frac{\Delta T}{M}}p-\sqrt{\frac{M}{\Delta T}}(a-L)$	$\frac{ML^2}{\Delta T}$
Ultraslow projectile	$\Theta\left(\frac{\Delta T}{M}P-a\right)-\Theta\left(X+\frac{\Delta T}{M}P-a\right)$	$\mathcal{S}[0,L]$	$-\sqrt{\frac{M}{\Delta T}}x$	$\sqrt{\frac{\Delta T}{M}}p+\sqrt{\frac{M}{\Delta T}}(x-a)$	$\frac{ML^2}{\Delta T}$
Quantum backflow	$\Theta\left(-X-\frac{\Delta T}{M}P\right)-\Theta(-X)$	$\mathcal{P}[0,\infty)$	$-\sqrt{\frac{\Delta T}{M}}p$	$-\sqrt{\frac{M}{\Delta T}}x$	∞

Table 2. Most of the optimization problems considered in this paper are of the form $\max_{\rho\in S}\operatorname{tr}(\rho\Omega)$, for some operator Ω and some set of states S. This table contains the definitions of each problem and the reversible transformations mapping the standard problem to any other. $\S(R)$ denotes the set of states with position support in $R\subset\mathbb{R}$, and $\mathrm{P}(R)$ denotes the set of states with momentum support in $R\subset\mathbb{R}$. We use the shorthand $\sigma(x):=\sigma(x,p)_1$ and $\sigma(p):=\sigma(x,p)_2$, and omit parentheses whenever R is an interval.

for some function ϕ such that $\int_0^\infty|\phi(p)|^2=1$. The probability flux at the origin is therefore

$$j(0,t)=\frac{1}{4M\pi}\int_0^\infty dpdq(p+q)e^{it(q^2-p^2)/2M}\phi(p)\overline{\phi(q)},$$

and thus the integrated flux at the origin from time 0 to time ΔT is

$$\int_0^{\Delta T}dtj(0,t)=\frac{1}{2\pi}\int_0^\infty dpdq\frac{e^{i\frac{\Delta T(q^2-p^2)}{2M}}-1}{i(q-p)}\phi(p)\overline{\phi(q)}.$$

Note the similarity with Equation (61). Guided by classical intuition, one would expect this integrated flux to be non-negative, since the particle is only moving to the right. However, for some quantum states $\phi(x,t)$, this magnitude can be negative: in that case, we speak of quantum backflow.

Alternatively, we can interpret quantum backflow as a *decrease* in the probability of detecting a particle particle with positive momentum in the region $[0,\infty)$. This is so because, by the continuity equation

$$\frac{\partial}{\partial t}|\psi(x)|^2=-\frac{\partial}{\partial x}j(x,t),$$

the integrated flux satisfies:

$$\int_0^{\Delta T}dtj(0,t)=\langle\psi|U^\dagger\Theta(X)U|\psi\rangle-\langle\psi|\Theta(X)|\psi\rangle,$$

where $|\psi\rangle=\int dx\psi(x,0)|x\rangle$ and $U=e^{-i\frac{P^2}{2M}\Delta T}$.

Scenario	Operator	Set of states	$\sigma(x)$	$\sigma(p)$	β
Extended standard problem, $\alpha = \infty$	$\Theta(P+X) - \Theta(P)$	$\mathcal{S}(-\infty, \beta]$	x	p	β
Generalized quantum backflow	$\Theta\left(-X - \frac{\Delta T}{M}P\right) - \Theta(-X)$	$\mathcal{P}[-\gamma, \infty)$	$-\sqrt{\frac{\Delta T}{M}}p$	$-\sqrt{\frac{\Delta T}{M}}x$	$\sqrt{\frac{\Delta T}{M}}\gamma$
Constant force quantum backflow	$\Theta\left(-X - \frac{\Delta T}{M}P + \frac{F\Delta T^2}{2M}\right) - \Theta(-X)$	$\mathcal{P}[0, \infty)$	$-\sqrt{\frac{\Delta T}{M}}\left(p - \frac{F\Delta T}{2}\right)$	$-\sqrt{\frac{M}{\Delta T}}x$	$\frac{F\Delta T}{2}$
Quantum reentry	$\Theta\left(l - X - \frac{t_2}{M}P\right) - \Theta\left(l - X - \frac{t_1}{M}P\right)$	$\mathcal{S}(-\infty, 0]$	$\sqrt{\frac{MC}{t_1}}(x-l)$	$\sqrt{\frac{M}{t_1 C}}\left(l - x - \frac{t_1}{M}p\right)$	l

Table 3. Some of the problems which are (anti-)metaplectically equivalent to the semi-infinite standard problem, with the same notation as in Table 2. In the last row, the normalization factor of the metaplectic transformation is $C := (t_2 - t_1)/t_2$.

Call $\mathcal{P}[0, \infty)$ the space of all states with positive momentum support. From all the above it follows that the maximum amount of backflow is given by

$$\sup_{\rho \in \mathrm{P}[0,\infty]} \operatorname{tr}\left(\rho\left(\Theta\left(-X - P\frac{\Delta T}{M}\right) - \Theta(-X)\right)\right) := c_{bm},$$

where we used the identity $\Theta(z) = 1 - \Theta(-z)$. The number c_{bm}, known in the literature as the *Bracken-Melloy constant* [BM94], is thus the solution a problem of the form $\sup_{\rho \in S} \operatorname{tr}(\rho A)$, for some space of states S and some operator A. In fact, this problem can be obtained from the standard problem with $\alpha = \infty$ via the anti-metaplectic transformation $\sigma(x, p) = (-p\sqrt{\Delta T/M}, -x\sqrt{M/\Delta T})$. Therefore, $c_{bm} = \varphi(\infty)$.

Going through the literature on quantum backflow, one finds that c_{bm} is conjectured to have the value 0.038452 [EFV05]. This figure is obtained by fitting many points of (an approximation to) the graph of $\varphi(\alpha)$ with the ansatz $r - s\alpha^{-1/2}$. To our knowledge, prior to our work there were no rigorous upper bounds on c_{bm}, and the best lower bound fell 41% short of the conjectured value of the constant [HGL+13]. Our results in the preceding section hence give mathematical support to the conjecture $c_{bm} \approx 0.038452$.

4. Some generalized effects

4.1. Generalized Quantum Backflow and related effects. In Table 3 we present another set of quantum effects that are mathematically equivalent, not to the standard problem, but to the extended standard problem with $\alpha = \infty$, which we express, via the transformation $\sigma(x, p) = (x - \beta, p + \beta)$, as an optimization of Ω over the set of states $\mathcal{S}(-\infty, \beta]$.

One of these effects is a variant of quantum backflow in which the particle evolves in the presence of a constant force [MB98]. That is, with the Hamiltonian given by $H = P^2/2M - FX$. In [Gou19] Goussev proves that this effect is at the same time equivalent to something he calls *quantum reentry*. Quantum reentry is a generalization of the famous diffraction in time of Moshinsky [Mos52], and consists in preparing a particle in $\mathcal{S}(-\infty, 0]$, letting it evolve and then measuring a negative probability flow in

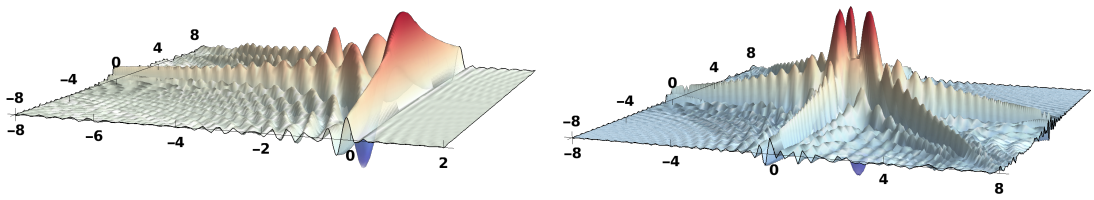

Figure 4. Wigner functions of (left) near-optimal state for the projectile scenario and (right) conjectured-optimal state for the constrained projectile scenario. Both states are obtained by truncating to the harmonic oscillator energy level $N = 170$. The left state is the eigenstate of $[\Theta(-X)]_{170}[(\Theta(X+P)-\Theta(P))]_{170}[\Theta(-X)]_{170}$ with eigenvalue 0.0331, where $[C]_N$ denotes the restriction of the operator C to the subspace spanned by the first $N+1$ number states. The right state is the eigenstate of $[\Theta(X+P)-\Theta(X)-\Theta(P)]_{170}$ with eigenvalue 0.1113.

some point $l \geq 0$. That is, the quantity under consideration is $-\int_{t_1}^{t_2} dtj(l,t)$ for some $t_2 > t_1 > 0$, which can again be easily transformed to the semi-infinite standard problem, as also shown in Table 3. In particular, the maximum probability transfer in both these effects is the same. This also answers the open question left in [Gou19] of physically understanding the equivalence between quantum reentry and quantum backflow in the presence of a constant force.

Finally, we note that the extended standard problem is equivalent to computing the maximum expression of quantum backflow when the initial momentum is in the region $[-\gamma, \infty)$ for some $\gamma \in \mathbb{R}$, as shown in Table 3. Thus, when the initial momentum is in this region, the probability "backflow" acts as if there were a constant force acting on the system, since these two problems are again equivalent. This seems to have gone unnoticed by Bracken, who studied the former effect in [Bra21a], despite having studied the latter in [MB98] together with Melloy.

4.2. Quantum Rockets. The low value of c_{bm} constitutes a severe obstruction to any practical application of quantum systems for transportation tasks. How to overcome this limit? A tempting idea is to consider scenarios where a transiting quantum projectile launches a second quantum projectile. Iterating this procedure, we arrive at the notion of a *quantum rocket*, i.e., a quantum mechanical system that, from time to time, throws away some fuel mass in the direction opposite to the intended motion. Since this rocket scenario encompasses the quantum projectile scenario, its maximum quantum advantage is lower-bounded by the Bracken-Melloy constant. Furthermore, one would imagine that, should we prepare the fuel in the right quantum state, the limited quantum advantage present in quantum projectiles could be somehow bootstrapped, hence increasing the overall advantage of the quantum rocket with respect to a classical rocket whose fuel combustion has an identical momentum distribution.

Unfortunately, this is not the case, at least for a large class of quantum rockets. Consider a minimal model for a quantum rocket, where, at time t, the rocket itself is regarded as a 1-dimensional particle of mass $M(t)$ and zero spin. The state of the rocket at time t is therefore specified through a trace-class positive semidefinite operator $\rho(t) : L^2(\mathbb{R}) \to L^2(\mathbb{R})$. For most of its flight, the rocket will be propagated by the kinetic

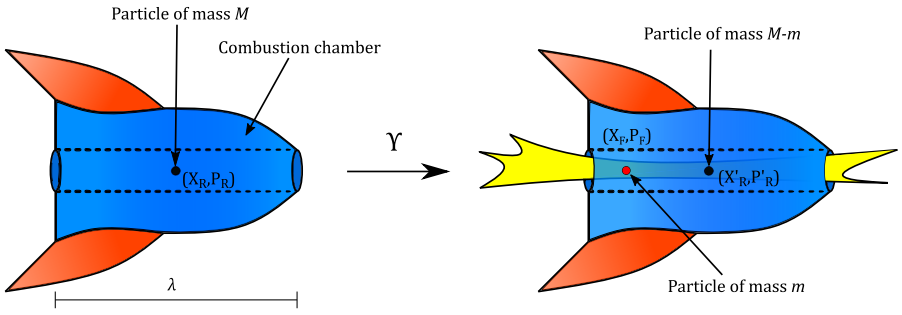

Figure 5. Action of the rocket-fuel splitting map Υ.

Hamiltonian $H = P_R^2/2M(t)$. At times $0 = t_1 < t_2 < ... < t_N$, though, the rocket's free evolution is interrupted: namely, at time t_j the rocket burns and releases a predetermined amount of fuel m_j instantaneously, thus decreasing its overall mass by the same amount.

To model the instantaneous combustion of fuel of mass $m < M$, we consider a CPTP map Υ that, acting on the rocket's state $\rho(t)$, returns a density matrix representing the joint state of the fuel F and that of the rest of the rocket R, whose mass is now $M - m$, see Figure 5.

Call X_F, P_F (X_R, P_R) the absolute position and momentum operators of the fuel (the rest of the rocket), and let X_{CM}, P_{CM} (X_{REL}, P_{REL}) denote the canonical variables of the center of mass (the relative coordinates between systems F and R), with:

$$\begin{aligned} X_{CM} &= \frac{M-m}{M} X_R + \frac{m}{M} X_F, \; P_{CM} = P_R + P_F, \\ X_{REL} &= X_F - X_R, P_{REL} = -\frac{m}{M} P_R + \frac{M-m}{M} P_F. \end{aligned} \tag{66}$$

Let $U_{M,m}$ be the (symplectic) unitary that switches between the R, F and CM, REL representations and define $\omega_{CM,REL} \equiv U_{M,m} \Upsilon(\rho) U_{M,m}^\dagger$. Since Υ is an internal and instantaneous operation, it cannot modify the rocket's center of mass degree of freedom. This means that $\mathrm{tr}_{REL}(\omega) = \rho$. For $\rho = |\psi\rangle\langle\psi|$, this last relation implies that $\omega = |\psi\rangle\langle\psi| \otimes \sigma_\psi$, for some quantum state σ_ψ.

However, σ_ψ must be independent of ψ. Otherwise, one could find two non-orthogonal vectors ψ, ψ' with the property that $\Upsilon(|\psi\rangle\langle\psi|), \Upsilon(|\psi'\rangle\langle\psi'|)$ are more easily distinguishable than $|\psi\rangle\langle\psi|, |\psi'\rangle\langle\psi'|$, which contradicts the contractivity of the trace norm under CPTP maps (Proposition 4.2.2). Putting all together, we find that any rocket-fuel splitting map Υ must be of the form

$$\Upsilon(\rho; \sigma, M, m) = U_{M,m}^\dagger (\rho \otimes \sigma) U_{M,m}, \tag{67}$$

where σ is the state of the relative system rocket-fuel. It must be noted that σ should have been prepared in the rocket's combustion chamber. If we assume that the combustion chamber is centered in the rocket's center of mass and has length λ, then σ must have spatial support in $[-\lambda/2, \lambda/2]$.

In describing the overall flight of the rocket, we assume that, at time t_j, the quantum rocket, with mass M_j, will release a mass m_j of fuel in state $\sigma_j \in \mathcal{S}[-\lambda/2, \lambda/2]$

(in the relative frame of reference). Hence, the mass and state of the rocket will be instantaneously updated to $M_{j+1} = M_j - m_j$, $\rho \to \mathrm{tr}_F(\Upsilon(\rho; \sigma_j, M_j, m_j))$.

We consider the probability to find the rocket at time $t_{N+1} > t_N$ in the region $[a, \infty)$. This is to be compared with the maximum probability that an analog classical rocket arrives at the same region in time t_{N+1}. Like in the projectile scenario, this classical rocket is assumed to have, at time t_1, the same initial mass, initial momentum distribution and initial spatial support as the quantum one. At time t_j, this classical rocket will burn a mass m_j of fuel, and the phase space distribution of the classical fuel in the fuel's reference frame relative to the rocket is demanded to have the same momentum distribution and spatial support as σ_j.

In these conditions, we now show that the difference between the quantum and classical arrival probabilities is also limited by c_{bm}. This no-go result crucially relies on Equation (67), which expresses the assumption that the fuel's interaction with the rocket is instantaneous. Physically, this corresponds to a configuration where the combustion chamber is open on both sides, i.e., the fuel is allowed to exit the rocket, not only against the rocket's direction of motion, but also towards it. Assumption (67) allows us to map the computation of the rocket's maximum quantum advantage to the standard problem (with further state constraints) through a metaplectic transformation.

Theorem 4.2.1. *The quantum advantage of rockets is also bounded by* c_{bm}

Proof. Under the assumption that the map (67) describes fuel combustion, consider a rocket that, most of the time, freely propagates through space, except at times $0 = t_1 < t_2 < t_3 < ... < t_N$, when the rocket burns fuel instantaneously. We assume that, initially, the state of the rocket's center of mass is $\rho \in \mathcal{S}([0, l])$, with canonical operators $X^{(0)}, P^{(0)}$. At time t_j, the rocket burns a fuel mass m_j, hence reducing its mass to $M_j = M - \sum_{k=1}^{j} m_j$, and experiencing a transformation $\rho \to \Upsilon(\rho; \sigma^{(j)}, M_{j-1}, m_j)$, where $\sigma^{(j)} \in \mathcal{S}\left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right]$ of the fuel in the rocket's reference frame, with canonical operators $X^{(j)}_{REL}, P^{(j)}_{REL}$. Between the times t_j and t_{j+1}, the rocket propagates freely and thus its canonical operators X_R, P_R experience the transformation

$$X_R \to X_R + \frac{t_{j+1} - t_j}{M_j} P_R, P_R \to P_R. \tag{68}$$

Call $X^{(j)}_R, P^{(j)}_R$ the canonical operators of the rocket at time t_j, just before the new fuel combustion. From Equations (66), (68) it is easy to see that they satisfy the relation

$$\begin{aligned} X^{(j)}_R &= X^{(j-1)}_R - \frac{m_j}{M_j} X^{(j)}_{REL} + \frac{1}{M_j} P^{(j-1)}_R - \frac{1}{M_j - m_j} P^{(j)}_{REL}, \\ P^{(j)}_R &= \frac{M_j - m_j}{M_j} P^{(j-1)}_R - P^{(j)}_{REL}. \end{aligned} \tag{69}$$

Through repeated iteration of Equation (69), we can express the rocket's final position operator $X^{(N)}_R$ as a linear combination of $X^{(0)}_R, P^{(0)}_R$ and $\{X^{(j)}_{REL}, P^{(j)}_{REL}\}$. That is, for some real vectors $\vec{c}, \vec{d}$, we have $X^{(N)}_R = \vec{c} \cdot \vec{X} + \vec{d} \cdot \vec{P}$, where $\vec{X} = (X^{(0)}_R, X^{(1)}_{REL}, ...)$ and $\vec{P} = (P^{(0)}_R, P^{(1)}_{REL}, ...)$. The probability of detecting the quantum rocket at time t_N in

$[a, \infty)$ and its classical counterpart is thus given by

$$\left\langle \Theta\left(\vec{c}\cdot\vec{X}+\vec{d}\cdot\vec{P}-a\right)\right\rangle_{\rho}, \tag{70}$$

where $\rho = \rho^{(0)} \otimes \bigotimes_{k=1}^{N} \sigma^{(k)}$. Since Equation (69) also holds for classical systems, so does Equation (70), when we understand ρ as a product of probability densities. We now consider a classical rocket with the same combustion schedule as the quantum one, and such that the probability densities for the classical moment variables $p_R^{(0)}, p_{REL}^{(1)}, p_{REL}^{(2)}, ...$ respectively coincide with those of the states $\rho^{(0)}, \sigma^{(1)}, \sigma^{(2)}, ...$ We further assume that the distributions of the initial position of the rocket and the fuel explosions respectively have supports $[0, l]$ and $\left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right]$, just like in the quantum case. Then, the maximum probability of detecting the classical rocket in $[a, \infty)$ at time t_N is

$$\left\langle \Theta\left(\vec{d}\cdot\vec{P}-(a-L^{+})\right)\right\rangle_{\rho},$$

where

$$L^{+} := l\max(0, c_0) + \frac{\lambda}{2}\sum_{k}|c_k|.$$

The maximum advantage φ_R of such a quantum rocket is thus the result of maximizing

$$\left\langle \Theta\left(\vec{c}\cdot\vec{X}+\vec{d}\cdot\vec{P}-a\right)-\Theta\left(\vec{d}\cdot\vec{P}-(a-L^{+})\right)\right\rangle_{\rho}, \tag{71}$$

over all separable states $\rho = \rho^{(0)} \otimes \bigotimes_{k=1}^{N} \sigma^{(k)}$ such that $\rho^{(0)} \in \mathcal{S}[0, l]$, $\sigma^{(j)} \in \mathcal{S}\left(\left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right]\right)$, for $j = 1, ..., N$. Call $\rho^{\star}$ the corresponding maximizer (if the maximizer does not exist, then the following argument still carries through if the average value of Equation (71) with $\rho = \rho^{\star}$ is $\phi_R - \epsilon$).

Now, consider the commutator $[\vec{c}\cdot\vec{X}, \vec{d}\cdot\vec{P}] = i\beta$, and assume that $\beta > 0$. Then, $X \equiv \vec{c}\cdot\vec{X} \to_S X$, $P \equiv \frac{1}{\beta}\vec{d}\cdot\vec{P}$, are canonically conjugated operators. Let $\tilde{\rho}$ be the result of tracing out all degrees of freedom of $\rho^{\star}$, but that corresponding to X, P. Then we have that

$$\varphi_R = \left\langle \Theta\left(X+\beta P-a\right)-\Theta\left(\beta P-(a-L^{+})\right)\right\rangle_{\tilde{\rho}},$$

with $\tilde{\rho} \in \mathcal{S}[L^{-}, L^{+}]$, with

$$L^{-} \equiv l\min(0, c_0) - \frac{\lambda}{2}\sum_{k}|c_k|.$$

Hence we end up computing φ under an extra restriction on the quantum states to be optimized. Through the metaplectic transformation $X \to X - L^{-}, P \to P$, we can map this problem to an optimization over the operator

$$\varphi_R = \left\langle \Theta\left(X+\beta P-a'\right)-\Theta\left(\beta P-(a'-L)\right)\right\rangle_{\tilde{\rho}},$$

over a constrained set of quantum states contained in $\mathcal{S}[0, L]$, with $L = L^{+} - L^{-}$, $a' = a - L^{-}$. This means that $\varphi_R \leq \varphi(\frac{L^2}{\beta}) \leq \varphi(\infty) \approx 0.038452$.

If $\beta < 0$, we apply the time-reversal anti-unitary operator $X_R^{(0)} \to X_R^{(0)}$, $X_{REL}^{(j)} \to X_{REL}^{(j)}$, $P_R^{(0)} \to -P_R^{(0)}$, $P_{REL}^{(j)} \to P_{REL}^{(j)}$ on the operator of Equation (71). This transformation does not affect the spatial support or separability of ρ, but effectively changes the sign of $\vec{d}$; and thus, of the commutator, in which case the argument above carries through.

Finally, if $\beta = 0$, then $\vec{c}\cdot\vec{X}, \vec{d}\cdot\vec{P}$ are commuting operators, in which case Equation (71) cannot have a value greater than 0.

The final conclusion is that a quantum rocket cannot be more advantageous than a quantum projectile. □

4.3. Another kind of projectile. In view of the last result, it would be reasonable not to expect significant gaps between the arrival probabilities of quantum and classical particles. As it turns out, though, a simple variation of the way we compare classical and quantum projectiles is enough to find quantum advantages for transportation way beyond the Bracken-Melloy constant.

We have compared the behavior of a quantum projectile (or a rocket) with respect to that of a classical one with the same momentum distribution and the same spatial support at time $t = 0$. Could the quantum advantage be amplified if we demanded further constraints on the initial position distribution $\mu(x)dx$ of the classical projectile, besides its support? In the extreme case, we could demand $\mu(x)dx$ to coincide with the position distribution of the quantum projectile.

Consider thus the following problem: let ρ denote the density matrix of a particle of mass M, and let $\mu(x)dx, \nu(p)dp$ be its position and momentum distributions at time $t = 0$. As before, we let the projectile evolve freely for time ΔT and then check whether the projectile is in $[a, \infty)$; call $p_q(\rho)$ the corresponding probability. How much does $p_q(\rho)$ differ from the maximum arrival probability of an analog classical particle, with initial position and momentum distributions $\mu(x)dx, \nu(p)dp$?

The maximum classical probability of arrival is

$$\begin{aligned} p_c^\star(\rho) = \sup & \int dxdpW(x,p)\Theta\left(x + p\frac{\Delta T}{M} - a\right) \\ \text{s.t.} \quad & \forall x, p, W(x,p) \geq 0, \\ & \int dpW(x,p) = \mu(x), \\ & \int dxW(x,p) = \nu(p), \end{aligned} \tag{72}$$

where $W(x,p)$ represents the probability distribution of the classical particle in phase space at time $t = 0$.

The maximum quantum-to-classical advantage in this projectile scenario is therefore $\Phi^\star = \sup_{\rho\in S} \mathcal{W}(\rho)$, where $\mathbb{W}(\rho) := p_q(\rho) - p_c^\star(\rho)$. This is a nested max-min optimization problem, whose solution can be proven independent of $a, M, \Delta T$ by using the metaplectic transformation $X \to \sqrt{\frac{\Delta T}{M}}X + a$, $P \to \sqrt{\frac{M}{\Delta T}}P$..

Suppose that there existed a linear operator Z such that

$$p_c^\star(\rho) \leq \mathrm{tr}(Z\rho), \tag{73}$$

for all states ρ. Then we could maximize the value

$$\mathbb{W}_Z(\rho) := \mathrm{tr}\left[\rho\left(\Theta\left(X + P\frac{\Delta T}{M} - a\right) - Z\right)\right] \tag{74}$$

over all density matrices with support on the first N number states. The result would provide us with a lower bound on $\Phi^\star$. In addition, if the maximizer $\rho^\star$ satisfied $\mathcal{W}_Z(\rho^\star) > 0$, then that state would be a good starting point for gradient ascent.

Now, how to identify an operator Z satisfying (73)? Let $f, g : \mathbb{R} \to \mathbb{R}$ be two functions such that

$$\Theta(x + p\Delta T/M - a) - f(x) - g(p) \leq 0, \tag{75}$$

for all x, p. Then, for any distribution $W(x,p)$ in phase space with marginals $\mu(x), \nu(p)$,

$$\begin{aligned} &\int dxdpW(x,p)\Theta(x + p\Delta T/M - a) \leq \\ &\int dxdpW(x,p)(f(x) + g(p)) = \\ &\int dx\mu(x)f(x) + \int dp\nu(p)g(p). \end{aligned} \tag{76}$$

It follows that the operator $Z = f(X) + g(P)$ fulfills condition (73). In fact, the dual of problem (72) is the maximum of the right-hand side of eq. (76) over all such functions f, g.

Take $M = \Delta T = 1, a = 0$. We observe that the functions $f = g = \Theta$ satisfy (75), and hence, the supremum of the spectrum of the operator $\Omega = \Theta(X+P) - \Theta(P) - \Theta(X)$ provides us with a lower bound for $\Phi^\star$, as $\mathrm{tr}(\rho\Omega) \geq \Phi^*$.

If we truncate this operator in the number basis, we are looking at the maximum eigenvalue of the matrix $(\mathcal{M}^{(N)}_{nm} : n, m = 0, ..., N)$, with

$$\mathcal{M}^{(N)}_{nm} = \langle n| \left(\Theta(X + P) - \Theta(X) - \Theta(P)\right) |m\rangle ,$$

For $N = 170$, the maximum eigenvalue of this matrix is 0.1113: the reader can find a plot of the Wigner function of the corresponding eigenvector in Figure 4 (right). Taking $N = 1700$, we obtain the tighter bound $\Phi^\star \geq 0.1228$. The maximum quantum advantage in this projectile scenario is therefore substantially greater than the conjectured value of c_{bm}, or even its upper bound 0.0725, derived in Section 3.2.

From all the above, it is thus natural to conjecture that the obtained value of 0.1228 is (close to) a local maximum of $\mathbb{W}$, at least among quantum states with support in $\{|n\rangle : n = 0, ..., 1700\}$.

On the other hand, note that after a suitable metaplectic transformation the problem $\sup_\rho \mathrm{tr}(\rho\Omega)$ becomes $\sup_\rho \mathrm{tr}\left(\rho\tilde{\Omega}\right)$, where

$$\tilde{\Omega} = \mathbb{1} - \sum_{k=0}^{2} \Theta(X_k) = -\frac{1}{2}\mathbb{1} - \frac{3}{2}\left(\frac{1}{3}\sum_{k=0}^{2} \mathrm{sgn}(X_k)\right)$$

with $X_k := \cos(2\pi k/3)X + \sin(2\pi k/3)P$. The operator $\sum_{k=0}^{2} \mathrm{sgn}(X_k)/3$ is the one studied by Tsirelson in [Tsi06]. The best known bounds for its spectrum are given in [ZALS22]. Using Equation (D20) in [ZALS22], one obtains that $\Phi^* \geq -0.5 + 1.5 \times \sqrt{0.17491} = 0.1262$. In particular, this shows how unreliable the numerical estimation of these quantities is, even after using a basis with 1700 number states, and thus the importance of getting good upper bounds as well as lower bounds.

5. Discussion

In this Chapter, we have investigated how the dynamics of quantum and classical projectiles differ, using the probability of arrival at a distant region of space as a figure of merit. We found that non-relativistic quantum particles can arrive at a distant region with higher or lower probability than any classical particle with the same initial spatial support and momentum distribution. Curiously enough, the maximum gap between quantum and classical probabilities is independent of the distance to the arrival region, and just depends on the mass M and spatial support L of the projectile and its flying time ΔT through the single parameter $\alpha = ML^2/\Delta T$. This is a new quantum effect taking place on mechanical systems, like quantum tunneling and quantum backflow.

The discrepancy between the quantum and classical arrival probabilities is, however, limited by the Bracken-Melloy constant $c_{bm} \approx 0.038452$. As we showed, the maximum quantum advantage of rockets with an open combustion chamber is also bounded by this value. Our no-go result does not apply, however, to rockets with a 1-side closed combustion chamber, which just allows the fuel to exit the rocket opposite to its direction of motion. Whether such rocket models are also limited by c_{bm}, or on the contrary, they can achieve arrival probabilities much higher than classical is an interesting topic for future research.

In a similar direction, we showed that considerable quantum-classical gaps of at least 0.1262 can be observed if we demand classical projectiles to reproduce the initial position distribution of the quantum projectile. It is an open problem whether this figure is indeed close to the maximum quantum advantage, and whether this effect can be exploited for real transportation tasks. This will be the topic of future work.

Chapter 4

Time translations

Sections 2 through 4 of this Chapter have been published as

- D. Trillo, B. Dive, M. Navascués.
  "Translating Uncontrolled Systems in Time"
  Quantum 4, 374 (2020),

which is Reference [TDN20]. It has been edited and reordered for inclusion in this Chapter. Unlike in the previous Chapters, the mathematical preliminaries (the tensor polynomials described in Sections 3.2 and 3.3) are an original contribution.

Section 5 of this Chapter has been published as

- D. Trillo, B. Dive, M. Navascués.
  "Universal quantum rewinding protocol with an arbitrarily high probability of success"
  Phys. Rev. Lett. 130, 110201 (2023),

which is Reference [TDN23]. It is mostly the same, with Section 5.1 being the main text of [TDN23] and Section 5.2 being the Appendix of [TDN23].

Since the papers are quite closely related, they have been merged into a single chapter. A single Section 1 has been included as an introduction to both papers, and a single Section 6 as a discussion to both papers.

1. Introduction

Transformations mapping a physical system to some other point on its free evolution curve (see Figure 1) are known as *time translations* [AAPV90]. In quantum theory, the effect of a time translation on a system evolving under a time-independent Hamiltonian H_0 can be described by a parameter $t \in \mathbb{R}$. Indeed, such a system evolves under the strongly continuous one-parameter unitary group $U(t) := e^{-iH_0 t}$, so that if the system starts in state $|\psi_0\rangle \in \mathrm{Dom}(H_0)$, the state at all times is described by

$$|\psi(t)\rangle = U(t)\,|\psi_0\rangle\,.$$

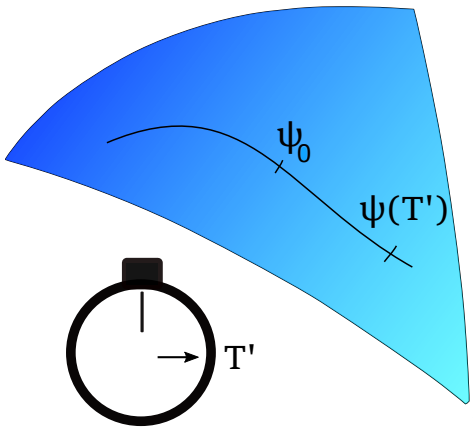

Figure 1. The normal time evolution of a system. On a Hilbert space (background blue), an initial state $|\psi_0\rangle$ travels along the curve $\{|\psi(t)\rangle \equiv U_t |\psi_0\rangle\}_{t\in\mathbb{R}}$, where U_t is a one-parameter strongly continuous unitary group with infinitesimal generator H_0, the Hamiltonian of the system. The parameter $t \in \mathbb{R}$ is the *evolution time*.

The goal of this Chapter is to study time translation protocols in the most general setting possible. In particular, we are interested in protocols that work for every initial state ψ_0, and every Hamiltonian H_0. It will turn out that the dimension of our system d is the most important parameter that we have to consider. In particular, we take d to be finite always. A nice consequence of this is that, in this Chapter, we will never encounter any domain issues.

Definition 1.0.1 (Time translations)**.** A time translation protocol $\mathcal{P} \equiv \mathcal{P}(d, T, T')$ is a completely positive (CP) map $\mathcal{P}$ such that given any d-dimensional system ρ_0 evolving under a Hamiltonian H_0, and interacting with other systems via an interaction Hamiltonian H_I,

$$\mathcal{P}(\rho_0) = pe^{-iH_0T}\rho_0 e^{iH_0T} \otimes |\text{success}\rangle\langle\text{success}| \,,$$

where $0 < p \equiv p(H_0, H_I) \leq 1$ is the probability that $\mathcal{P}$ has succeeded, and the CP map $\mathcal{P}$ is implemented in such a way that it takes a time T' for the output to appear.

The last part of the definition is crucial. Our goal is to use such a protocol to get the system to parts of its evolution curve which are not accessible just by waiting a certain time T', so we need to consider how long it takes for the protocol to be implemented. Since our protocols are probabilistic by nature, we will consider them to be heralded.

For a general system, we consider four kinds of time translations, pictured in Figure 2. It is important to note that finite dimensional systems usually exhibit a periodic time evolution, such as that of a rotating spin. In this case, evolution time is cyclic, and so there really is no distintion between rewinding or fast-forwarding. However, since we require our protocols to work for all Hamiltonians, we can always assume that we are in a regime where this doesn't happen.

There is much literature about time translations. The original paper [AAPV90] considers a time-translation machine based on the concept of superoscillations. In this case, the final state is only approximately in the evolution curve. Other techniques to implement specific time-translations include refocusing [SCHL16], Hamiltonian amplification

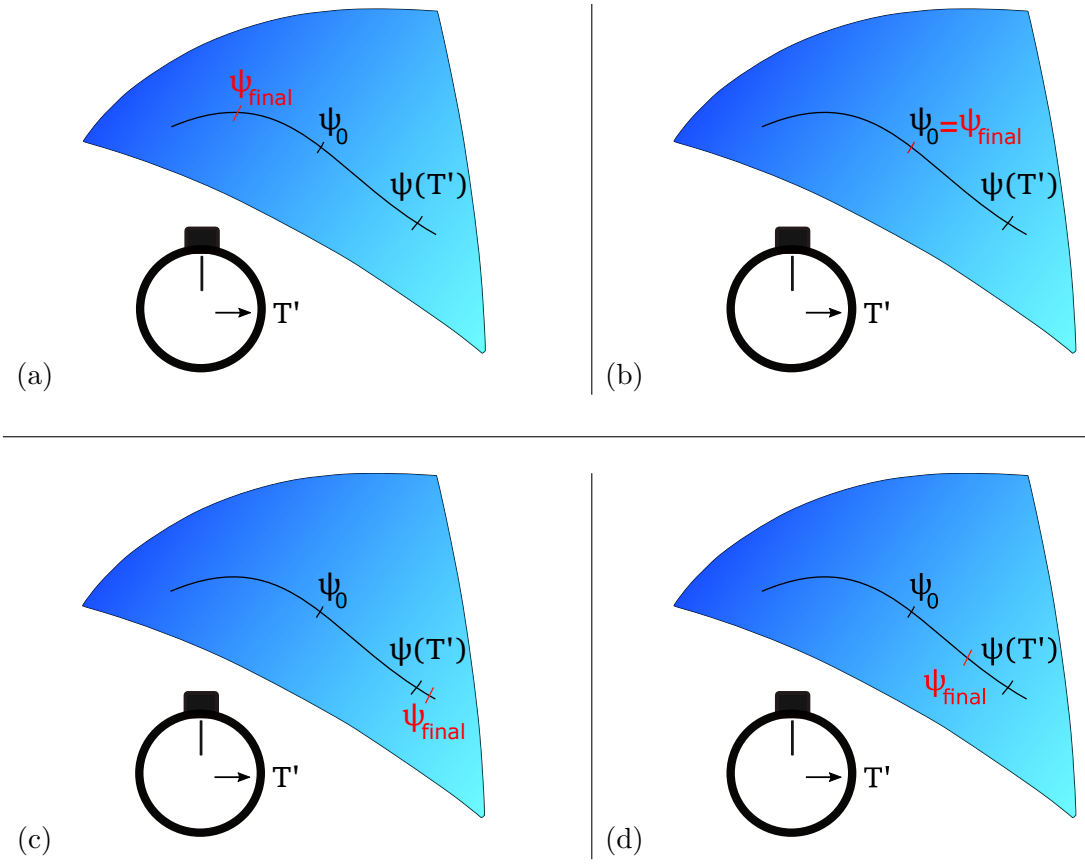

Figure 2. (a) A time translation corresponding to *rewinding*. That is, the final state is equal to $|\psi(-T)\rangle$ for some $T \geq 0$. (b) A special type of rewinding happens when the final state of the protocol is equal to the initial state. This is called *resetting*. (c) A time translation corresponding to *fast-forward*. That is, the final state corresponds to $|\psi(T)\rangle$ for $T > T'$. (d) The case where the final state of the protocol is $|\psi(T)\rangle$ with $0 < T \leq T'$ is a *slow-down*, and easy to achieve by waiting a certain amount of time and then performing one of the other protocols.

[ABB+20], and the recent work on unitary inversions [QDS+19a; QDS+19b]. In all these works, there is the implicit assumption that we can exert any operation we want on our system. That is, the systems on which we act are *controlled*, meaning that we can implement any interaction Hamiltonian that we want (or equivalently, arbitrary unitary operations). We do away with this assumption, and refer to such protocols are *universal*. A universal time-translation protocol must work even if we can only exert some unknown unitaries on our system.

Furthermore, in [QDS+19a; QDS+19b], evolution time is substituted by the proxy of number of gate applications, as it is usual in the quantum circuit model of quantum computing. In particular, the system is assumed not to evolve unless being acted upon by a unitary operator. We consider a more general model, although our results can always be reformulated in the circuit model, since we always take care in having a minimal amount of time for which we let our system evolve, and can thus think of the evolution of our system as the application of a unitary a repeated number of times. This will become clear in the formulation of the protocols.

This Chapter is structured as follows. In Section 2, we introduce the model we work with. We show how to obtain a purely mathematical problem from it. In Section 3, we study this mathematical model; that is, matrix polynomials. We prove the existence of many new kinds of matrix polynomials. In Section 4 we characterize all universal time translation protocols. Finally, in Section 5, we prove that qubits have rewinding protocols with arbitrarily high probability of success, at the cost of unbounded running time.

2. The model

In order to describe our time translation protocols we consider an idealized model of scattering particles, as shown in Figure 3. We consider a scenario where the experimental setup consists of two parts: a *controlled lab*, where we assume that we have complete control over quantum systems - which we call *probes*, and a *scattering region*, where n unknown uncontrollable *target systems* live. We assume the following:

(1) All the target systems have a finite radius of interaction, and they are sufficiently separated from one another so that they do not interact.

(2) All the target systems are identical. That is, they evolve under the same Hamiltonian H_0 and interact with the probes under the same interaction Hamiltonian H_I.

(3) All the target systems stay in place throughout the whole process, even while interacting with the probes.

(4) All the probes also have a finite radius of interaction, and there are paths starting and finishing in the lab where a probe would interact with exactly one target system.

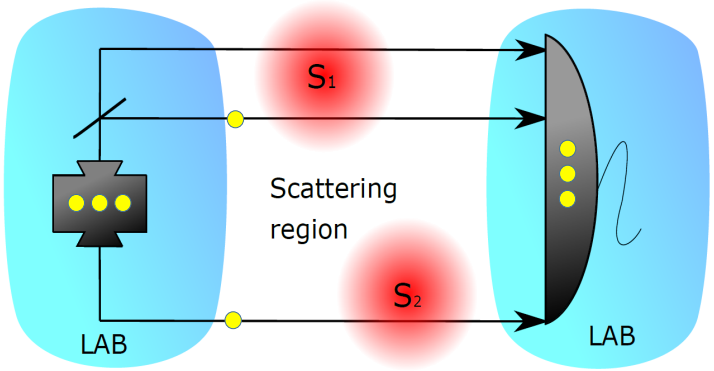

Figure 3. The larg red discs labelled S_1 and S_2 are the target systems and their interaction radius, which are well separated to ensure that the probes (small yellow discs) only interact with one at a time. The left section is the *preparation* part of the lab, where the probes are prepared in a desired quantum state. After scattering with one of the systems, these probes are then measured in the rightmost section of the lab in order to herald a successful run of the experiment.

In this model, we can describe a time warping protocol as a procedure in which we prepare probes, allow them to interact with the target systems and then measure them.

2.1. From scattering experiments to matrix polynomials and back. In this section we describe a family of scattering protocols that we will call *canonical*. As we will soon see, the effect of all such protocols on the target systems $1, ..., n$ is in one-to-one correspondence with some mathematical entities that we dub homogeneous tensor matrix polynomials.

We start by dividing a probe's Hilbert space into the factors $\mathcal{H}_p, \mathcal{H}_r$, where $\mathcal{H}_p$ denotes the probe's internal degree of freedom; and $\mathcal{H}_r = L^2(\mathbb{R}^3)$ is used to model the position of the probe's center of mass. Let $\{|i\rangle\}_{i=1}^{d_p}$ be an orthonormal basis for $\mathcal{H}_p$, and let R denote a qubit register within the lab.

Suppose that we prepare a probe in a superposition of states, one inside and another one outside the lab, controlled by a qubit register R_1. That is, the probe is in the state $\frac{1}{\sqrt{2}}(|0\rangle_R |\varphi\rangle_{pr} + |1\rangle_R |1\rangle_p |\Phi\rangle_r)$, where $|\varphi\rangle$ is the state that allows the probe to interact with system 1 or else be absorbed, $|\Phi\rangle$ is some bounded state within the lab. Such a state exists because of the assumptions 1-4.

The world line marked by state $|0\rangle_R$ hence propagates through the scattering region, interacting with system 1, initially in state $|\psi_1\rangle$, until it re-enters the lab after time Δt. When the probe re-enters the lab, its spatial degree of freedom is projected onto the state $|\tilde{\varphi}\rangle_r$, which we subsequently transform to the state $|\Phi\rangle$. The final (unnormalized) joint state of system 1 and the lab is thus

$$\frac{1}{\sqrt{2}} \left(\sum_{j=1}^{d_p} X_j |\psi_1\rangle |0\rangle_R |j\rangle_p + X_0 |\psi_1\rangle |1\rangle_R |1\rangle_p \right) \otimes |\Phi\rangle_r , \tag{77}$$

where the $d \times d$ matrices $X_0, ..., X_{d_p}$ are given by

$$\begin{aligned} X_0 &= e^{-iH_0 \Delta t}, \\ X_j &= \langle j|_p \langle \tilde{\varphi}|_r e^{-i(H_0 + H_P + H_I)\Delta t} |\varphi\rangle_{pr} , \end{aligned} \tag{78}$$

for $j = 1, ..., d_p$.

Since the state $|\Phi\rangle$ of the probe's center of mass plays no further role, we will omit it. Similarly, the states of the register and the internal degree of freedom of the probe, can be combined into a single label j: the final state of the joint system can thus be rewritten as $\frac{1}{\sqrt{2}} \sum_{j=0}^{d_p} X_j |\psi_1\rangle |j\rangle$.

The first step of a canonical scattering protocol consists in simultaneously sending a probe to each target system $1, ..., n$ in the above fashion. We invoke the targeting assumption to ensure that each probe interacts only with its targeted system. This step is iterated m times. Hence, at time $T' = m\Delta t$, each target system has interacted with m probes.

The final step of a canonical scattering protocol consists in post-selecting the lab's degree of freedom to the pure state

$$\sum_{\vec{\jmath}^{\,1},\ldots,\vec{\jmath}^{\,n}} g^*_{\vec{\jmath}^{\,1},\ldots,\vec{\jmath}^{\,n}} \left|j_1^1,\ldots,j_m^1\right\rangle \otimes \cdots \otimes \left|j_1^n,\ldots,j_m^n\right\rangle, \tag{79}$$

where each of the factors $k = 1, ..., n$ above represents the part of the lab system entangled with target system k due to the latter's interaction with m probes.

Let $|\psi_{1,\ldots,n}\rangle$ be the the initial state of the target systems $1, ..., n$. Then, the final state of systems 1 to n after a canonical scattering protocol is verified to be $G(X)\,|\psi_{1,\ldots,n}\rangle$, with $G(X)$ given by

$$\sum_{\vec{\jmath}^{\,1},\ldots,\vec{\jmath}^{\,n}} g_{\vec{\jmath}^{\,1},\ldots,\vec{\jmath}^{\,n}} X_{j_m^1} \cdots X_{j_1^1} \otimes \cdots \otimes X_{j_m^n} \cdots X_{j_1^n}. \tag{80}$$

Expressions of the form $G(X)$ will in the following be called *homogeneous tensor matrix polynomials of degree* m (for $n = 1$, those reduce to the old notion of homogeneous matrix polynomials of degree m). More formally, we define

Definition 2.1.1 (Tensor polynomial)**.** A homogeneous tensor polynomial f on non-commuting variables $X_1, ..., X_k$ of *tensor degree* $\deg^{\otimes}(f) = n$ with a tensor factors is an expression of the form

$$f(X_1, ..., X_k) := \sum_{i=1}^{p} c_i f_i^{(1)}(X_1, ..., X_k) \otimes \cdots \otimes f_i^{(a)}(X_1, ..., X_k),$$

where f_i^j is a word of length exactly n in the alphabet given by $\Sigma = \{X_1, ..., X_k\}$, and $c_i \in \mathbb{C}$.

Tensor polynomials have only recently been picked up by the mathematics community [Pro20]. There, it is also interesting to consider non-homogeneous polynomials, such as expressions of the form of $X_1 \otimes X_2 X_3$. Their notion of degree is then slightly different.

We just established that the effect of any canonical protocol is to propagate the target systems by a homogeneous tensor matrix polynomial. Conversely, for any homogeneous tensor matrix polynomial $G(X)$ of degree m, one can, by choosing the post-selection state (79) appropriately, devise a canonical scattering protocol that makes systems $1, ..., n$ leap to a state proportional to $G(X)\,|\psi_{1,\ldots,n}\rangle$. Thus, there exists a correspondence between tensor polynomials and canonical scattering protocols with the following properties:

Number of targets -1	$\longleftrightarrow$	Number of tensor products
Dimension of target	$\longleftrightarrow$	Size of matrix variables
Dimension of probes $+1$	$\longleftrightarrow$	Number of variables
Number of probes (per target)	$\longleftrightarrow$	(Tensor) degree of the polynomial

With the above definition, the duration of a canonical scattering protocol is proportional to the number of probes that we send to each system. More concretely,

$$T' = \deg^{\otimes}(f)\Delta t, \tag{81}$$

where Δt is the time that we allow the target systems to evolve freely or interact with the probes at each step of the protocol.

Note that there may exist other scattering protocols (not necessarily canonical) which map to the same tensor polynomial. In fact, some of those will have a higher probability

of success, or are experimentally preferable. This is illustrated with an example in Section 4.3.

With this map, we may reduce the problem of finding interesting universal protocols to the problem of finding interesting tensor polynomials. For example, suppose that there is a tensor polynomial $\Omega(X, Y)$ with two tensor factors such that, for any $d \times d$ matrices A, B,

$$\Omega(A, B) = \texttt{SWAP}.$$

Then, under the assumptions $1-4$, this means that there is a protocol $\mathcal{P}_\Omega$ that takes any d-dimensional states $|\psi_0\rangle, |\varphi_0\rangle$ from $|\psi_0\rangle \otimes |\varphi_0\rangle$ to $|\varphi_0\rangle \otimes |\psi_0\rangle$. That is, we are able to perform a universal permutation operation, no matter how the systems evolve or interact with other systems.

2.2. Other implementations. As we have anticipated, one should not take seriously the model presented in Section 2 as a scattering model. Our assumptions 1-4 can at most be satisfied approximately, and in that case the protocols that we describe will at most only work approximately. Rather, one should see this probe model as a motivation on how to physically realize the matrix polynomials. As long as one can think of a way to do this with their system of choice, then the protocols are physical and the theorems hold.

For example, maybe one is limited on the dimensionality of the probes they have access to. The index j of each variable X_j in the matrix polynomials only varies over $0, ..., d_p$, but for our constructions below we require a potentially very large number of matrix variables. This can be achieved by considering the action of several probes one after the other on the same system as the effect of a single virtual probe of much larger dimension. That is, by sending D probes one after the other to a single system, we get access to matrix products of the form $X_{j_D} ... X_{j_1}$, which we can relabel as $\tilde{X}_k$ where k now varies over $0, ..., (d_p + 1)^D - 1$. By also rescaling the unit of time in the protocols from Δt to $D\Delta t$, we can construct polynomials as if the probes had arbitrary dimension, even if the physical probes have no internal degrees of freedom at all ($d_p = 1$).

In fact, canonical scattering protocols can always be introduced for the case $d_p = 1$. In this predicament, there are just two physical matrix variables to play with: X_0 (absence of a probe) and X_1 (presence of a probe), so that the protocol consists on switching some Hamiltonians on and off. By regarding $D = 2$ consecutive probes as a single virtual probe, our new matrix variables are X_0X_0, X_0X_1, X_1X_0, X_1X_1: for $d > 1$ and generic matrices X_0, X_1, such matrix products are expected to span a 4-dimensional matrix space. Furthermore, by taking D sufficiently high, for generic X_0, X_1 we expect the resulting matrix products to span the whole set of $d \times d$ matrices. At that point, we can model fully generic matrix variables by linear combinations of $\{\tilde{X}_k\}_k$. That is how one can propagate systems $1, ..., n$ by tensor matrix polynomials of arbitrarily many variables.

3. Matrix polynomials

As we have seen, our protocols are equivalent to matrix polynomials. Thus, we turn to the mathematical problem of finding interesting matrix polynomials. The key to study polynomials in matrix algebras is the Cayley-Hamilton theorem:

Theorem 3.0.1 (Cayley-Hamilton)**.** *Let $A \in \mathrm{M}_n(\mathbb{C})$. Then A is a zero of its characteristic polynomial. That is,*

$$p_A(A) := \sum_{i=0}^{n} c_i A^i = 0,$$

where c_i is the coefficient of λ^i in $p_A(\lambda) := \det(\lambda \mathbb{1} - A)$.

Proof. Indeed, if v is an eigenvector of A with eigenvalue λ, then $p_A(A)v = p_A(\lambda)v = 0$. Therefore, if A is diagonalizable, one has $p_A(A) = 0$. The set of diagonalizable matrices is dense, and since the map $X \mapsto p_X(X)$ is polynomial in the entries of X, it is continuous. Therefore, $p_A(A) = 0$ for all $A \in \mathrm{M}_n(\mathbb{C})$. □

3.1. Central polynomials. Two classes of non-commutative polynomials are specially important to us.

Definition 3.1.1 (Polynomial identities)**.** Let $p \in \mathbb{C}[X_1, ..., X_k]$ be a non-zero non-commutative polynomial, $\mathcal{A}$ an algebra. We say that p is a *polynomial identity* (PI) for $\mathcal{A}$ if, for all $A_1, ..., A_k \in \mathcal{A}$,

$$p(A_1, ..., A_k) = 0.$$

Definition 3.1.2 (Central polynomials)**.** Let $p \in \mathbb{C}[X_1, ..., X_k]$ be a non-commutative polynomial, $\mathcal{A}$ an algebra. We say that $p \in \mathbb{C}[X_1, ..., X_k]$ is a *central polynomial* if $p(A_1, ..., A_k)$ is in the center of $\mathcal{A}$ for all $A_1, ..., A_k \in \mathcal{A}$ and $p(A_1, ..., A_k) \neq 0$ for some $A_1, ..., A_k \in \mathrm{M}_n(\mathbb{C})$.

Since the center of the ring of matrices is just those matrices proportional to the identity, we have that, if $p(X_1, ..., X_k)$ is a central polynomial for $\mathrm{M}_d(\mathbb{C})$, then

$$p(A_1, ..., A_k) \propto \mathbb{1}.$$

Therefore, central polynomials correspond to *resetting protocols*, under the correspondence explain in Section 2.1.

Example 3.1.3. The non-commutative polynomial

$$p(X, Y) := [X, Y]^2 = XYXY - XYYX - YXXY + YXYX$$

is central for $\mathrm{M}_2(\mathbb{C})$. Indeed, the Cayley-Hamilton theorem 3.0.1 for 2×2 matrices reads

$$p_A(A) = A^2 - \mathrm{tr}(A)A + \det(A)\mathbb{1} = 0,$$

and therefore $[A, B]^2 = -\det([A, B])\mathbb{1} \propto \mathbb{1}$ for all $A, B \in \mathrm{M}_2(\mathbb{C})$. Note that the proportionality constant depends on the matrices, and that is is non-zero except in a measure-zero set. These are usual properties of these polynomials.

Another important property of central polynomials which we will use later is that they are polynomial identites when restricted to a smaller dimensional matrix space:

Proposition 3.1.4. *Let $p(X_1, ..., X_k)$ be a central polynomial for $\mathrm{M}_d(\mathbb{C})$. Then p is a PI for $\mathrm{M}_{d'}(\mathbb{C})$ as long as $d' < d$.*

Proof. Indeed, consider the algebra embedding $\varphi : \mathrm{M}_{d'}(\mathbb{C}) \hookrightarrow \mathrm{M}_d(\mathbb{C})$ which takes a matrix A to $A \oplus 0_{d-d'}$. It follows that $\lambda \mathbb{1} = p(\varphi(A_1), ..., \varphi(A_k)) = \varphi(p(A_1, ..., A_k)) = p(A_1, ..., A_k) \oplus 0_{d-d'}$ therefore, it must be $\lambda = 0$ for all $A_1, ..., A_k \in \mathrm{M}_{d'}(\mathbb{C})$ and thus $p(A_1, ..., A_k) = 0$ for all $A_1, .., A_k \in \mathrm{M}_{d'}(\mathbb{C})$. □

It was shown for the first time in [For72] that there is a central polynomial in $\mathrm{M}_d(\mathbb{C})$ for all $d \in \mathbb{N}$. The construction is as follows.

Theorem 3.1.5 ([For72]). *For all $d \in \mathbb{N}$, there exists a homogeneous noncommutative polynomial $F(Y, X_1, ..., X_d)$ which is central in $\mathrm{M}_d(\mathbb{C})$, linear in the variables $X_1, ..., X_d$ and of degree d^2.*

Proof. We first define the linear map

$$\begin{aligned}\varphi : \mathbb{C}[x_1, ..., x_{d+1}] &\longrightarrow \mathbb{C}[X, Y_1, ..., Y_d] \\ x_1^{a_1} \cdots x_{d+1}^{a_{d+1}} &\mapsto X^{a_1} Y_1 X^{a_2} Y_2 \cdots X^{a_d} Y_d X^{a_{d+1}}.\end{aligned}$$

Let $G(X, Y_1, ..., Y_d)$ be the image under this map of the polynomial

$$g(x_1, ..., x_{d+1}) := \prod_{2 \leq i \leq d} (x_1 - x_i)(x_{d+1} - x_i) \prod_{2 \leq j < k \leq d} (x_j - x_k)^2.$$

Note that there are exactly $2(d-1) + (d-1)(d-2)$ terms in this product, and therefore G has degree $(d-1)d + d = d^2$. Then, Formanek's polynomial is

$$F(X, Y_1, ..., Y_d) := \sum_{i=1}^{d} G(X, Y_{\sigma^i(1)}, ..., Y_{\sigma^i(d)}),$$

where σ is the cyclic permutation of $(1, 2..., d)$. Indeed, this polynomial is linear in $Y_1, ..., Y_d$, homogeneous in X, and of degree d^2. We just need to check now that it is a central polynomial. As before, we argue by continuity that it is enough to check this for X diagonalizable. Also by linearity, we only need to evaluate Y_k to some elementary matrix $|i_k\rangle\langle j_k|$. We begin by evaluating X to be a diagonal matrix $\Lambda := \mathrm{diag}(\lambda_1, ..., \lambda_d)$. We have

$$\Lambda^{a_1} |i_1\rangle\langle j_1| \Lambda^{a_2} |i_2\rangle \cdots \langle j_{d-1}| \Lambda^{a_d} |i_d\rangle\langle j_d| \Lambda^{a_{d+1}} = \lambda_{i_1}^{a_1} \cdots \lambda_{i_d}^{a_d} \lambda_{j_d}^{a_{d+1}} |i_1\rangle\langle j_1| \cdots |i_d\rangle\langle j_d|,$$

and therefore

$$G(\Lambda, |i_1\rangle\langle j_1|, ..., |i_d\rangle\langle j_d|) = g(\lambda_{i_1}, ..., \lambda_{i_d}, \lambda_{j_d}) |i_1\rangle\langle j_1| \cdots |i_d\rangle\langle j_d|.$$

However, g was explicitely constructed to that $g(\lambda_{i_1}, ..., \lambda_{i_d}, \lambda_{j_d})$ is zero unless perhaps $j_d = i_1$ and $i_1, ..., i_d$ are all distinct, in which case

$$g(\lambda_{i_1}, ..., \lambda_{i_d}, \lambda_{j_d}) = \prod_{1 \leq i < j \leq d} (\lambda_i - \lambda_j)^2 =: D(\Lambda)$$

is the discriminant of $\{\lambda_1, ..., \lambda_d\}$. On the other hand, $|i_1\rangle\langle j_1| \cdots |i_d\rangle\langle j_d|$ is non-zero iff $j_1 = i_2, ..., j_{d-1} = i_d$. Since we have the additional constrain that $j_d = i_1$, in this situation we have that

$$G(\Lambda, |i_1\rangle\langle j_1|, ..., |i_d\rangle\langle j_d|) = D(\Lambda) |i_1\rangle\langle i_1|$$

and

$$F(\Lambda, |i_1\rangle\langle j_1|, ..., |i_d\rangle\langle j_d|) = D(\Lambda) \sum_{k=1}^{d} |i_k\rangle\langle i_k| = D(\Lambda) \mathbb{1}.$$

By linearity, $F(\Lambda, B_1, ..., B_d) \propto \mathbb{1}$ for all $B_1, ..., B_d \in \mathrm{M}_d(\mathbb{C})$. Now note that for all invertible A, one has

$$F(A\Lambda A^{-1}, AB_1A^{-1}, ..., AB_dA^{-1}) = AF(\Lambda, B_1, ..., B_d)A^{-1} \propto \mathbb{1}.$$

Therefore, $F(\Lambda, B_1, ..., B_d)$ holds when Λ is a diagonalizable matrix and, by continuity, for any matrix. □

There are today more illuminating proofs of the existence of central polynomials, but we are particularly interested in the properties of Formanek's polynomial, since they will allow us to construct optimal rewinding protocols. As a corollary, we get that there exist central polynomials of arbitrarily high degree. These correspond to rewinding protocols of arbitrarily high duration.

Corollary 3.1.5.1. *For all $m \geq d^2$, there is a homogeneous central polynomial for $\mathrm{M}_d(\mathbb{C})$ of degree m with the same number of variables.*

Proof. Indeed, suppose $\mathcal{C}(X_1, ..., X_k)$ is a central polynomial of degree n_d, linear in X_1. Then,

$$\mathcal{C}'(Y_1, ..., Y_k) := \mathcal{C}(Y_1^{m-n_d+1}, Y_2, ..., Y_k)$$

is also a central polynomial. To see that it is not a PI, consider some diagonalizable matrices $A_1, ..., A_k \in \mathrm{M}_d(\mathbb{C})$ such that $\mathcal{C}(A_1, ..., A_k) \neq 0$. Then,

$$\mathcal{C}'\left(A_1^{\frac{1}{m-n_d+1}}, A_2, ..., A_k\right) \neq 0.$$

□

3.2. Permutation polynomials. After central polynomials, the next most natural object are polynomials that are proportional to a fixed matrix other than the identity, since the existence of such polynomials would allow us to probabilistically implement fixed gates other than time translations. Of particular interest are multipartite gates like the `SWAP`, which permutes two systems. In general, if we have more than 2 target systems, one can perform any permutation using `SWAP` gates. In this section, we'll see that this is actually the most general fixed transformation that we can do with tensor polynomials.

Theorem 3.2.1 (Schur-Weyl duality)**.** *Let V be some finite dimensional complex vector space. Consider the following natural representation of the permutation group S_n on $V^{\otimes n}$ given by the linear extension of*

$$\sigma(v_1 \otimes \cdots \otimes v_n) := v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(n)}, \quad \sigma \in S_n, v_1, ..., v_n \in V.$$

and the diagonal representation of $\mathrm{GL}(V)$, given by the linear extension of

$$g(v_1 \otimes \cdots \otimes v_n) := gv_1 \otimes \cdots \otimes gv_n, \quad g \in \mathrm{GL}(V), v_1, ..., v_n \in V.$$

Then, $\mathbb{C}[\mathrm{GL}(V)]$ and $\mathbb{C}[S_n]$ are centralizers of each other in $\mathrm{L}(V^{\otimes n})$.

Proof. An excellent exposition is given in [Aub]. □

Theorem 3.2.2. *Let $p(X_1, ..., X_k)$ be a tensor polynomial with n tensor products, such that $p(A_1, ..., A_k) \propto M$ for all $A_1, ..., A_k \in \mathrm{M}_d(\mathbb{C})$, with $M \in \mathrm{M}_d(\mathbb{C})^{\otimes n}$. Then, M is a linear combination of permutation matrices or $p(A_1, ..., A_k) = 0$ for all $A_1, ..., A_k \in \mathrm{M}_d(\mathbb{C})$.*

Proof. We are going to suppose that there exists $A_1, ..., A_k$ such that $p(A_1, ..., A_k) \neq 0$. Note that any tensor polynomial $p(X_1, ..., X_k)$ satisfies

$$p(UX_1U^{-1}, ..., UX_kU^{-1}) = U^{\otimes n} p(X_1, ..., X_k) U^{\otimes -n}$$

for all $U \in \mathrm{GL}_d$. Therefore, if $p(A_1, ..., A_k) \propto M$ for all $A_1, ..., A_k \in \mathrm{M}_d(\mathbb{C})$, where $M \in \mathrm{M}_d(\mathbb{C})^{\otimes n}$ is a fixed matrix, then since there is a set of matrices for which the proportionality constant is not zero, we must have $U^{\otimes n} M U^{\otimes -n} = \lambda_U M$ for all invertible U as well. Note that

$$1 = \lambda_{UU^{-1}} = \lambda_U \lambda_{U^{-1}},$$

so that $\lambda : \mathrm{GL}_d \to \mathbb{C}^\times$ is a group homomorphism. However, all such maps factor through the determinant. Indeed, suppose that $x, y \in \mathrm{GL}_d$. We have $\lambda_{xyx^{-1}y^{-1}} = \lambda_x \lambda_y \lambda_{x^{-1}} \lambda_{y^{-1}} = \lambda_{xx^{-1}yy^{-1}} = 1$, so the commutator subgroup $[\mathrm{GL}_d, \mathrm{GL}_d]$, defined as the subgroup generated by all words of the form $xyx^{-1}y^{-1}$, is in the kernel of λ. Therefore, there is a unique homomorphism f such that the following diagram commutes.

$$\begin{array}{ccc} \mathrm{GL}_d & \xrightarrow{\lambda} & \mathbb{C}^\times \\ {\scriptstyle \pi}\downarrow & \nearrow_{f} & \\ \mathrm{GL}_d/[\mathrm{GL}_d, \mathrm{GL}_d] & & \end{array}$$

Here, π is the canonical projection onto the quotient group. However, $[\mathrm{GL}_d, \mathrm{GL}_d] = \mathrm{SL}_d$. Since $\mathrm{GL}_d/\mathrm{SL}_d = \mathbb{C}^\times$, with the canonical projection given by the determinant, we have as expected that λ factors through the determinant.

$$\begin{array}{ccc} \mathrm{GL}_d & \xrightarrow{\lambda} & \mathbb{C}^\times \\ {\scriptstyle \det}\downarrow & \nearrow_{f} & \\ \mathbb{C}^\times & & \end{array}$$

On the other hand $\lambda_{\alpha U} = \lambda_U$ for every $\alpha \in \mathbb{C}^\times$. Since there is always an α that makes αU have determinant 1, we have that λ is a constant function, and thus $\lambda_U = 1$. But that means that $U^{\otimes n} M = M U^{\otimes n}$ for all $U \in \mathrm{GL}_d$. By Schur-Weyl duality 3.2.1, it must be that $M \in \mathbb{C}[S_d]$. That is, M is a linear combination of permutation matrices. □

That is, the most general gates we can implement with out protocols are at most linear combinations of permutation matrices.

Definition 3.2.3 (Permutation polynomials)**.** Let $p(X_1, ..., X_k)$ be a tensor polynomial with n tensor products. We say that p is a *permutation polynomial* if p is not a PI, and there is an $M \in \mathbb{C}[S_n] \subset \mathrm{L}((\mathbb{C}^d)^{\otimes n})$ such that

$$p(A_1, ..., A_k) \propto M, \quad \forall A_1, ..., A_k \in \mathrm{M}_d(\mathbb{C}).$$

Since the identity is a permutation, this concept is a generalization of central polynomials. Indeed, note that Theorem 3.2.2 says that, when there are no tensor products, the only matrix polynomials which, when evaluated with matrices of a certain size, are always proportional to a fixed matrix are central polynomials or polynomial identities.

Example 3.2.4 (A `SWAP` polynomial)**.** Recall first that

$$\texttt{SWAP} = \mathbb{1} \otimes \mathbb{1} + \sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z. \tag{82}$$

The idea is that, since we know how to construct polynomials which end up being proportional to $\mathbb{1}$ when evaluated on 2×2 matrices, to find polynomials that are proportional to the different Pauli matrices. We know from Theorem 3.2.2 that they don't exist, but we can work around this by allowing our polynomials to output the pauli matrices in different bases. We are going to look for three matrix polynomials $f_x(X,Y), f_y(X,Y), f_z(X,Y)$ which have the properties of the Pauli matrices when evaluated on 2×2 matrices. That is, that they square to the identity and that they anticommute. From the Cayley-Hamilton theorem 3.0.1 we know that any traceless 2×2 matrix squares to the identity up to a proportionality constant. Therefore, we consider commutators. We have for all $A, B \in \mathrm{M}_2(\mathbb{C})$ that

$$\{[A,B],[[A,B],B]\} = [A,B]^2B - B[A,B]^2 \propto B - B = 0.$$

Thus, we can define $f_x(X,Y) := [X,Y]$, $f_y(X,Y) := [[X,Y],Y]$ and $f_z := f_xf_y$. The image of these polynomials on 2×2 matrices is a projective representation of the Pauli group which lifts to an irreducible injective representation, known to be unitarily equivalent to the Pauli matrices. If we now call $f_0(X,Y) := [X,Y]^2$, we have what we were looking for. However, the proportionality constant in all these polynomials is in principle different, so we cannot add them together and recover the `SWAP`. In order to correct this, we make use of the fact that all our polynomials squared are central polynomials. Remembering that, we see that

$$f_{\texttt{SWAP}} := f_xf_xf_0 \otimes f_yf_yf_0 + f_xf_yf_y \otimes f_xf_0f_0 + f_yf_xf_x \otimes f_yf_0f_0 + f_xf_yf_0 \otimes f_xf_yf_0.$$

This is a non-homogeneous tensor polynomial in two variables X, Y. This stems from the fact that f_x and f_y do not have the same degree, and therefore can be corrected by redefining $f_x(X,Y) := [X^2, Y^2]$ and $f_y(X,Y) := [[X^2,Y],Y]$. The anticommutation relations still hold thanks to the Cayley-Hamilton theorem 3.0.1. Therefore, with these definitions, for all $A, B \in \mathrm{M}_2(\mathbb{C})$,

$$f_{\texttt{SWAP}}(A,B) = \lambda_x^2\lambda_y^2\lambda_0^2(A,B)\texttt{SWAP}.$$

The polynomial we have just constructed has degree $12+12$. Comparing this example to Example 3.1.3, which shows the simplest central polynomial, one hopes that there will be simpler `SWAP` polynomials. It turns out not so be so easy. For us, having a homogeneous tensor polynomial is quite important, since these are the polynomials that we can implement in our model, and this forces an increase in the degree of the polynomial. We can numerically determine that the simplest `SWAP` polynomial has 40 summands and is of degree 5. If one is not concerned with that, however, one can do better, as shown by [Pro22]. For 2-variables homogeneous tensor polynomials, we believe the simplest `SWAP` polynomial to be

$$
\begin{aligned}
(83) \qquad \tilde{\Omega}(V,W) :=& VWWVW \otimes VWWVW - VWWVW \otimes WWVVW \\
&- VWWWV \otimes VWWVW + VWWWV \otimes VWWWV \\
&+ VWWWV \otimes WVWVW - VWWWV \otimes WWVWV \\
&- VWWWW \otimes VWVWV + VWWWW \otimes WVVWV \\
&- WVWWV \otimes VWWVW - WVWWV \otimes VWWWV \\
&+ WVWWV \otimes WVWVW + WVWWV \otimes WWWVV \\
&- WVWWW \otimes VWVWV + WVWWW \otimes VWWVV \\
&+ WVWWW \otimes WVWVV - WVWWW \otimes WWVVV \\
&+ WWVWV \otimes VWWVW - WWVWV \otimes VWWWV \\
&- WWVWV \otimes WVWVW + WWVWV \otimes WVWWV \\
&- WWVWW \otimes VWVVW + WWVWW \otimes VWVWV \\
&+ WWVWW \otimes WVVVW - WWVWW \otimes WVVWV \\
&+ WWWVV \otimes VWWWV - WWWVV \otimes WVWVW \\
&+ WWWVV \otimes WWVVW - WWWVV \otimes WWWVV \\
&+ WWWVW \otimes VWVWV - WWWVW \otimes VWWVV \\
&- WWWVW \otimes WVWVV + WWWVW \otimes WWVVV \\
&- WWWWV \otimes VVWWV + WWWWV \otimes VWVVW \\
&+ WWWWV \otimes VWVWV - WWWWV \otimes WVVVW \\
&- WWWWV \otimes WVWVV + WWWWV \otimes WWVVV \\
&+ WWWWW \otimes VVWVV - WWWWW \otimes VWVVV.
\end{aligned}
$$

Such a complicated polynomial will necessarily give a very low probability of success. Nonetheless, we can generalize Theorem 3.1.5 for permutation polynomials. That is, we prove that they exist. An alternative less pedestrian proof of existence of `SWAP` polynomials can be found in [Pro22].

Theorem 3.2.5 (Existence of permutation polynomials)**.** *For all $d, n \in \mathbb{N}$, and $M \in S_n \subset \mathrm{L}((\mathbb{C}^d)^{\otimes n})$ there exists a homogeneous tensor polynomial $\mathcal{P}_M$ with n tensor products which is proportional to M when evaluated on $\mathrm{M}_d(\mathbb{C})$.*

Proof. The basic construction block of the proof is the existence of a `SWAP` polynomial for all dimensions d and $n = 2$ parties. Indeed, let Ω be such a `SWAP` polynomial. Then, consider the tensor polynomial

$$\Omega^{ij} := \Omega \otimes \bigotimes_{k \neq i,j} \mathcal{C}_k,$$

where $\mathcal{C}_k$ is a central polynomial of the same degree as Ω. This polynomial exists by Corollary 3.1.5.1. If necessary, we first artificially increase the degree of Ω. The polynomial Ω^{ij} now is always proportional to a `SWAP` in systems i, j and the identity in the other systems. Given that we can decompose any permutation as a product of swaps, we may equally construct a product of polynomials Ω^{ij} to recover a polynomial which is proportional to any permutation π when evaluated on $\mathrm{M}_d(\mathbb{C})$.

Conversely, in order to build a SWAP polynomial, we are going to construct polynomials which project to the symmetric and the antisymmetric subspaces of $\mathbb{C}^d \otimes \mathbb{C}^d$. That is, consider the symmetric and the antisymmetric subspaces, defined as

$$\mathrm{Sym}^2(\mathbb{C}^d) := \{z \in \mathbb{C}^d \otimes \mathbb{C}^d \mid \mathtt{SWAP} z = z\} \quad \Lambda^2(\mathbb{C}^d) := \{z \in \mathbb{C}^d \otimes \mathbb{C}^d \mid \mathtt{SWAP} z = -z\}.$$

They have dimensions

$$d_S = \frac{d(d+1)}{2} \quad d_A = \frac{d(d-1)}{2},$$

respectively. Indeed, a basis of $\mathrm{Sym}^2(\mathbb{C}^d)$ is $\{e_i \otimes e_i, e_i \otimes e_j + e_j \otimes e_i\}_{j>i=1}^d$, and a basis of $\Lambda^2(\mathbb{C}^d)$ is given by $\{e_i \otimes e_j - e_j \otimes e_i\}_{j>i=1}^d$. Let us call $\Pi_{\mathrm{Sym}^2(\mathbb{C}^d)}, \Pi_{\Lambda^2(\mathbb{C}^d)}$ the projectors onto these subspaces. Note that

$$\mathtt{SWAP} = \Pi_{\mathrm{Sym}^2(\mathbb{C}^d)} - \Pi_{\Lambda^2(\mathbb{C}^d)}.$$

This decomposition, rather than Equation (82) is what we are going to use, but in general following the steps of Example 3.2.4.

Let us then call $\mathcal{C}^S(Z_1, ..., Z_{m_S}), \mathcal{C}^A(Z_1, ..., Z_{m_A})$ central polynomials for matrices of dimensions d_S, d_A, respectively. Let us now consider matrices of size $d^2 \times d^2$ acting on $\mathbb{C}^d \otimes \mathbb{C}^d = \mathrm{Sym}^2(\mathbb{C}^d) \oplus \Lambda^2(\mathbb{C}^d)$. Note that any matrix C in the linear span of $\{V \otimes V\}_{V \in \mathrm{M}_d(\mathbb{C})}$ commutes with the SWAP, and preserves the symmetric and antisymmetric subspaces, so that we may write

$$S = \begin{pmatrix} S_S & 0 \\ 0 & S_A \end{pmatrix}, \tag{84}$$

where $S_S \in \mathrm{M}_{d_S}(\mathbb{C})$ and $S_A \in \mathrm{M}_{d_A}(\mathbb{C})$, and we consider a basis given by a basis of $\mathrm{Sym}^2(\mathbb{C}^d)$, and then a basis of $\Lambda^2(\mathbb{C}^d)$.

Now, given a certain amount of matrix variables $X_1, ..., X_m$, consider linear combinations $S_i(X_1, ..., X_m)$ of the letters $\{X_j \otimes X_j, X_j \otimes X_k + X_k \otimes X_j\}_{j,k=1}^m$ such that

$$\mathcal{S}(X_1, ..., X_m) := \mathcal{C}^S(S_1(X_1, ..., X_m), ..., S_{m_S}(X_1, ..., X_m)) \tag{85}$$

is not a PI. Such combinations must exist, as we will see a posteriori. Now, for all $A_1, ..., A_m \in \mathrm{M}_d(\mathbb{C})$, we have

$$\mathcal{S}(A_1, ..., A_m) =$$
$$\begin{pmatrix} \mathcal{C}^S(S_1(A_1, ..., A_m)_S, ..., S_{m_S}(A_1, ..., A_m)_S) & 0 \\ 0 & \mathcal{C}^S(S_1(A_1, ..., A_m)_A, ..., S_{m_S}(A_1, ..., A_m)_A) \end{pmatrix}$$
$$\propto \begin{pmatrix} \mathbb{1}_S & 0 \\ 0 & 0_A \end{pmatrix} = \Pi_{\mathrm{Sym}^2(\mathbb{C}^d)},$$

where we have used Proposition 3.1.4 together with the fact that $d_A < d_S$. The proportionality constant is not zero, because the span of $\{X_j \otimes X_j, X_j \otimes X_k + X_k \otimes X_j\}_{j,k=1}^m$ is equal to the matrices of the form given by Equation (84), as a quick dimension count shows. Therefore, using linear combinations of said matrices, we can construct any matrix in $\mathrm{M}_{d_S}(\mathbb{C})$ as $S_i(A_1, ..., A_m)_S$, and $\mathcal{C}^S$ is not a PI, so there must be some combination which gives a non-zero proportionality constant.

Similarly, we can construct a tensor polynomial which plays the role of the antisymmetric projector. To this effect, we now consider the antisymmetric part of $\mathrm{M}_{d^2}(\mathbb{C})$. That is linear combinations of the letters $\{Y_j \otimes Y_k - Y_k \otimes Y_j\}_{j,k=1}^m$. Such matrices anticommute

with the SWAP, and therefore interchange the spaces $\mathrm{Sym}^2(\mathbb{C}^d)$ and $\Lambda^2(\mathbb{C}^d)$. That is, they correspond to matrices of the form

$$A = \begin{pmatrix} 0_S & A_1 \\ A_2 & 0_A \end{pmatrix}.$$

Conversely, every matrix of this form, can be written as such a linear combination, again since the dimensions of both spaces are the same. Let us call, as usual, $Y_j \wedge Y_k := Y_j \otimes Y_k - Y_k \otimes Y_j$. Then, consider the polynomials

$$P_{ijkl}(Y_1, ..., Y_m) := (Y_i \wedge Y_j)\mathcal{S}(Y_1, ..., Y_m)(Y_k \wedge Y_l),$$

where $\mathcal{S}$ is the polynomial defined in Equation (85). It follows that, when evaluated on $d \times d$ matrices, one obtains

$$P_{ijkl}(A_1, ..., A_m) \propto (A_i \wedge A_j) \begin{pmatrix} \mathbb{1}_S & 0 \\ 0 & 0 \end{pmatrix} (A_k \wedge A_l) = \begin{pmatrix} 0_S & 0 \\ 0 & A \end{pmatrix}$$

for some $A \in \mathrm{M}_{d_A}(\mathbb{C})$. Conversely, any matrix of this form can be obtained in this way, and therefore, the polynomial

$$\mathcal{A}(Y_1, ..., Y_m) := \mathcal{C}^A(A_1(Y_1, ..., Y_m), ..., A_{m_A}(Y_1, ..., Y_m))$$

is not a PI for some linear combinations of the letters $\{P_{ijkl}(Y_1, ..., Y_m)\}_{i,j,k,l=1}^m$, which we denote by $A_i(Y_1, ..., Y_m)$. It follows that

$$\mathcal{A}(A_1, ..., A_m) \propto \Pi_{\Lambda^2(\mathbb{C}^d)}$$

for all $A_1, ..., A_m \in \mathrm{M}_d(\mathbb{C})$.

In order to substract this polynomials to obtain a SWAP polynomial, we need to make sure that they have the same proportionality constants in $\mathcal{S}, \mathcal{A}$, and that they have the same degree. Otherwise, we won't get a homogeneous tensor polynomial. To correct this second part we may use the same trick as in Example 3.2.4. That is, using the polynomials $\mathcal{S}(I_S \otimes I_S)$ and $\mathcal{A}(I_A \otimes I_A)$, where I_S,I_A are central polynomials for $\mathrm{M}_d(\mathbb{C})$ of sufficiently high degree, so that we can make $\deg(\mathcal{S}) + \deg(I_S) = \deg(\mathcal{A}) + \deg(I_A)$ .These exist due to Corollary 3.1.5.1. We rename $\mathcal{S}$ and $\mathcal{A}$ to be such polynomials.

In order to correct the first part, we consider yet another polynomial $F(T_1, ..., T_m)$, this time central for matrices of size $d^2 \times d^2$. The new polynomial

$$\begin{aligned}\tilde{F}(X_1, ..., X_m, Y_1, ..., Y_m, Z_1, ..., Z_{2m}) := \\ F([\mathcal{S}(X_1, ..., X_m) + \mathcal{A}(Y_1, ..., Y_m)][Z_1 \otimes Z_2][\mathcal{S}(X_1, ..., X_m) + \mathcal{A}(Y_1, ..., Y_m)], ..., \\ [\mathcal{S}(X_1, ..., X_m) + \mathcal{A}(Y_1, ..., Y_m)][Z_{2p-1} \otimes Z_{2p}][\mathcal{S}(X_1, ..., X_m) + \mathcal{A}(Y_1, ..., Y_m)]).\end{aligned}$$

is a homogeneous central tensor polynomial for $\mathrm{M}_d(\mathbb{C})$, since F is a homogeneous central polynomial and all matrices are in the span of elementary tensors. On the other hand, we may write it as

$$\tilde{F} = \mathcal{S}F^1\mathcal{S} + \mathcal{S}F^2\mathcal{A} + \mathcal{A}F^3\mathcal{S} + \mathcal{A}F^4\mathcal{A}$$

for some tensor polynomials F^1, F^2, F^3, F^4.

Now, let $f(X_1, ..., X_m, Y_1, ..., Y_m, Z_1, ..., Z_{2p})$ be such that

$$\tilde{F}(A_1, ..., A_m, B_1, ..., B_m, C_1, ..., C_{2p}) = f(A_1, ..., A_m, B_1, ..., B_m, C_1, ..., C_{2p})\mathbb{1}_{d^2}$$

for all $A_1, ..., A_m, B_1, ..., B_m, C_1, ..., C_{2p} \in \mathrm{M}_d(\mathbb{C})$. We have that

$$\begin{aligned} f(A_1, ..., C_{2p})\Pi_{\mathrm{Sym}^2(\mathbb{C}^d)} &= \Pi_{\mathrm{Sym}^2(\mathbb{C}^d)}\tilde{F}(A_1, ..., C_{2p})\Pi_{\mathrm{Sym}^2(\mathbb{C}^d)} \\ &= \mathcal{S}(A_1, ..., A_m)F^1(A_1, ..., C_m)\mathcal{S}(A_1, ..., A_m) \end{aligned}$$

Analogously,

$$\mathcal{A}(B_1, ..., B_m)F^4(X_1, ..., Z_{2p})\mathcal{A}(B_1, ..., B_m) = f(A_1, ..., C_{2p})\Pi_{\Lambda^2(\mathbb{C}^d)}$$

Therefore, we define

$$\Omega := \mathcal{S}F^1\mathcal{S} - \mathcal{A}F^4\mathcal{A}.$$

□

3.3. Time translating polynomials. We proceed now to construct polynomials which perform a wide variety of universal time translation protocols. As we will see in Theorem 4.0.1, these polynomials are going to provide a full list of allowed universal time translation protocols.

3.3.1. *Rewinding polynomials.* Consider Formanek's polynomials $F(X, Y_1, ..., Y_d)$ constructed in Theorem 3.1.5. We construct a new family of "rewinding polynomials" implicitely by evaluating F on X and $Y_i = Z_i X^s$. It follows that

$$F(X, Z_1X^s, ..., Z_dX^s) = R(X, Z_1, ..., Z_d)X^s \propto \mathbb{1}.$$

for some other polynomial R. Since F has degree $d^2 + 1$, and is linear in $Y_1, ..., Y_d$, we get that R must have degree $d^2 + (d-1)s$ and is linear in $Z_1, ..., Z_d$. On the other hand, if we evaluate R on $d \times d$ matrices $V, W_1, ..., W_d$, we have

$$R(V, W_1, ..., W_d)V^s \propto \mathbb{1}.$$

Therefore,

$$R(V, W_1, ..., W_d) \propto V^{-s}, \quad W_1, ..., W_d \in \mathrm{M}_d(\mathbb{C}), V \in \mathrm{GL}_d.$$

For $n = 2$ parties, we consider the following "fast-rewinding" tensor polynomials:

$$\begin{aligned} D(X, Y_1, ..., Y_m) := (R(X, Y_1, ..., Y_m) \otimes \mathbb{1})\Omega(X, Y_1, ..., Y_m) \\ \cdot (\mathbb{1} \otimes R(X, Y_1, ..., Y_m))\Omega(X, Y_1, ..., Y_m). \end{aligned}$$

When evaluated on $d \times d$ matrices $V, W_1, ..., W_m$, we have

$$D(V, W_1, ..., W_m) \propto (V^{-s} \otimes \mathbb{1})\mathtt{SWAP}(\mathbb{1} \otimes V^{-s})\mathtt{SWAP} = (V^{-2s} \otimes \mathbb{1}).$$

In general, we define for n parties the polynomials

$$D^j := R_j\Omega^{jk}R_k\Omega^{jk},$$

where R_j is a short-hand to introduce the polynomial R in system j: $\mathbb{1}\otimes\cdots\otimes R\otimes\cdots\otimes\mathbb{1}$. This tensor polynomial, when evaluated on $d \times d$ matrices $V, W_1, ..., W_m$, satisfies

$$D^j(V, W_1, ..., W_m) = V_j^{-ns}.$$

3.3.2. *Fast-forwarding protocols.* Consider the 2-tensor polynomal defined as

$$E(X, Y_1, ..., Y_m) := (X^s \otimes \mathbb{1})\Omega(X, Y_1, ..., Y_m)(\mathbb{1} \otimes X^s)\Omega(X, Y_1, ..., Y_m), \tag{86}$$

where Ω is a SWAP polynomial, proven to exist in Theorem 3.2.5. Note that this is a polynomial of tensor degree $s + 2\deg\Omega$. Evaluated on $d \times d$ matrices $V, W_1, ..., W_m$, we obtain a matrix proportional to

$$E(V, W_1, ..., W_m) \propto (V^s \otimes \mathbb{1})\texttt{SWAP}(\mathbb{1} \otimes V^s)\texttt{SWAP} = V^{2s} \otimes \mathbb{1}.$$

In general, we define for n parties,

$$E^j(X, Y_1, ..., Y_m) := X_j^s \prod_{k \neq j} \Omega^{jk}(X, Y_1, ..., Y_m) X_k^s \Omega^{jk}(X, Y_1, ..., Y_m),$$

where X_j^s again denotes the expression $\mathbb{1} \otimes \cdots X^s \otimes \cdots \mathbb{1}$, where X^s is in position j. This tensor polynomial, when evaluated on $d \times d$ matrices $V, W_1, ..., W_m$, satisfies

$$E^j(V, W_1, ..., W_m) \propto V_j^{ns}.$$

4. All the time translations

We are now ready to state the main result of this Chapter. A theorem characterizing all the possible universal time translation protocols:

Theorem 4.0.1. [⇒] [⇐] *Let $\mathcal{P}(d, (T_1, ..., T_n), T')$ be a universal time translation protocol of duration T' acting on n target systems of dimension d, each evolving freely under a Hamiltonian H_0 and interacting with a probe system with a bounded Hamiltonian H_I, such that system j is sent to state $e^{-iH_0T_j}|\psi_j\rangle$. Then, it must be that*

$$\sum_{i:T_i>0} T_i + \sum_{i:T_i<0} |T_i|(d-1) \leq nT'. \tag{87}$$

Conversely, if $(d, (T_1, ..., T_n), T')$ satisfy Equation (87) with a strict inequality, then there is a universal time translation protocol with such parameters that works for almost all H_0, H_I with a non-zero probability of success.

Before proceeding to prove this theorem, let us look more closely at the meaning. If we only have one target system ($n = 1$), then Equation (87) means that

$$-\frac{T'}{d-1} \leq T_1 \leq T'.$$

That is, a protocol applied to a single system cannot perform fast-fowarding, and the amount of time that it can rewind is bounded by the dimension. However, if we have more than one system, the more interesting protocols become available. For example, we may achieve target times of $(nT', 0, ...0)$ or $(-nT'/(d-1), 0, ..., 0)$ with a protocol of running time T'. That is, we are transfering the time from the extra target systems to the first one in order to break the limits of a protocol acting on a single system.

Equivalently, we may rephrase Equation (87) as a set of properties that evolution time must follow.

- Evolution time cannot be created.
- Evolution time can be transferred between two identical systems at no cost.
- Evolution time of a system of dimension d can be inverted at a cost $(d-1)$.

- Evolution time can be destroyed.

All the universal time translation protocols that exist are the ones that allowed by this "axioms".

4.1. Existence. Let us begin by proving the "only if" part of Theorem 4.0.1.

Proof of Theorem 4.0.1, $\Leftarrow$. First note that to get all protocols satisfying Equation (87), one only needs to achieve the "extremal" protocols with output times $(0, ..., -nT'/(d-1), ...0)$ and $(0, ..., nT', ...0)$, as every other output time can be achieved by waiting for a while and then performing some of these protocols for a shorter period of time T''.

Using the canonical correspondence described in Section 2.1, we obtain for each of the polynomials described in this Section protocols by taking $V = e^{-iH_0\Delta t}$, and suitable W_i that require the same amount of time Δt to be implemented. These protocols take an amount of time given by Equation (81), and so we must consider the degree of our polynomials. We have that

$$\deg^{\otimes}(D^j) = 2\deg^{\otimes}(\Omega) + \deg^{\otimes}(R) = 2\deg^{\otimes}(\Omega) + d^2 + (d-1)s.$$

Therefore, we have a protocol of duration

$$T' = (2\deg^{\otimes}(\Omega) + d^2)\Delta t + (d-1)s\Delta t \tag{88}$$

with output times $(0, ..., -ns\Delta t, ..., 0)$.

On the other hand, for the fast-forward polynomial we have

$$\deg^{\otimes}(E^j) = 2\deg^{\otimes}(\Omega) + s,$$

from which we obtain a protocol of duration

$$T' = 2\deg^{\otimes}(\Omega)\Delta t + s\Delta t \tag{89}$$

that achieves output times $(0, ..., ns\Delta t, ..., 0)$.

Choosing a sufficiently small interaction time Δt and a sufficiently big s we can get Equations (88 and 89) to indeed describe the total duration of the protocol T', if it was fixed before hand. Making Δt even smaller if necessary, we can get any protocol given by Equation (87).

□

4.2. Necessary limits. To prove that Equation (87) is also necessary, we resort to the Dyson series:

Theorem 4.2.1 (Dyson). *Let H_0, H_I be self-adjoint operators acting on $\mathcal{H}$, with $H_I \in \mathrm{B}(\mathcal{H})$. Then, the operator*

$$U(t) = \sum_{j=0}^{\infty}(-i)^j \int_{t\geq t_1\geq ...\geq t_j\geq 0} dt_1 \cdots dt_j e^{-iH_0(t-t_1)} H_I e^{-iH_0(t_1-t_2)} H_I \cdots \cdots e^{-iH_0(t_{j-1}-t_j)} H_I e^{-iH_0 t_j} \tag{90}$$

is a strongly continuous one-parameter unitary group, which describes the evolution of a system evolving under $H_0 + H_I$.

Proof of Theorem 4.0.1, ⇒. We consider first the case where we have just one target system. Before we switch on the interaction, we consider that the target system and probe i are evolving under the independent hamiltonians H_0 and $H_P^{(i)}$, respectively. In particular, we assume that H_0 and $H_P := \sum_{i=1}^N H_P^{(i)}$ commute. We consider Dyson's series 4.2.1 with free Hamiltonian $H_0 + \sum_{i=1}^N H_P^{(i)}$.

Whatever post-selection we might be applying to the probes will just affect system 1 through the bipartite terms H_I in Equation (90). Taking into account that $e^{-i(H_P+H_0)t} = e^{-iH_Pt}e^{-iH_0t} = e^{-iH_0t}e^{-iH_Pt}$, we may group the terms of the form $e^{iH_Pt_k}H_Ie^{-iH_Pt_{k+1}}$ and so after measuring we are left with a continuous linear combination of terms of the form

$$\Pi_1 e^{-iH_0t_1}\Pi_2 e^{-iH_0t_2}..., \tag{91}$$

where $\sum_i t_i = t$, and the Π_j depend on the Hamiltonians H_I, H_P and the measurement outcome.

In general in a protocol we take the segment $[0, T']$ and choose some sequence of times $0 = T_0 < T_1 < ... < T_M = T'$. Between times T_i and T_{i+1} we can decide whether we send some probes to interact with the target system of not. If we do not, the target system evolves via the operator $e^{-iH_0(T_{i+1}-T_i)}$. Otherwise, it evolves via an operator that we can decompose as before in terms of the form of Equation (91). At the end of the protocol, after the post-selection, we thus have a linear combination of terms of the form of Equation (91), but with $\sum_i t_i = T'$.

Therefore, the target system will have a state proportional to

$$\sum_j A_j \left|\psi_1\right\rangle\!\left\langle\psi_1\right| A_j^\dagger,$$

where $|\psi_1\rangle$ is the initial state of the target, the index j labels the measurement outcomes that we post-select to and A_j is an operator which can be expressed as a continuous linear combination of terms of the form of Equation (91) with $\sum_i t_i = T'$.

Suppose now that we have one target, and we are interested in effecting the transformation $|\psi_1\rangle \mapsto e^{iH_0T}|\psi_1\rangle$. We must therefore have

$$\sum_j A_j \left|\psi_1\right\rangle\!\left\langle\psi_1\right| A_j^\dagger \propto e^{-iH_0T_1}\left|\psi_1\right\rangle\!\left\langle\psi_1\right| e^{iH_0T_1},$$

with $T_1 \equiv -T < 0$. It follows by convexity that the above equation can hold for some non-zero proportionality scalar only when all the non-zero A_j's are proportional to e^{iH_0T}. Let then A be one of such non-zero terms $\{A_j\}_j$ corresponding to one possible measurement result of the probes. From the above, we have that

$$A \propto e^{iH_0T}. \tag{92}$$

for all Hamiltonians H_0, H_I, H_C.

Suppose then that there exists such an operator A, and let H_0 be any generic Hamiltonian such that the proportionality constant of Equation (92) does not vanish, for some fixed H_I, H_P. Since H_0 is a generic operator acting on a d-dimensional Hilbert space, it must admit a Jordan decomposition of the form $B^{-1}H_0B = \sum_{i=1}^d \alpha_i |i\rangle\!\langle i|$, with $\{\alpha_k\}_{k=1}^d \subset \mathbb{C}$, for some invertible $d \times d$ matrix B. Since $A = f(\vec{\alpha})e^{iH_0T}$ for some scalar $f(\vec{\alpha})$, we have that

$$\langle j|\,\tilde{A}\,|j\rangle = f(\vec{\alpha})e^{i\alpha_j T}, \tag{93}$$

for $j = 1, ..., d$, where here $\tilde{A}$ denotes $B^{-1}AB$. On the other hand, braketing Equation (91) between $\langle j|\,B^{-1} \bullet B\,|j\rangle$ we obtain

$$\sum_{\ell_1,\ell_2,\ldots} \langle j|\,\tilde{\Pi}_1\,|\ell_1\rangle\,\langle \ell_1|\,\tilde{\Pi}_2\,|\ell_2\rangle \cdots\; e^{-i(\alpha_{\ell_1}t_1+\alpha_{\ell_2}t_2+\cdots)},$$

where we are again denoting $\tilde{\Pi}_i := B^{-1}\Pi_i B$. By grouping the terms in the exponent with the same α_i we may write each term as some coefficient times $e^{-i\vec{\alpha}\cdot\vec{t}}$, where now $\vec{t}$ lives in $\mathbb{R}^d$, but still satisfies $\sum_i t_i = T'$ and positivity. Therefore, the full linear combination can be written as

$$\langle j|\,\tilde{A}\,|j\rangle = \int_{\mathcal{T}} c^j(\vec{t})e^{-i\vec{\alpha}\cdot\vec{t}}d\vec{t}, \tag{94}$$

by choosing appropiate coefficients $c^j(\vec{t})$, and where $\mathcal{T} := \{\vec{t} \in \mathbb{R}^d, t_i \geq 0, \sum_i t_i = T'\}$. From Equations (93) and (94), it follows that $f(\vec{\alpha})$ admits the decompositions

$$f(\vec{\alpha}) = \int_{\mathcal{T}_j} \tilde{c}^j(\vec{t})e^{-i\vec{\alpha}\cdot\vec{t}}d\vec{t},$$

for $j = 1, ..., d$. Here $\mathcal{T}_j = \{\vec{t} \in \mathbb{R}^d, t_i \geq T\delta_{ij}, \sum_i t_i = T' + T\}$.

Now, express the vector $\vec{\alpha} \equiv \vec{\beta} + i\vec{\gamma}$ in terms of its real and imaginary parts $\vec{\beta}, \vec{\gamma}$. Fixing $\vec{\gamma}$, we have that the above expressions depend on $\vec{\beta}$ as

$$f(\vec{\alpha}) = \int_{\mathcal{T}_j} \tilde{c}^j(\vec{t}, \vec{\gamma})e^{-i\vec{\beta}\cdot\vec{t}}d\vec{t}.$$

for some new coefficients $\tilde{c}$. In particular, for all $k, j \in \{1, ..., d\}$,

$$\int_{\mathcal{T}_j} \tilde{c}^j(\vec{t}, \vec{\gamma})e^{-i\vec{\beta}\cdot\vec{t}}d\vec{t} = \int_{\mathcal{T}_k} \tilde{c}^k(\vec{t}, \vec{\gamma})e^{-i\vec{\beta}\cdot\vec{t}}d\vec{t}.$$

This holds for all $\vec{\beta} \in \mathbb{R}^d$. Multiplying the above expression by $e^{i\vec{\beta}\cdot\vec{t}_0}$ for some $\vec{t}_0$ and integrating with respect to $\vec{\beta}$ one can see that $\tilde{c}^j(\vec{t}_0, \vec{\gamma})$ must vanish if $\vec{t}_0 \in \mathcal{T}_j \setminus \mathcal{T}_k$. Therefore, we may write

$$f(\vec{\alpha}) = \int_{\cap_j \mathcal{T}_j} \tilde{c}(\vec{t}, \vec{\gamma})e^{-i\vec{\beta}\cdot\vec{t}}d\vec{t}.$$

Any $\vec{t} \in \cap_j \mathcal{T}_j$ satisfies $t_i \geq T$ and $\sum_{i=1}^d t_i = T + T'$. Combining these two expressions, we conclude that, for all $\vec{\gamma}$ for which $f(\vec{\alpha})$ does not vanish, $dT \leq \sum_{i=1}^d t_i = T + T'$. That is,

$$(d-1)|T_1| \leq T'. \tag{95}$$

Note that the argument above does not invoke at any point the uncontrollability of system 1: it holds even if we know the form of the operators $\{\Pi_i\}_i$ in Equation (91). In fact, it holds if we further know the similarity transformation that diagonalizes H_0.

We can now build from this result to the general scenario. Suppose that, through a scattering experiment of duration T', we were able to induce a transformation of the type $A \propto U(T_1, ..., T_n) := \bigotimes_{j=1}^n e^{-iH_0T_j}$, for some times $T_1, ..., T_n$. Let us assume, w.l.o.g., that $T_1, ..., T_k < 0$, and $T_{k+1}, ..., T_n \geq 0$.

Now, let B be the similarity transformation that diagonalizes H_0, i.e., $B^{-1}H_0B = \sum_i \alpha_i |i\rangle\langle i|$, and consider the operator

$$\tilde{A} := \prod_{j=1}^{d-1} \left(B^{\otimes n}\Gamma^j (B^{-1})^{\otimes n}\right) A \left(B^{\otimes n}\Gamma^{-j}(B^{-1})^{\otimes n}\right),$$

with $\Gamma = \mathbb{1}^{\otimes k} \otimes \tilde{\Gamma}^{\otimes n-k}$, where $\tilde{\Gamma} := \sum_{i=1}^{d} |i\rangle \langle i \oplus 1|$ and $\oplus$ is addition modulo d. Noting that, for any $d \times d$ diagonal matrix Y, $\prod_{j=1}^{d-1} \tilde{\Gamma}^j Y \tilde{\Gamma}^{-j} = \det(Y) Y^{-1}$, we have that $\tilde{A} \propto U((d-1)T_1, ..., (d-1)T_k, -T_{k+1}, ..., -T_n)$.

Finally, define the linear map $\Lambda : M_d^{\otimes n} \to M_d$ by

$$\Lambda(X) := \sum_{i_1,\ldots,i_{n-1}} (\mathbb{1} \otimes \langle i_1| \otimes \cdots \otimes \langle i_{n-1}|) X (|i_1\rangle \otimes \cdots \otimes |i_{n-1}\rangle \otimes \mathbb{1}).$$

This map implements the linear extension of the operation

$$\Lambda(X_1 \otimes \cdots \otimes X_n) = X_1 \cdots X_n,$$

so it follows that

$$\Lambda(\tilde{A}) \propto U\left((d-1)\sum_{i=1}^{k} T_i - \sum_{i=k+1}^{n} T_i\right),$$

i.e., $\Lambda(\tilde{A})$ is a rewinding transformation for 1 target system. Clearly, $\Lambda(\tilde{A})$ can be expressed as linear combinations of product operators of the form Equation (91) for $n = 1$, with the particularity that $\sum_i t_i = (d-1)nT'$. By Equation (95) we have, then, that the total rewinding time $(d-1)\sum_{i=1}^{k} |T_i| + \sum_{i=k+1}^{n} |T_i|$ is upper bounded by nT'. That is

$$\sum_{i:T_i>0} T_i + \sum_{i:T_i<0} |T_i|(d-1) \leq nT'$$

as stated in Equation (87). □

4.3. Different protocols implementing the same polynomials. The goal of this section is to show by means of a concrete examples all the steps involved in designing a scattering protocol from a given polynomial.

Let $R(V, W) = [W, V]V^s[W, V]$ be the polynomial we want to implement. We know that this effects the transformation V^{-s} on qubits (see Eq.(6)). We will show three different ways to propagate targte system 1 by $R(V, W)$. First, we expand the terms to write

$$R = WVV^sWV - VWV^sWV - WVV^sVW + VWV^sVW. \tag{96}$$

By the correspondence given in Section B, we know we can implement this with a canonical scattering protocol that sends $s + 4$ one-dimensional probes.

This corresponds to sequentially preparing $s + 4$ probes in the state

$$\frac{1}{\sqrt{2}}\left(|0\rangle_R |\varphi\rangle_r + |1\rangle_R |\Phi\rangle_r\right) \tag{97}$$

(we assume that $d_p = 1$, so we do not bother introducing system r). Then, conditioned on all such probes having returned to the lab, we post-select the lab register to the many-body state

$$\frac{1}{\sqrt{4}}(|01\rangle|0\rangle^{\otimes s}|01\rangle - |01\rangle|0\rangle^{\otimes s}|10\rangle - |10\rangle|0\rangle^{\otimes s}|01\rangle + |10\rangle|0\rangle^{\otimes s}|10\rangle)|\phi\rangle_p . \tag{98}$$

This is very inefficient, as by inspection we see that for $s\Delta t$ units of time we could just let the system evolve naturally. So the second implementation uses only 4 probes: sending the s probes of the intermediate steps will only decrease the probability of success because we need to post-select to the state of not having sent them.

Lastly, it is possible to implement the same polynomial with only 2 probes. Instead of using the initial state of a probe as a superposition of sending it and not sending it, we put it in a superposition of sending it at time 0 and at time Δt. We again label these with a qubit register $|0\rangle_R$ and $|1\rangle_R$, respectively. After $2\Delta t$ time units, the state of the joint system target-register will be

$$\frac{1}{\sqrt{2}}(VW|\psi_1\rangle|0\rangle_R + WV|\psi_1\rangle|1\rangle_R) \tag{99}$$

where we ignore the internal state of the probe. If we do this once, let the system evolve for $s\Delta t$ time units, and then send another probe in the same initial state, we get polynomial R after post-selecting the lab register on state

$$(\frac{1}{\sqrt{2}}(|0\rangle_R - |1\rangle_R))^{\otimes 2}. \tag{100}$$

We now move on to the implementation of a protocol that involves two different target systems, specifically, one that transfers time. A family of polynomials to do this is given by Equation (86):

$$E(V,W) := (V^s \otimes \mathbb{1})\Omega(V,W)(\mathbb{1} \otimes V^s)\Omega(V,W), \tag{101}$$

where Ω is any polynomial which is proportional to a `SWAP`. For example, we can use the polynomial given by Equation (83). The time warping polynomial we obtain this way is very long, but expanding it we get

$$E(V,W) = V^sVWWVWVWWVW \otimes VWWVWV^sVWWVW + \cdots \tag{102}$$

A possible way to obtain from E(V,W) a time translation protocol would be to use the correspondence of Section 2.1, to send $10 + s$ one-dimensional probes to each system. This is done by sequentially preparing each of the $s + 10$ pairs of probes in the state

$$\frac{1}{\sqrt{2}}\big(|0\rangle_R|\varphi_1\rangle_r + |1\rangle_R|\Phi\rangle_r\big) \otimes \frac{1}{\sqrt{2}}\big(|0\rangle_R|\varphi_2\rangle_r + |1\rangle_R|\Phi\rangle_r\big) \tag{103}$$

where $|\varphi_1\rangle, |\varphi_2\rangle$ are the states required to make a probe interact only with target system 1 and 2, respectively. We remind the reader that such states exist by virtue of our working assumptions.

Conditioned on the return of all the probes, we post-select the lab register on the state

$$\frac{1}{40}(|1011010110\rangle|0\rangle^{\otimes s} \otimes |10110\rangle|0\rangle^{\otimes s}|10110\rangle + \cdots \tag{104}$$

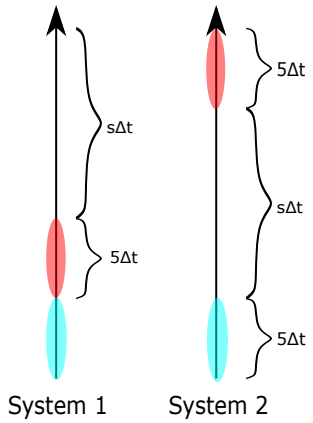

Figure 4. A schematic representation of a fast-forwarding protocol. The vertical arrows represent unperturbed evolution of the target systems. Blue and red regions correspond, respectively, to times in which we are sending probes to apply a protocol that performs a SWAP gate.

However, there is a better way to do this, as in the previous example, with only 10 probes for each system. To do this we consider implementing a single SWAP, we do this by preparing 5 pairs of probes in the state (103) and post selecting the lab register on the (normalized) state that encodes the polynomial. For example, for the polynomial $\tilde{\Omega}$ this state would be

$$|\phi_{\tilde{\Omega}}\rangle = \frac{1}{\sqrt{40}}\big(|10110\rangle_R \otimes |10110\rangle_R - |10110\rangle_R \otimes |10011\rangle_R - ... \tag{105}$$

Our protocol is schematically shown in Figure 4. First we sequentially prepare and release 5 probe pairs in state (103) to systems $1, 2$ at times $\{0, \Delta t, ..., 4\Delta t\}$. Then, while we let system 2 evolve naturally for s time units, we sequentially release five more probes in state $\frac{1}{\sqrt{2}}(|0\rangle_R |\varphi_1\rangle_r + |1\rangle_R |\Phi\rangle_r)\big)$ to interact with system 1, at times $\{5\Delta t, ..., 9\Delta t\}$. After that, we let system 1 evolve naturally for s time units. Meanwhile, at times $\{5\Delta t + s\Delta t, ..., 9\Delta t + s\Delta t\}$, we sequentially prepare five more probes in state $\frac{1}{\sqrt{2}}(|0\rangle_R |\varphi_2\rangle_r + |1\rangle_R |\Phi\rangle_r)\big)$; those will interact with system 2. After time $10\Delta t + s\Delta t$ has elapsed, we postselect the registers in the lab to the state $|\phi_\Omega\rangle \otimes |\phi_\Omega\rangle$.

The final state of the system after the post-selection, and keeping track of the normalization of the states, is now

$$\frac{1}{40}\frac{1}{2^{10}} E(V, W) |\psi_{1,2}\rangle = \lambda V^{2s} \otimes \mathbb{1} |\psi_{1,2}\rangle , \tag{106}$$

where λ is a scalar that depends on V and W. That is, by following this protocol, we will have made system 1 evolve by $T_1 = 2s\Delta$ time units in time $s\Delta t + 10\Delta t$, and the probability of success of the protocol is $|\lambda|^2$.

Using the SWAP polynomial given in Equation (83) to construct $|\phi_{\tilde{\Omega}}\rangle$ and averaging over V and W according to the Haar measure, we get an average success probability of approximately 2×10^{-4}. While this may be too small to be of practical importance, it is still large enough to be detectable experimentally. Furthermore, it may be possible to boost this probability significantly by increasing the number of probes used to implement the SWAP. It would also be possible to increase the probability of success by post-selecting

on a larger dimensional space that is spanned by multiple vectors encoding different SWAP polynomials.

5. High probability rewinding for qubits

As we have seen, the probability of successfully time warping a system can become extremely low. A natural question arises, can we increase the probability of success of our protocols? If one takes a look at the protocol given in [QDS+19a; QDS+19b], it is immediate to notice that, although it is probabilistic, a failure can be corrected. In the following, we show that a similar scheme can work in the uncontrolled setting for a certain rewinding protocol on qubits.

5.1. The protocol. Given a time unit $\Delta t > 0$, we introduce a universal physical process that rewinds any two-level quantum system by any amount $T = s\Delta t$, where s is an arbitrary natural number. This process, acting on a target system with free Hamiltonian H_0, will propagate the target's initial quantum state by $e^{iH_0 s\Delta t}$, thus leaving the target on the state it had $s\Delta t$ time units before the experiment started.

The basic building block of the protocol is the gate Q depicted in Figure 5 (a). Denoting by $|\psi\rangle$ the state of the target, this gate performs the transformation

$$Q\,|\psi\rangle\,|{\rightarrow}\rangle \propto [V,W]\,|\psi\rangle\,|{\uparrow}\rangle + \{V,W\}\,|\psi\rangle\,|{\rightarrow}\rangle\,. \tag{107}$$

Here $W := e^{-iH_0\Delta t}$ and V denotes an unknown linear map, detailed below. The kets $|{\rightarrow}\rangle, |{\uparrow}\rangle$ respectively label a left-to-right and a bottom-up trajectory of the target system, as seen in Figure 5 when the letter Q is upright. If, right after implementing Q, we measure the target's motion degree of freedom in the $\{|{\uparrow}\rangle, |{\rightarrow}\rangle\}$ basis, the target will be propagated by either $\{V,W\}$ or $[V,W]$, depending on the measurement result.

We next dedicate some lines to explain how to universally realize the gate Q, also known in the literature as SWITCH [CDPV13], for some uncharacterized matrix V. Let O be a repeatable physical operation (e.g.: switching on a magnetic field, releasing an electron) of duration τ, whose effect on the target's internal degree of freedom is to propagate its ket by some unknown operator V. This can be achieved by, e.g., switching on an interacting Hamiltonian for some time τ, or, as described in Section 2, by making the target unitarily interact with a probe, which is post-selected onto a given pure state after the interaction. Note that V will be a unitary matrix in the first case and non-unitary in the second.

Given the ability to conduct any such operation O, one can implement the gate Q by playing with the motion degree of freedom of the target: it suffices to put the latter in an equal superposition of two paths. In the first path, the target is allowed to evolve freely for time Δt and then we act on it with O for time τ. In the second path, we first act on the target with O for time τ and then we let it evolve freely for time Δt. The state of the target at this stage will thus be proportional to $VW\,|\psi\rangle\,|\gamma_1\rangle + WV\,|\psi\rangle\,|\gamma_2\rangle$, where γ_1, γ_2 denote the two trajectories. Next we make the two trajectories interfere, by conducting the unitary operation

$$|\gamma_1\rangle \mapsto \frac{1}{\sqrt{2}}\left(|{\rightarrow}\rangle - |{\uparrow}\rangle\right), \qquad |\gamma_2\rangle \mapsto \frac{1}{\sqrt{2}}\left(|{\rightarrow}\rangle + |{\uparrow}\rangle\right)$$

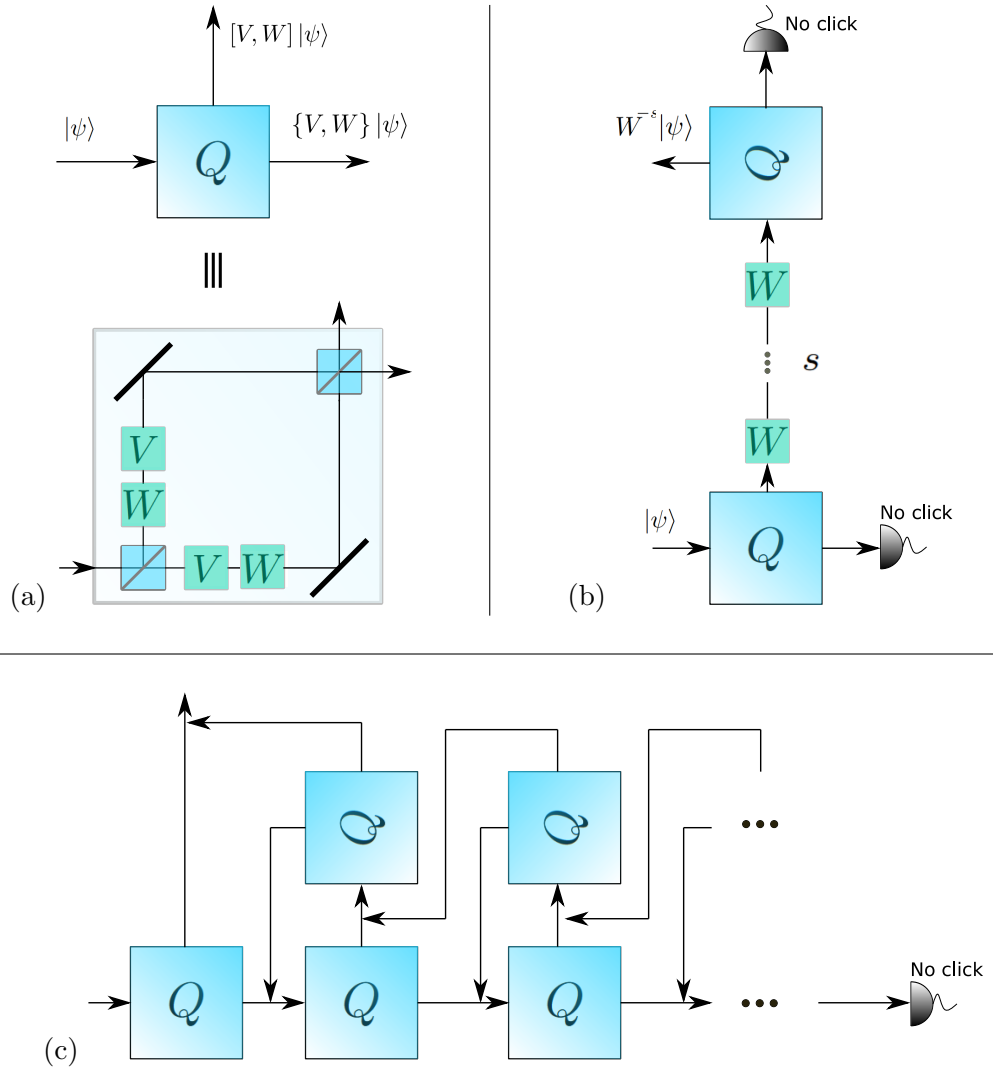

Figure 5. (a) The gate Q specified by Eq.(107). This quantum operation, the building block of the whole protocol, can be implemented in different ways depending on the physical systems under consideration. The figure shows a possible way to implement Q on photons with an interferometer. (b) A schematic of the full protocol without corrections. (c) A schematic of a level n correction. This figure substitutes each Q in part (b). Following any path in the figure that ends up in the top left outputs a state proportional to $[V,W]\,|\psi\rangle$.

(in optical systems, this can be achieved with a balanced beam splitter), arriving at equation (107).

The gist of the protocol is to apply the gate Q over and over to the target until it reaches a state proportional to $W^{-s}\,|\psi\rangle = e^{iH_0 s\Delta t}\,|\psi\rangle$. To achieve this goal, we rely on three general properties of 2×2 matrices.

Lemma 5.1.1. *Let V,W be 2×2 matrices and $n\in\mathbb{N}_0$. Then,* $\mathrm{tr}([V,W]\{V,W\}^n)=0$.

Proof. By the Caley-Hamilton theorem, for $n \geq 2$, the 2×2 matrix $\{V, W\}^n$ is a linear combination of $\mathbb{1}, \{V, W\}$. Hence, it is enough to show that the lemma holds for $n = 0, 1$, and this is a simple consequence of the cyclicity of the trace. □

Proposition 5.1.2. *Let V, W be arbitrary 2×2 matrices, and define $x \equiv [V, W], y \equiv \{V, W\}$. Then we have that:*

(a) $x^2 \propto \mathbb{1}_2$.

(b) If W is invertible, then, for any natural number s, $xW^s x \propto W^{-s}$.

(c) For any natural number n, $y^n x y^n \propto x$.

Proof. Now, note that any 2×2 traceless matrix can be written as a linear combination of the Pauli matrices $\sigma_X, \sigma_Y, \sigma_Z$, and thus its square is proportional to the identity matrix. This applies to the commutator $[V, W]$, the matrix polynomial $W^s[V, W]$ and, by Lemma 5.1.1, to $[V, W]\{V, W\}^n$. We thus have that, for all 2×2 matrices,

$$\begin{aligned} &[V, W]^2 \propto \mathbb{1}, \\ &W^s[V, W]W^s[V, W] \propto \mathbb{1}, \\ &[V, W]\{V, W\}^n[V, W]\{V, W\}^n \propto \mathbb{1}. \end{aligned} \tag{108}$$

If W is invertible, then we can multiply the second expression by W^{-s} on the left and arrive at the identity

$$[V, W]W^s[V, W] \propto W^{-s}. \tag{109}$$

Similarly, multiplying the third line of equation (108) by $[V, W]$ on the left and invoking the first line, we arrive at

$$\{V, W\}^n[V, W]\{V, W\}^n \propto [V, W]. \tag{110}$$

Note that the last step is only rigorous if the proportionality factor in the expression $[V, W]^2 \propto \mathbb{1}$ is non-zero. As it turns out, by the Cayley-Hamilton theorem, this factor is $-\det([V, W])$. Let us then prove that the relation also holds for $\det([V, W]) = 0$.

Define the matrices $x \equiv [V, W]$, $y \equiv \{V, W\}$ and $z \equiv y^n x y^n$. The matrices x and z have in this case rank at most 1. Since both x, z have zero trace (z, by virtue of Lemma 5.1.1), it follows that one can write them as $x = \lambda \left|\phi\right\rangle\!\left\langle\phi^\perp\right|$, $z = \nu \left|\varphi\right\rangle\!\left\langle\varphi^\perp\right|$, where $\langle\phi|\phi^\perp\rangle = \langle\varphi|\varphi^\perp\rangle = 0$. Now, by the third line of equation (108), $xz \propto \mathbb{1}$. Since the left-hand side of the relation has rank at most one, it follows that $xz = 0$. This is only possible if $\nu = 0$ (note that $\lambda = 0$ implies $\nu = 0$), in which case $z = 0 \propto x$; or if $\lambda, \nu \neq 0$ and $\langle\phi^\perp|\varphi\rangle = 0$, from which $z \propto x$. In either case, relation (110) holds.

The first line of equations (108), (109) and (110) are, respectively, the 2×2 matrix relations (a), (b), (c) claimed to hold in Proposition 5.1.2. This finishes the proof.

□

We remark that the proportionality factors on the right-hand sides of equations (a)-(c) are functions of the entries of the matrices V, W, and might vanish for some values of V, W.

Proposition 5.1.2 suggests a simple method to bring the target system to state $W^{-s}|\psi\rangle$ through consecutive uses of gate Q. First, we aim to effect the transformation $|\psi\rangle \to x|\psi\rangle$. Once there, all we have to do is wait for time $s\Delta t$ and manage to enforce the transformation x once more. The final state will then be $xW^s x|\psi\rangle$, that, by relation (b), is proportional to the state $W^{-s}|\psi\rangle = e^{iH_0 s\Delta t}|\psi\rangle$. In that case, the target will have been translated by $-s\Delta t$ time units.

The shortest way to rewind the system hence requires two applications of gate Q, see Figure 5 (b). Provided that, after measuring in the $\{|\uparrow\rangle, |\rightarrow\rangle\}$ basis, the target system emerges from gate Q through its vertical output port, the system will have been acted upon by x. Next, we wait for time s and then we input the system in Q again. If, once more, the target exits the gate through its vertical port, then we can guarantee that the rewinding process took place.

It could happen, though, that the target exits the first gate through its horizontal port. In that case, the system will be propagated by y instead of x. To proceed with the rewinding protocol, we must eliminate this operator. A possible path out is given by taking $n = 1$ in relation (c), namely, by the identity $yxy \propto x$. It follows that, if we make the system pass through two more Q gates and it exits the first one through the vertical port; and the second one, through the horizontal port, the system will end up in a state proportional to $x|\psi\rangle$. The situation is thus the same as if the target had exited through the vertical port in the original Q gate, see Figure 5 (c). Hence we can wait for $s\Delta t$ time units before trying to effect another transformation x on the system.

By virtue of relations (a)-(c) in Proposition 5.1.2, whichever sequence of ports the system happens to exit will propagate the target by an operator of the form xy^n or y^n. In the first case, n consecutive exits through the horizontal port of gate Q will propagate the system by x. In the second one, a vertical detection, followed by n consecutive horizontal ones, will have the same effect. Hence, no matter how advanced the protocol is, there always exists a chance of bringing the target to the terminal configuration $xW^s x|\psi\rangle$, as sketched in Figure 5 (c). Note that relations (a)-(c) hold even if the matrices V, W are not unitary. The protocol can thus be used, e.g., to rewind a two-level system undergoing a continuous decay governed by a non-Hermitian Hamiltonian, such as a neutral kaon [Kab68]. In this case, however, the terminal configuration is not ensured.

Notice as well that, should we enforce any limit m on the number of times that gate Q can be applied, the running time of the protocol would be upper bounded by $T' = m(\Delta t + \tau) + s\Delta t$ (recall that τ is the time it takes to implement the operation O). On the other hand, the protocol, if successful, would rewind the target system by an amount $T = s\Delta t$. Hence $T' = T + O(1)$ and, by the results of Sections 2 and 3, this implies that such an 'm-trimmed' universal rewinding protocol runs on asymptotically minimal time in the limit $T \to \infty$.

It remains to be seen how likely it is that the (trimmed or untrimmed) protocol succeeds. In principle it could be that, even allowing an unlimited number of uses of gate Q, the system never reaches a terminal configuration $xW^s x|\psi\rangle$. In this regard, note that, if the physical operation O has no effect whatsoever on the target (namely, if

$V = e^{-iH_0\tau}$), then the latter will keep evolving unperturbed, no matter how many times we act on it with the Q gate. More generally, one can see that the rewinding protocol will fail with certainty whenever $[V, W] = 0$.

The condition $[V, W] = 0$, violated by generic interactions V, W, requires a high degree of fine-tuning if the experimental setup is capable of perturbing the target system at all. One therefore wonders what the chances of success are when $[V, W] \neq 0$.

Using techniques from probability theory [Wil91], we prove in Section 5.2 that, provided that V, W are unitary and $[V, W] \neq 0$, the target will reach the pattern $xW^s x\,|\psi\rangle$ after a random finite number of uses of Q with probability 1. Moreover, we show how to compute the probability of success of an m-trimmed protocol given a lower bound on $\|[V, W]\|$. See Figure 6.

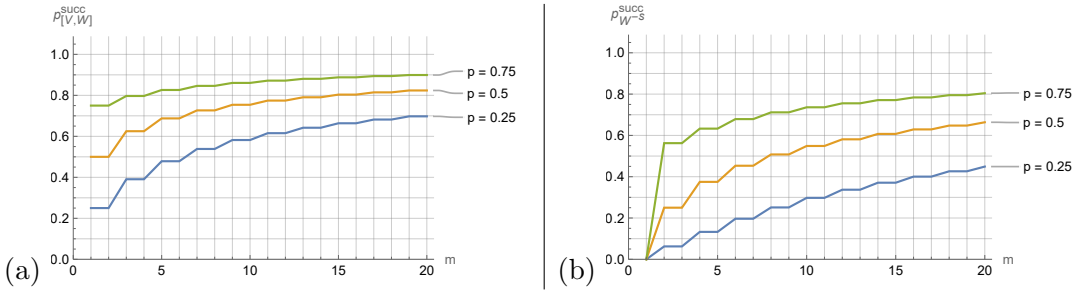

Figure 6. The protocol's probability of success p_A^{succ} of implementing operation A (vertical axis) as a function of the number m of uses of gate Q (horizontal axis), for different values of the probability p of exiting the Q gate through the vertical port (as proven in Section 5.2, p depends on V, W, but not on the state of the target). (a) The probability of successfully implementing $[V, W]$ (b) The probability of successfully rewinding the system for the full adaptive protocol. Note that this is not the square of (a).

5.2. Proofs. We begin by defining and proving the validity of a completely abstract protocol, without considering any implementation details. In this protocol, the system of interest is $\mathbb{C}^2$, and the other system acts as an ancilla that heralds the success of the protocol. We need to have control over the ancillary system.

Let V, W be 2×2 matrices. We define a gate Q acting on $\mathbb{C}^2 \otimes \ell^2(\mathbb{Z} \times \{-1, 1\})$ as

$$Q\,|\psi\rangle\,|n, z\rangle = \frac{1}{2}[V, W]\,|\psi\rangle\,|n, -z\rangle + \frac{1}{2}\{V, W\}\,|\psi\rangle\,|n - z, z\rangle\,.$$

We now show that Protocol 5.2.1 does what we claim. This will require several intermediate results. We begin by showing that if the protocol terminates, the output is indeed the rewound state that we are after. In fact, we show something slightly more general. Let us call, as in the previous section $x \equiv [V, W]$ and $y \equiv \{V, W\}$.

Lemma 5.2.2. *In Protocol 5.2.1, the state in line* 6*, after having obtained outcome* (n, z) *is*

$$|\psi\rangle \propto \theta(n) x^{\frac{z+1}{2}} y^n\,|\psi_0\rangle + (1 - \theta(n)) x^{-\frac{z-1}{2}} y^{-n} W^s x\,|\psi_0\rangle\,,$$

where $\theta(n) = 1$ *if* $n > 0$ *and* 0 *otherwise. In particular, the output of Protocol 5.2.1 is proportional to* $xW^s x\,|\psi_0\rangle$*, and thus equal to* $W^{-s}\,|\psi_0\rangle$*.*

Protocol 5.2.1 An adaptive protocol to rewind qubits

Input: Initial target state $|\psi_0\rangle \in \mathbb{C}^2$. Matrices $V, W \in M_2(\mathbb{C})$ such that $[V, W] \neq 0$, and W is invertible, $s \in \mathbb{N}$.

1: Initialize the target $|\psi\rangle \leftarrow |\psi_0\rangle$.
2: Initialize the ancilla $|\varphi\rangle \leftarrow |0, -1\rangle$.
3: Apply Q to the combined state $|\psi\rangle |\varphi\rangle$.
4: Measure the ancilla in the basis $\{|m, w\rangle\}_{m\in\mathbb{Z}, w\in\{-1,1\}}$.
5: $(n, z) \leftarrow$ Outcome of measurement
6: $|\psi\rangle \leftarrow$ Target state after measurement.
7: **if** $(n, z) = (0, 1)$ **then**
8: Apply W^s to $|\psi\rangle$.
9: $|\varphi\rangle \leftarrow |0, 1\rangle$.
10: **else if** $(n, z) = (0, -1)$ **then**
11: **return** $|\psi\rangle$.
12: **else**
13: $|\varphi\rangle \leftarrow |n, z\rangle$.
14: **end if**
15: **goto** 3.

Proof. Recall the relations shown in Proposition 1 of the main text: $x^2 \propto \mathbb{1}$, $xW^s x \propto W^{-s}$, $y^n x y^n \propto x$. These give rise to a word problem that we now proceed to solve. Note also that, it is only possible to get a measurement outcome (n, z) with $n > 0$ before obtaining the measurement outcome $(0, 1)$ and from that moment onwards it is only possible to get $n < 0$ until the protocol terminates. Therefore, the problem is naturally divided in the cases $n > 0$ and $n < 0$. Let us proof the result in the case $n > 0$, the other one being completely symmetric, but using $W^s x |\psi_0\rangle$ as the initial state.

We proceed by induction in the amount of times m we have reached line 6. For $m = 1$, the outcome obtained is either $(1, -1)$ or $(0, 1)$, and the target state updates as we say in line 6. Suppose now that we have obtained outcome (n, z) with $n > 0$ in loop iteration $m > 1$. This means that in loop $m - 1$ we had obtained either outcome $(n + z, z)$ or outcome $(n, -z)$, which also have $n > 0$. Then the update rule says that, in the first case, we have by induction and the simplification rules that

$$y x^{\frac{z+1}{2}} y^{n+z} |\psi_0\rangle \propto \begin{cases} y^n |\psi_0\rangle & \text{if } z = -1, \\ x y^n |\psi_0\rangle & \text{if } z = 1. \end{cases}$$

In the second case, from the update rule and the simplification rules, we have

$$x x^{\frac{-z+1}{2}} y^n |\psi_0\rangle \propto \begin{cases} y^n & \text{if } z = -1, \\ x y^n & \text{if } z = 1, \end{cases}$$

as required. □

Note that after each loop the ancillary system is always in a particular state $|n, z\rangle$ and so the measurement is always a two outcome measurement. In fact, we can always implement this protocol with a qubit and something to keep track of the measurement outcomes. We now show that the probability of either outcome is always independent of the state of the system.

Lemma 5.2.3. *Let V, W be 2×2 unitary matrices. Then, the probability*

$$p = \langle n, -z| \, Q \, |\psi\rangle \, |n, z\rangle$$

is independent of $|\psi\rangle$.

Proof. We apply the Cayley-Hamilton theorem to the 2×2 matrix $VWV^\dagger W^\dagger$, obtaining

$$(VWV^\dagger W^\dagger)^2 - \mathrm{tr}(VWV^\dagger W^\dagger) VWV^\dagger W^\dagger + \det\Big(VWV^\dagger W^\dagger\Big) = 0, \tag{111}$$

The last determinant equals 1. Thus, Multiplying the above expression by WV on the left and by $W^\dagger V^\dagger$ on the right, we find that

$$V^\dagger W^\dagger VW + W^\dagger V^\dagger WV = \mathrm{tr}(VWV^\dagger W^\dagger)\mathbb{1}. \tag{112}$$

Now, the probability p of measuring the ancillary system and obtaining result (n, z) is

$$\frac{1}{4} \langle \psi| \, [V, W]^\dagger [V, W] \, |\psi\rangle = \frac{1}{4} \langle \psi| \, 2\mathbb{1} - V^\dagger W^\dagger VW - W^\dagger V^\dagger WV \, |\psi\rangle \tag{113}$$

By Equation (112), the last expression just depends on the invariant $\mathrm{tr}\big(VWV^\dagger W^\dagger\big)$ and not on the state itself. □

This independence on the state and the fact that the state only depends on the outcome measuremt allows us to model the evolution of the target state $|\psi\rangle$, at each loop of the protocol, as a classical particle undergoing a random walk in the directed graph shown in Figure 7.

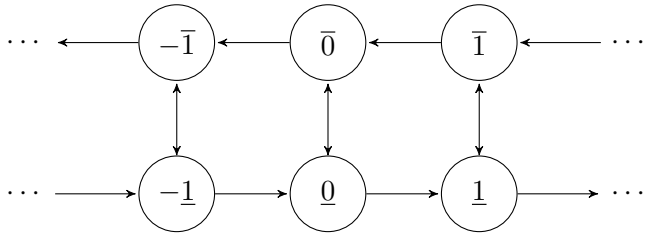

Figure 7. A random walk modelling the word problem. In this directed graph we consider the following situation. We label $(n, -1) \equiv \underline{n}$ and $(n, 1) \equiv \overline{n}$. Thus, starting in position $\underline{0}$ at $t = 0$, at each loop of the protocol the classical particle moves along a vertical edge with probability p and along a horizontal one with probability $1 - p$. The goal is to get back to $\underline{0}$ at a positive finite time. A move along a horizontal edge corresponds to the operation $|\psi\rangle \mapsto \{V, W\} \, |\psi\rangle$ on the target system, and a move along a vertical edge to the operation $|\psi\rangle \mapsto [V, W] \, |\psi\rangle$.

After each loop of the protocol, the particle can move in the horizontal direction with probability $1 - p$: this corresponds to measuring the quantum target leaving a Q gate and obtaining outcome $(n - z, z)$, hence propagating the current quantum state by the operator $\{V, W\}$. Alternatively, with probability p, the classical particle will move in the vertical direction of the graph. This corresponds to measuring the target leaving the Q gate and obtaining outcome $(n, -z)$, which propagates its state by $[V, W]$.

If the initial position of the classical particle is $\underline{0}$, then, by the time the particle reaches $\overline{0}$, the quantum target system will have been propagated by $[V, W]$. This is independent of the graph path taken by the classical particle, by virtue of Lemma 5.2.2.

Once the classical particle is in $\overline{0}$, we would stop the random walk momentarily and let the quantum target system evolve freely for $s\Delta t$ time units. Then we would act again with the Q gate on the system, thus continuing the random walk until the classical particle arrives at $\underline{0}$, at which point the target has been propagated by $xW^s x$ and hence it would have been rewound.

We will next prove that the probability that the classical particle comes back during its random walk to $\underline{0}$ in finite time is 1. For this, we will use standard notation and techniques which can be found for example, in [Wil91]. This proves that Protocol 5.2.1 terminates in a finite amount of time with probability 1. Since the waiting time does not affect the probability of success of the protocol, we set $s = 0$ for the remainder of the discussion.

Let $\{S_n\}_{n\geq 0}$ be the sequence of random variables which describe this random walk when starting in the state S_0. We define the hitting time of node a from state b as the random variable

$$T_{b\to a} := \inf\{n > 0 \mid S_n = a, S_0 = b\},$$

which takes values in $\mathbb{N}\cup\{+\infty\}$. The probability of successfully finishing the protocol in m steps is therefore given by $\mathrm{P}(T_{\underline{0}\to\underline{0}} = m)$. To calculate this probability we will exploit the symmetries of the graph. All basic concepts about random walks that we use in the proof can be found, for example, in [Wil91].

Theorem 5.2.4. *For all natural $m \geq 1$ and all $p \in [0, 1]$, with the convention that $0^0 = 1$,*

$$\mathrm{P}\left(T_{\underline{0}\to\overline{0}} = 2m-1\right) = \sum_{n=1}^{m}(-1)^{n+1}\binom{1/2}{n}\binom{1-2n}{m-n}(2p)^{2n-1}(2p-1)^{m-n},$$

$$\mathrm{P}\left(T_{\underline{0}\to\overline{0}} = 2m\right) = 0.$$

Remark 5.2.5. The use of $0^0 = 1$ in Theorem 5.2.4 is justified in two ways. First, when $p \neq 1/2$ there is no indeterminacy in the formula, which is a continuous function of p. Taking the limit $p \to 1/2$ we obtain

$$\mathrm{P}\left(T_{\underline{0}\to\overline{0}} = 2m-1\right) = (-1)^{m+1}\binom{1/2}{m}, \tag{114}$$

so regarding 0^0 as 1 is the natural choice to make the probability continuous on p. Furthermore, equation (114) is also the probability for the hitting time of the origin in a simple symmetric random walk of the integers starting at the origin. Closer inspection of our graph reveals that indeed, when $p = 1/2$ (and only in this case), these two processes are equivalent for the purposes of this random variable. This is therefore the correct formula for $p = 1/2$.

We proceed to prove Theorem 5.2.4.

Proof. We define the generating function

$$f(\alpha) := \mathbb{E}\left[\alpha^{T_{\underline{0}\to\overline{0}}}\right] = \sum_{n=1}^{\infty} \mathrm{P}\left(T_{\underline{0}\to\overline{0}} = n\right)\alpha^n.$$

Note that in principle this is only correctly defined for $\alpha < 1$, as the probability of having an infinite hitting time could be non-zero. The correctness of the last equality is justified a posteriori, when we will see that $\lim_{\alpha\to 1^-} f(\alpha) = 1$. For the time being, assume that $0 < \alpha < 1$. We have

$$\begin{aligned} f(\alpha) &= \mathbb{E}\left[\alpha^{T_{\underline{0}\to\overline{0}}} \mid S_1 = \overline{0}\right]\mathrm{P}\left(S_1 = \overline{0}\right) + \mathbb{E}\left[\alpha^{T_{\underline{0}\to\overline{0}}} \mid S_1 = \underline{1}\right]\mathrm{P}\left(S_1 = \underline{1}\right) \\ &= p\mathbb{E}\left[\alpha^{T_{\underline{0}\to\overline{0}}} \mid S_1 = \overline{0}\right] + (1-p)\,\mathbb{E}\left[\alpha^{T_{\underline{0}\to\overline{0}}} \mid S_1 = \underline{1}\right] \\ &= p\alpha + (1-p)\,\mathbb{E}\left[\alpha^{T_{\underline{0}\to\overline{0}}} \mid S_1 = \underline{1}\right]. \end{aligned}$$

However, the graph is invariant under horizontal translations, so

$$\mathrm{P}\left(T_{\underline{0}\to\overline{0}} = n \mid S_1 = \underline{1}\right) = \mathrm{P}\left(T_{\underline{1}\to\overline{0}} = n-1\right) = \mathrm{P}\left(T_{\underline{0}\to-\overline{1}} = n-1\right).$$

Therefore,

$$\mathbb{E}\left[\alpha^{T_{\underline{0}\to\overline{0}}} \mid S_1 = \underline{1}\right] = \sum_{n=2}^{\infty} \mathrm{P}\left(T_{\underline{1}\to\overline{0}} = n-1\right)\alpha^n = \alpha\mathbb{E}\left[\alpha^{T_{\underline{0}\to-\overline{1}}}\right].$$

We can divide the process of getting to $-\overline{1}$ from $\underline{0}$ in two parts: by going for the first time to $\overline{0}$ from $\underline{0}$ and then visiting $-\overline{1}$ from $\overline{0}$ also for the first time. The probabilities are decomposed as follows:

$$\begin{aligned} \mathrm{P}\left(T_{\underline{0}\to-\overline{1}} = n\right) &= \mathrm{P}\left(\inf\{k > T_{\underline{0}\to\overline{0}} \mid S_k = -\overline{1}, S_{T_{\underline{0}\to\overline{0}}} = \overline{0}, S_0 = \underline{0}\} = n\right) \\ &= \sum_{m<n} \mathrm{P}\left(T_{\underline{0}\to\overline{0}} = m\right)\mathrm{P}\left(\inf\{k > m \mid S_k = -\overline{1}, S_m = \overline{0}\} = n\right) \\ &= \sum_{m<n} \mathrm{P}\left(T_{\underline{0}\to\overline{0}} = m\right)\mathrm{P}\left(T_{\overline{0}\to-\overline{1}} = n-m\right) \\ &= \mathrm{P}\left(T_{\underline{0}\to\overline{0}} + T_{\overline{0}\to-\overline{1}} = n\right), \end{aligned}$$

where we have used the Markov property in the third step. Note that by the strong Markov property, $\inf\{k > T_{\underline{0}\to\overline{0}} \mid S_k = -\overline{1}, S_{T_{\underline{0}\to\overline{0}}} = \overline{0}\}$ (which has the same distribution as $T_{\overline{0}\to-\overline{1}}$) is independent of $T_{\underline{0}\to\overline{0}}$. In particular,

$$\mathbb{E}\left[\alpha^{T_{\underline{0}\to-\overline{1}}}\right] = \mathbb{E}\left[\alpha^{T_{\underline{0}\to\overline{0}}}\right]\mathbb{E}\left[\alpha^{T_{\overline{0}\to-\overline{1}}}\right] = \mathbb{E}\left[\alpha^{T_{\underline{0}\to\overline{0}}}\right]\mathbb{E}\left[\alpha^{T_{\underline{0}\to\underline{1}}}\right],$$

where the last equality follows from the reflection symmetry of the graph.

Repeating the arguments made at the beginning for $T_{\underline{0}\to\overline{0}}$ we get that

$$\mathbb{E}\left[\alpha^{T_{\underline{0}\to\underline{1}}}\right] = \alpha p\mathbb{E}\left[\alpha^{T_{\underline{0}\to-\overline{1}}}\right] + (1-p)\alpha.$$

Combining everything,

$$f(\alpha) = p\alpha + \frac{(1-p)^2\alpha^2 f(\alpha)}{1 - pf(\alpha)\alpha}$$

or the second degree equation

$$\alpha p f(\alpha)^2 + (\alpha^2 - 2p\alpha^2 - 1)f(\alpha) + p\alpha = 0,$$

which has the solutions

$$\frac{1 + 2p\alpha^2 - \alpha^2 \pm \sqrt{(1 + 2p\alpha^2 - \alpha^2)^2 - 4p^2\alpha^2}}{2p\alpha}.$$

The correct behaviour as $\alpha \to 0^+$ is obtained with the minus sign in front of the square root, so let us expand this one as a power series on α centered at zero:

$$\begin{aligned}
f(\alpha) &= \frac{1 + 2p\alpha^2 - \alpha^2 - \sqrt{(1 + 2p\alpha^2 - \alpha^2)^2 - 4p^2\alpha^2}}{2p\alpha} \\
&= \frac{-(1 + 2p\alpha^2 - \alpha^2)\sum_{n=1}^{\infty}(-1)^n\binom{1/2}{n}\left(\frac{2p\alpha}{1+2p\alpha^2-\alpha^2}\right)^{2n}}{2p\alpha} \\
&= \sum_{n=1}^{\infty}(-1)^{n+1}\binom{1/2}{n}\left(\frac{2p\alpha}{1 + 2p\alpha^2 - \alpha^2}\right)^{2n-1} \\
&= \sum_{n=1}^{\infty}(-1)^{n+1}\binom{1/2}{n}(2p)^{2n-1}\alpha^{2n-1}\left(\frac{1}{1 + (2p-1)\alpha^2}\right)^{2n-1} \\
&= \sum_{n=1}^{\infty}(-1)^{n+1}\binom{1/2}{n}(2p)^{2n-1}\alpha^{2n-1}\sum_{k=0}^{\infty}\binom{1-2n}{k}(2p-1)^k\alpha^{2k} \\
&= \sum_{m=1}^{\infty}\sum_{n+k=m}(-1)^{n+1}\binom{1/2}{n}\binom{1-2n}{k}(2p)^{2n-1}(2p-1)^k\alpha^{2n+2k-1} \\
&= \sum_{m=1}^{\infty}\sum_{n=1}^{m}(-1)^{n+1}\binom{1/2}{n}\binom{1-2n}{m-n}(2p)^{2n-1}(2p-1)^{m-n}\alpha^{2m-1},
\end{aligned}$$

from which the statement follows. Note that

$$\lim_{\alpha\to 1^-} f(\alpha) = \frac{2p - \sqrt{(2p)^2 - 4p^2}}{2p} = 1,$$

so that $\mathrm{P}\left(T_{\underline{0}\to\overline{0}} < +\infty\right) = 1$, like we had anticipated.

□

To get now the probability of successfully resetting at a particular time, we can just use the formula we just got and compute (as we did in the previous proof for $T_{\underline{0}\to\overline{-1}}$):

$$\mathrm{P}\left(T_{\underline{0}\to\underline{0}} = t\right) = \sum_{k+l=t}\mathrm{P}\left(T_{\underline{0}\to\overline{0}} = k\right)\mathrm{P}\left(T_{\underline{0}\to\overline{0}} = l\right).$$

The result is the following formula:

Corollary 5.2.5.1. *For all natural $m \geq 1$ and all $p \in [0,1]$, with the convention that $0^0 = 1$,*

$$\mathrm{P}\left(T_{\underline{0}\to\underline{0}} = 2m\right) = \sum_{k=1}^{m}\sum_{i=1}^{k}\sum_{j=1}^{m-k+1}(-1)^{i+j}\binom{1/2}{i}\binom{1/2}{j}\binom{1-2i}{k-i}\cdots$$
$$\cdots\binom{1-2j}{m-k+1-j}(2p)^{2(i+j-1)}(2p-1)^{m+1-(i+j)},$$
$$\mathrm{P}\left(T_{\underline{0}\to\underline{0}} = 2m-1\right) = 0.$$

Note in particular that $\mathrm{P}(T_{\underline{0}\to\underline{0}} < \infty) = 1$.

For completeness, we now combine all the statements we have proven.

Theorem 5.2.6. *Let H_0 and H_I be hermitian operators acting on $\mathbb{C}^2$, $|\psi_0\rangle \in \mathbb{C}^2$ an initial state, $T > 0$ an amount of time to rewind, and $V := e^{-iH_I\tau}, W := e^{-iH_0T/s}$ for some $s \in \mathbb{N}, \tau > 0$, such that $[V,W] \neq 0$.*

Then, for every $0 < q < 1$ there is a natural number $m \in \mathbb{N}$ such that Protocol 5.2.1 terminates after m iterations with probability bigger or equal to q, outputing the state $e^{iH_0T}|\psi_0\rangle$.

Proof. Indeed, from Lemmas 5.2.2 and 5.2.3 we deduce that the probability that the random walk described in Figure 7 with parameter $p = 1/2 - \mathrm{tr}\left(VWV^\dagger W^\dagger\right)$ starting at $\underline{0}$ goes back to $\underline{0}$ in m steps coincides with the probability that protocol 5.2.1 stops after a m iterations. This probability tends to one, since p is not zero due to the fact that $[V,W] \neq 0$. We can then use Corollary 5.2.5.1 to compute an m such that $\mathrm{P}(T_{\underline{0}\to\underline{0}} < \infty) \geq q$. □

6. Discussion

We have studied time-translations in the most general setting possible. That is, of an uncontrolled system with uncontrolled operations. This is a generalization to the uncontrolled setting of other work [QDS+19a; QDS+19b]. One wonders which results from the controlled world are generalizable to the uncontrolled setting. We have seen that all the possible time translations attainable in this setting are the ones that satisfy Theorem 4.0.1. The conditions in 4.0.1 are equivalent to the following postulates:

- Evolution time cannot be created.
- Evolution time can be transferred between two identical systems at no cost.
- Evolution time of a system of dimension d can be inverted at a cost $(d-1)$.
- Evolution time can be destroyed.

We are led to conclude that evolution time behaves like a sort of resource that we can use for our operations. Furthermore, our work has introduced tensor polynomials, garnering the interest of mathematicians working in PI-rings [Pro20]. Many of the problems that we want to solve, including generalizing our adaptive protocol for higher dimensional systems, and for SWAP polynomials, other kinds of time translations, etc. can be rewritten in this algebraic language. Even though this mapping makes it seem like answering these questions is hopeless, it also opens a new avenue to tackle such problems, which we will undertake in future work.

Some of these protocols have been implemented with optical systems in the lab [GXL+20; LYW+20; SST+23]. In particular, in [SST+23] we further show that our qubit protocol exhibits a quantum advantage over universal time translation protocols where the probes are classical systems.

Recently, the breakthrough paper [YSM22] has shown that there is, in the controlled setting for qubits, a deterministic protocol which inverts a unitary in finite time. This is something that can never be translated to the uncontrolled setting, which is by nature purely probabilistic and therefore our qubit results is equally optimal. Left to do is to see if this can also be achieved for higher dimensional systems, and for time translations other than rewindings.

Chapter 5

Discussion

In this thesis, we have learned that real Hilbert spaces are not enough to describe tripartite nonlocality experiments, analogous to how local hidden variable theories are not enough to describe Bell experiments. We have found a new quantum effect experienced by the simplest mechanical system of all: a particle moving freely in a line. We proved that this effect is as powerful as quantum backflow and that, in some sense, they are equivalent. This has helped us give new upper bounds on the strength of quantum backflow. Finally, we have studied time translations in a completely uncontrolled system. We have seen that many of the results that work in the controlled case also work in this setting, although by using much more complicated protocols.

We have discussed each result and its implications in more detail in the respective discussion Sections on each Chapter. Roughly, all of these are new effects which shed a bit more light about the role of space and time in quantum mechanics, and have already led to many interesting experiments [CWL+22; LMW+22; WJG+22; GXL+20; LYW+20; SST+23] and mathematical developments [Pro20; Pro22]. However, there is still much to be done.

Bibliography

[AAPV90] 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.*, vol. 64, no. 25, 1990 (cit. on pp. 85, 86).

[ABB+20] C. Arenz, D. I. Bondar, D. Burgarth, C. Cormick, and H. Rabitz, "Amplification of quadratic Hamiltonians," *Quantum*, vol. 4, 2020 (cit. on p. 87).

[ABW13] A. Aleksandrova, V. Borish, and W. K. Wootters, "Real-vector-space quantum theory with a universal quantum bit," *Phys. Rev. A*, vol. 87, no. 5, 2013 (cit. on p. 17).

[AGP16] F. Albarelli, T. Guaita, and M. G. Paris, "Quantum backflow effect and nonclassicality," *International Journal of Quantum Information*, vol. 14, no. 07, 2016 (cit. on p. 75).

[All69] G. R. Allcock, "The time of arrival in quantum mechanics ii. the individual measurement," *Annals of Physics*, vol. 53, no. 2, 1969 (cit. on pp. 47, 75).

[APVW16] A. Acín, S. Pironio, T. Vértesi, and P. Wittek, "Optimal randomness certification from one entangled bit," *Phys. Rev. A*, vol. 93, no. 4, 2016 (cit. on p. 27).

[Aub] G. Aubrun, *A naive look at schur–weyl duality*, Web page, Accessed: 22-04-2023 (cit. on p. 94).

[BB22] A. Bednorz and J. Batle, "Optimal discrimination between real and complex quantum theories," *Phys. Rev. A*, vol. 106, no. 4, 2022 (cit. on p. 46).

[Bel64] J. S. Bell, "On the einstein podolsky rosen paradox," *Physics Physique Fizika*, vol. 1, no. 3, 1964 (cit. on p. 7).

[BG21] M. Barbier and A. Goussev, "On the experiment-friendly formulation of quantum backflow," *Quantum*, vol. 5, 2021 (cit. on p. 48).

[BM94] A. J. Bracken and G. F. Melloy, "Probability backflow and a new dimensionless quantum number," *J. Phys. A: Math. Gen.*, vol. 27, no. 6, 1994 (cit. on pp. 47, 48, 75, 77).

[Bra21a] A. J. Bracken, "Probability flow for a free particle: New quantum effects," *Phys. Scr.*, vol. 96, no. 4, 2021 (cit. on p. 78).

[Bra21b] J. B. Brask, "Gaussian states and operations – a quick reference," arXiv: 2102.05748, 2021 (cit. on p. 59).

[BŠCA18] J. Bowles, I. Šupić, D. Cavalcanti, and A. Acín, "Self-testing of Pauli observables for device-independent entanglement certification," *Phys. Rev. A*, vol. 98, no. 4, 2018 (cit. on pp. 28, 30).

[CDPV13] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, "Quantum computations without definite causal structure," *Phys. Rev. A*, vol. 88, no. 2, 2013 (cit. on p. 108).

[CFR00] C. M. Caves, C. A. Fuchs, and P. Rungta, "Entanglement of Formation of an Arbitrary State of Two Rebits," arXiv:quant-ph/0009063, 2000 (cit. on p. 18).

[CWL+22] M.-C. Chen, C. Wang, F.-M. Liu, J.-W. Wang, C. Ying, Z.-X. Shang, Y. Wu, M. Gong, H. Deng, F.-T. Liang, Q. Zhang, C.-Z. Peng, X. Zhu, A. Cabello, C.-Y. Lu, and J.-W. Pan, "Ruling out real-valued standard formalism of quantum theory," *Phys. Rev. Lett.*, vol. 128, no. 4, 2022 (cit. on pp. 46, 121).

[EFV05] S. P. Eveson, C. Fewster, and R. Verch, "Quantum inequalities in quantum mechanics," *Ann. Henri Poincaré*, vol. 6, 2005 (cit. on pp. 48, 75, 77).

[EPR35] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" *Phys. Rev.*, vol. 47, no. 10, 1935 (cit. on p. 18).

[EZB20] Y. Eliezer, T. Zacharias, and A. Bahabad, "Observation of optical backflow," *Optica*, vol. 7, no. 1, 2020 (cit. on p. 48).

[For72] E. Formanek, "Central polynomials for matrix rings," *Journal of Algebra*, vol. 23, no. 1, 1972 (cit. on p. 93).

[FR20] C. J. Fewster and K. Rejzner, "Algebraic quantum field theory," in *Progress and Visions in Quantum Theory in View of Gravity*, F. Finster, D. Giulini, J. Kleiner, and J. Tolksdorf, Eds., Springer International Publishing, 2020 (cit. on p. 11).

[Fri12] T. Fritz, "Beyond bell's theorem: Correlation scenarios," *New J. Phys.*, vol. 14, no. 10, 2012 (cit. on p. 8).

[Gou19] A. Goussev, "Equivalence between quantum backflow and classically forbidden probability flow in a diffraction-in-time problem," *Phys. Rev. A*, vol. 99, no. 4, 2019 (cit. on pp. 77, 78).

[Gro76] A. Grossmann, "Parity operator and quantization of δ-functions," *Commun.Math. Phys.*, vol. 48, no. 3, 1976 (cit. on p. 64).

[GXL+20] 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 Inf*, vol. 6, no. 1, 2020 (cit. on pp. 119, 121).

[Haa12] R. Haag, *Local quantum physics: Fields, particles, algebras*. Springer Science & Business Media, 2012 (cit. on p. 9).

[Hal13] B. C. Hall, *Quantum theory for mathematicians*. Springer, 2013 (cit. on pp. 12, 49, 52, 60, 61).

[HGL+13] J. J. Halliwell, E. Gillman, O. Lennon, M. Patel, and I. Ramirez, "Quantum backflow states from eigenstates of the regularized current operator," *J. Phys. A: Math. Theor.*, vol. 46, no. 47, 2013 (cit. on p. 77).

[JNV+20] Z. Ji, A. Natarajan, T. Vidick, J. Wright, and H. Yuen, "MIP*=RE," arXiv:2001.04383, 2020 (cit. on pp. 9, 41).

[Kab68] P. K. Kabir, *The CP Puzzle: Strange Decays of the Neutral Kaon*. Academic Press, 1968 (cit. on p. 111).

[KM19] M. Krumm and M. P. Müller, "Quantum computation is the unique reversible circuit model for which bits are balls," *npj Quantum Inf*, vol. 5, 7, 2019 (cit. on p. 7).

[Lau09] M. Laurent, "Sums of squares, moment matrices and optimization over polynomials," in *Emerging applications of algebraic geometry*, M. Putinar and S. Sullivant, Eds. Springer, 2009 (cit. on p. 42).

[LMW+22] Z.-D. Li, Y.-L. Mao, M. Weilenmann, A. Tavakoli, H. Chen, L. Feng, S.-J. Yang, M.-O. Renou, D. Trillo, T. P. Le, N. Gisin, A. Acín, M. Navascués, Z. Wang, and J. Fan, "Testing Real Quantum Theory in an Optical Quantum Network," *Phys. Rev. Lett.*, vol. 128, no. 4, 2022 (cit. on pp. 46, 121).

[LYW+20] 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, 2020 (cit. on pp. 119, 121).

[Mar21] D. Martínez-Rubio, *Latex and math notation with links to its definition*, Web page, Accessed: 17-04-2023, 2021 (cit. on p. 2).

[MB98] G. Melloy and A. Bracken, "The velocity of probability transport in quantum mechanics," *Annalen der Physik*, vol. 510, no. 7-8, 1998 (cit. on pp. 48, 77, 78).

[MBL+13] T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann, and O. Gühne, "Device-Independent Entanglement Quantification and Related Applications," *Phys. Rev. Lett.*, vol. 111, no. 3, 2013 (cit. on p. 44).

[MME07] G. Muga, R. S. Mayato, and I. Egusquiza, Eds., *Time in quantum mechanics*. Springer Science & Business Media, 2007, vol. 734 (cit. on p. 47).

[MMG09] M. McKague, M. Mosca, and N. Gisin, "Simulating Quantum Systems Using Real Hilbert Spaces," *Phys. Rev. Lett.*, vol. 102, no. 2, 2009 (cit. on p. 19).

[Mos52] M. Moshinsky, "Diffraction in time," *Phys. Rev.*, vol. 88, no. 3, 1952 (cit. on pp. 48, 77).

[MV97] R. Meise and D. Vogt, *Introduction to functional analysis*. Clarendon Press, 1997 (cit. on p. 18).

[MYDP21] M. Miller, W. C. Yuan, R. Dumke, and T. Paterek, "Experiment-friendly formulation of quantum backflow," *Quantum*, vol. 5, 2021 (cit. on p. 48).

[MYS12] M. McKague, T. H. Yang, and V. Scarani, "Robust self-testing of the singlet," *J. Phys. A: Math. Theor.*, vol. 45, no. 45, 2012 (cit. on pp. 32, 38).

[NPA07] M. Navascués, S. Pironio, and A. Acín, "Bounding the set of quantum correlations," *Phys. Rev. Lett.*, vol. 98, no. 1, 2007 (cit. on p. 42).

[NPA08] M. Navascués, S. Pironio, and A. Acín, "A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations," *New J. Phys.*, vol. 10, 2008 (cit. on p. 42).

[PBR12] M. F. Pusey, J. Barrett, and T. Rudolph, "On the reality of the quantum state," *Nature Phys*, vol. 8, no. 6, 2012 (cit. on p. 24).

[PGKW05] M. Penz, G. Grübl, S. Kreidl, and P. Wagner, "A new approach to quantum backflow," *J. Phys. A: Math. Gen.*, vol. 39, no. 2, 2005 (cit. on p. 48).

[Pro20] C. Procesi, "Tensor fundamental theorems of invariant theory," arXiv: 2011.10820, 2020 (cit. on pp. 90, 118, 121).

[Pro22] C. Procesi, "A construction of swap or switch polynomials," *Advances in Mathematics*, vol. 400, 2022 (cit. on pp. 96, 97, 121).

[PTMM13] M. Palmero, E. Torrontegui, J. G. Muga, and M. Modugno, "Detecting quantum backflow by the density of a bose-einstein condensate," *Phys. Rev. A*, vol. 87, no. 5, 2013 (cit. on p. 48).

[PV08] K. F. Pál and T. Vértesi, "Efficiency of higher-dimensional Hilbert spaces for the violation of Bell inequalities," *Phys. Rev. A*, vol. 77, no. 4, 2008 (cit. on p. 19).

[QDS+19a] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, "Probabilistic exact universal quantum circuits for transforming unitary operations," *Phys. Rev. A*, vol. 100, no. 6, 2019 (cit. on pp. 87, 108, 118).

[QDS+19b] 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.*, vol. 123, no. 21, 2019 (cit. on pp. 87, 108, 118).

[RTW+21] M.-O. Renou, D. Trillo, M. Weilenmann, T. P. Le, A. Tavakoli, N. Gisin, A. Acín, and M. Navascués, "Quantum theory based on real numbers can be experimentally falsified," Nature, vol. 600, 2021 (cit. on pp. 2, 7).

[SCHL16] I. S. B. Sardharwalla, T. S. Cubitt, A. W. Harrow, and N. Linden, "Universal Refocusing of Systematic Quantum Noise," arXiv:1602.07963, 2016 (cit. on p. 86).

[Sie17] A. F. Siegel, *Counterexamples in probability and statistics*. Routledge, 2017 (cit. on p. 11).

[SST+23] P. Schiansky, T. Strömberg, D. Trillo, V. Saggio, B. Dive, M. Navascués, and P. Walther, "Demonstration of universal time-reversal for qubit processes," *Optica*, vol. 10, no. 2, 2023 (cit. on pp. 119, 121).

[Stu60] E. C. G. Stueckelberg, "Quantum theory in real hilbert space," *Helv. Phys. Acta*, vol. 33, no. 727, 1960 (cit. on p. 17).

[TDN20] D. Trillo, B. Dive, and M. Navascués, "Translating Uncontrolled Systems in Time," *Quantum*, vol. 4, 2020 (cit. on pp. 2, 85).

[TDN23] D. Trillo, B. Dive, and M. Navascués, "Universal Quantum Rewinding Protocol with an Arbitrarily High Probability of Success," *Phys. Rev. Lett.*, vol. 130, no. 11, 2023 (cit. on pp. 2, 85).

[TLN22] D. Trillo, T. P. Le, and M. Navascués, "Quantum supremacy in mechanical tasks: projectiles, rockets and quantum backflow," arXiv:2209.00725, 2022 (cit. on pp. 2, 47).

[Tsi06] B. Tsirelson, "How often is the coordinate of a harmonic oscillator positive?," arXiv:quant-ph/0611147, 2006 (cit. on pp. 47, 59, 60, 83).

[VB96] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, 1996 (cit. on p. 42).

[vNeu31] J. v. Neumann, "Die eindeutigkeit der schrödingerschen operatoren," *Mathematische Annalen*, vol. 104, no. 1, 1931 (cit. on pp. 61, 62).

[Wat18] J. Watrous, *The theory of quantum information*. Cambridge University Press, 2018 (cit. on p. 40).

[WB05] J. Wood and A. Bracken, "Bounds on integrals of the wigner function: The hyperbolic case," *J. Math. Phys.*, vol. 46, no. 4, 2005 (cit. on p. 73).

[Wer87] R. Werner, "Local preparability of states and the split property in quantum field theory," *Lett Math Phys*, vol. 13, no. 4, 1987 (cit. on p. 10).

[Wer88] R. F. Werner, "Wigner quantisation of arrival time and oscillator phase," *J. Phys. A: Math. Gen.*, vol. 21, no. 24, 1988 (cit. on pp. 72, 73, 75).

[Wil91] D. Williams, *Probability with martingales*. Cambridge university press, 1991 (cit. on pp. 112, 115).

[WJG+22] D. Wu, Y.-F. Jiang, X.-M. Gu, L. Huang, B. Bai, Q.-C. Sun, X. Zhang, S.-Q. Gong, Y. Mao, H.-S. Zhong, M.-C. Chen, J. Zhang, Q. Zhang, C.-Y. Lu, and J.-W. Pan, "Experimental refutation of real-valued quantum mechanics under strict locality conditions," *Phys. Rev. Lett.*, vol. 129, no. 14, 2022 (cit. on pp. 46, 121).

[YSM22] S. Yoshida, A. Soeda, and M. Murao, "Reversing unknown qubit-unitary operation, deterministically and exactly," arXiv:2209.02907, 2022 (cit. on p. 119).

[ZALS22] L. H. Zaw, C. C. Aw, Z. Lasmar, and V. Scarani, "Detecting quantumness in uniform precessions," *Phys. Rev. A*, vol. 106, no. 3, 2022 (cit. on pp. 47, 83).

[ZZHE93] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, ""event-ready-detectors" bell experiment via entanglement swapping," *Phys. Rev. Lett.*, vol. 71, no. 26, 1993 (cit. on p. 28).