universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Quantum information approach to the foundations of quantum physics and spacetime“

verfasst von / submitted by

Marius Ernst Hagen Krumm, M.Sc. B.Sc.

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

Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, 2021 / Vienna 2021

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 Dipl.-Phys. Dr. Markus P. Müller

Acknowledgements

During my PhD, I encountered a lot of people that had a strong impact on me as a person and as a scientist.

The first person I wish to thank is one of my supervisors, Markus P. Müller. Markus entered my life exactly at the right time. Even long before studying physics, I was obsessed with understanding the deepest foundations of nature and reality. I was fascinated by the idea that gravity is really caused by the curvature of spacetime and that our classical reality is merely a shadow of the quantum world. However, during my Master studies I developed a deep sense of disappointment and frustration about how general relativity and quantum physics are actually justified and constructed. Luckily, at that time Markus became a group leader in Heidelberg. He introduced me to quantum information theory and its application to the foundations of physics. The operational approach fascinated me and the reinterpretation of physics as information processing guided me to Wheeler's famous *It from Bit* philosophy.

After being my Master thesis supervisor, he also became my PhD supervisor. Through all this time he sacrificed a lot of time and energy to support and guide me. I extremely admire both his spectacular creativity for finding fundamentally new project ideas and his immense mathematical capabilities that allow him to develop these ideas into a rigorous framework. Markus became a role model for me, not just scientifically, but also as a person. He has the aura of a benevolent father who often patiently listened to my sorrow and worries. His genuinely good heart kept me sane through all the severe existential and immense political pressure in academia and many other hardships.

I also have to thank my other supervisor, Časlav Brukner. Due to Časlav's help, my PhD studies could progress without any major disruption. Without him, this would have been impossible. I am thankful that Časlav welcomed me into his group and its activities. But most importantly, I am extremely impressed by Časlav's scientific program that he allowed me to be part of. He and his group made me realize and appreciate the crucial importance of what I like to call *operational quantum gravity*: The insight that unlocking fundamentally new secrets of nature likely requires to approach quantum gravity, but with an operational perspective focussing on physical protocols such that we do not get lost in artifacts of formalisms. This insight convinced me to adopt operational quantum gravity as one of my main future research endeavors. I expect that Časlav and his former and current group

members will make great discoveries that shape our future concepts of both quantum foundations and quantum spacetime. My hope is that I can contribute to these discoveries.

Through Časlav's group I had the honor to meet and discuss with many inspiring people, including Philippe Allard Guérin, Luca Apadula, Mateus Araújo, Veronika Baumann, Ämin Baumeler, Luis Barbado, Costantino Budroni, Esteban Castro-Ruiz, Borivoje Dakić, Flavio Del Santo, Flaminia Giacomini, Anne-Catherine de la Hamette, Viktoria Kabel, Ilya Kull, Martin Renner, Giulia Rubino, and Marco Túlio Quintino.

I am extremely thankful for my fruitful collaborations with Philippe Allard Guérin, Veronika Baumann, Costantino Budroni, Roberto D. Baldijão, Časlav Brukner, Andrew J. P. Garner, Philipp A. Höhn, and Markus P. Müller.

Also, I'd like to thank all the people at IQOQI Vienna, including Jessica Bavaresco, Yelena Guryanova, Marcus Huber, Benjamin Skarabella, and Miguel Navascués. I really enjoyed the relaxed and friendly atmosphere you provided. In particular, I thank the people that I shared an office with: Edgar A. Aguilar, Roberto D. Baldijão, Benjamin Dive, Andrew J. P. Garner, Mark F. Rogers, Le Phuc Thinh, David Trillo and Mirjam Weilenmann. I am especially happy that David became a good friend who gave me a life with hobbies beyond academia.

I am also happy that I got to share my office with David Schmid during my time at Perimeter Institute. It is very uncommon to meet someone who is as obsessed with quantum information theory and quantum foundations as I am.

Furthermore, I wish to express my thanks for Hans Briegel and his group, and all their effort and time they invested to explain their projects and endeavors to me.

Also, I gratefully acknowledge the hospitality of Renato Renner's group during my visit of their group, and in particular I thank Raban Iten for working with me on machine learning.

I also thank my parents Dittmar Krumm and Astrid Emmel, my brother Daniel Emmel, my dog Linda, and my cat Fenix for supporting and believing in me. Soon I will be home again.

I apologize to everyone I forgot to mention in these acknowledgements.

Academic CV

- **Since July 2017:** PhD student in physics at *University of Vienna* and *Institute for Quantum Optics and Quantum Information (IQOQI) Vienna*.
  Supervisors: Markus P. Müller and Časlav Brukner
- **06/2017:** Move from Canada to Vienna due to move of supervisor. I left the PhD program in Canada early to become a PhD student in Vienna instead.
- **09/2016-06/2017:** Resident Graduate Student at *Perimeter Institute for Theoretical Physics*, Waterloo ON, Canada.
  09/2015-06/2017: PhD student in applied mathematics at *University of Western Ontario*, London ON, Canada.
  Supervisor: Markus P. Müller
- **2012-2015:** Master studies in physics at *Ruprecht-Karls-Universität Heidelberg*.
  Degree: Master of Science (M.Sc.)
- **2009-2012:** Bachelor studies in physics at *Ruprecht-Karls-Universität Heidelberg* . Degree: Bachelor of Science (B.Sc.)

Research visits: Visited Renato Renner's group at ETH Zurich for two weeks in August 2019

Organization of scientific events: I was one of the organizers of the *Vienna Summer School 2020 on Gravitational Quantum Physics*.

Selection of conference/workshop presentations and posters:

- **03/2021:** Talk *Quantum reference frame transformations as symmetries and the paradox of the third particle* at *APS March Meeting 2021*, USA
- **06/2019:** Talk *Qubits of higher dimension cannot compute* at the conference *Quantum Information Revolution: Impact to Foundations 2019*, Växjö, Sweden
- **05/2019:** Poster and Talk *Physics beyond the quantum model: General probabilistic theories* at the *International Student Workshop on Quantum Technologies*, University of Science and Technology of China, Shanghai, China
- **09/2018:** Poster *The input problem of quantum machine learning* at the conference *Quantum Machine Learning Plus*, Innsbruck, Austria

Table of Contents

Abstract

This thesis applies methods and concepts from quantum information theory (QIT) to investigate the foundations of quantum physics and spacetime.

The process matrix formalism allows to describe indefinite causal structures. This thesis develops a systematic combination of process matrices and quantum clocks. Furthermore, it is proved that there cannot exist a universal formalism to describe the parallel application of indefinite causal structures.

Furthermore, this thesis investigates a scenario in which a quantum system is probed by a lab with an incomplete external reference frame. For the example of a missing origin, the observables and symmetries are characterized.

The setting of device-independence considers abstract devices described only via their input-output-correlations. This thesis investigates a modification in which the abstract inputs are exchanged with spacetime parameters. Non-classical bipartite correlations for the cases of angles or directions as inputs are characterized.

The framework called General Probabilistic Theories allows to develop theories of physics that are not classical or quantum. This thesis investigates objectivity of measurement outcomes in such theories by generalizing the phenomenology of quantum Darwinism. Furthermore, it is proved that theories with Bloch balls with dimension different from three cannot have interactions.

Zusammenfassung

Diese Doktorarbeit wendet Methoden und Konzepte der Quanteninformationstheorie (QIT) an um die Grundlagen der Quantenphysik und Raumzeit zu untersuchen.

Der Prozess-Matrix-Formalismus ermöglicht die Beschreibung von indefiniten kausalen Strukturen. Diese Doktorarbeit entwickelt eine systematische Kombination von Prozess-Matrizen und Quantenuhren. Weiterhin wird gezeigt, dass es keinen allgemeingültigen Formalismus geben kann, um die parallele Nutzung von indefiniten kausalen Strukturen zu beschreiben.

Außerdem betrachtet diese Doktorarbeit ein Szenario, in dem ein Quantensystem von einem Labor mit einem unvollständigen Bezugssystem untersucht wird. Für das Beispiel eines fehlenden Ursprungs werden die Observablen und Symmetrien charakterisiert.

Im Feld der Gerät-Unabhängigkeit werden abstrakte Geräte betrachtet, die nur über ihre Eingabe-Ausgabe-Korrelationen beschrieben werden. Diese Doktorarbeit untersucht eine Modifikation, in der die abstrakten Eingaben mit Raumzeit-Parametern ersetzt werden. Für Beispiele von Winkeln oder Richtungen als Eingabe-Parameter werden die nicht-klassischen bipartiten Korrelationen charakterisiert.

Das Feld der Allgemeinen Probabilistischen Theorien ermöglicht es, physikalische Theorien zu entwickeln, die weder klassisch noch quantenmechanisch sind. Diese Doktorarbeit untersucht die Objektivität von Messergebnissen in solchen Theorien, indem das Phänomen namens Quanten-Darwinismus angepasst wird. Außerdem wird gezeigt, dass Theorien mit Bloch-Bällen, die eine andere Dimensionalität als drei haben, keine Interaktionen haben können.

Chapter 1

Introduction

1.1 Motivation and goal

Quantum theory [1] is the foundation of large parts of modern physics. It allows to accurately describe atomic and molecular physics as well as quantum optics and condensed matter systems and ultra-cold gases. Combining quantum theory with special relativity theory [2] led to quantum field theory [3], which provides the framework for both high energy particle physics and nuclear physics.

Notwithstanding these scientific achievements, the exploration of the quantum aspects of our universe is far from over. Indeed, while quantum theory enjoys excellent experimental confirmation in the aforementioned fields, there exist extreme scales and exotic scenarios far beyond the reach of current experimental technology. Perhaps the most important experimentally unexplored regime is that of quantum gravity[4–6], which combines quantum physics with another crucial part of modern physics, namely general relativity [7].

When formulating theories and frameworks for new regimes and scenarios, it is not always clear how to extract observable predictions from the basic building blocks of the theory. To avoid this problem it can be fruitful to adapt the operational perspective of quantum information theory. This perspective imagines agents who use probes or other physical systems to implement protocols. For example, in the context of investigating new regimes of physics, probes can be used to extract information about an environment or new phenomena can be used as a resource for information processing and transmission tasks [8–18].

Quantum information theory [19, 20] does not assume any particular physical implementation (like electron spins, photon polarization, ...) and therefore is very general. Nonetheless, it provides a powerful formalism that allows to describe quantum phenomena using density matrices, *Positive Operator Valued Measurements (POVMs)* and channels. The quantum information point of view leads to information-theoretic protocols that allow to extract useful observable quantities such as the probability to confuse two similar physical scenarios with each other, or the amount of entanglement that can be extracted. Entropic information measures such as von Neumann entropy can help to find connections to statistical physics and thermodynamics. Super-strong quantum correlations, such as in Bell scenarios [21, 22], offer a way to test whether the physical phenomenon is indeed governed by quantum physics.

The goal of this thesis is to apply the quantum information approach to new or exotic physical scenarios in the foundations of quantum physics and spacetime. Many of the scenarios encountered in this thesis confront quantum physics and spacetime with each other in unconventional ways beyond the standard textbook settings of quantum field theory or non-relativistic phase space quantum mechanics. Therefore, this approach challenges our understanding of quantum spacetime in new or unusual settings and formalisms, and mastering these challenges may give us important hints about quantum gravity.

This is achieved by the use of several frameworks that originated in or are closely related to quantum information theory. All of these frameworks relax or modify different aspects of the typical settings in quantum information theory or physics in general:

The framework of *General Probabilistic Theories (GPTs)/Operational Probabilistic Theories (OPTs)* [23–29], allows to construct theories of physics that are neither classical not quantum. Mathematically, it generalizes quantum information theory by replacing density matrices with general vectors in a convex set, the state space, and adapts operational quantum information concepts like channels and entropies to these new state spaces.

The setting of *device-independent information processing* [30–33] reduces physical devices to black boxes that receive an abstract input value and randomly produce an output value. These devices are only modeled via their input-output statistics. Nonetheless, there exist spacetime scenarios like Bell tests [21] in which the statistics of such devices can prove that the devices have quantum correlations that may be useful for quantum information processing.

The concept called *indefinite causal structure* [9–11] generalizes our usual notion of cause-and-effect relations. The application of quantum physics to causality may give rise to exotic causal structures, such as superpositions of causal orders [34–36]. The framework of *process matrices* [9] models causal interventions via quantum instruments from quantum information theory and the process matrix encapsulates the causal structure by mapping these quantum instruments to the actual outcome probabilities.

In the following sections, we will give a more detailed introduction to these frameworks. Afterwards, we will explain the structure of this thesis and how these frameworks get applied.

1.2 Physics beyond the quantum model: General/operational probabilistic theories

General probabilistic theories (GPTs)/operational probabilistic theories (OPTs) is the name of a framework that allows to formulate probabilistic theories of physics that are neither classical nor quantum [23–29]. For this introduction to the framework we follow [24, 25, 27]. To approach physics beyond quantum theory from a meaningful perspective, one imagines lab experiments in the following way, see Figure 1.1: One considers a preparation device that upon activation outputs a single physical system. The system may experience some evolution and at the end it gets measured. The measurement is assumed to yield a definite outcome, so also GPTs/OPTs have a

collapse of the state. The state ω of a physical system can be defined operationally as a vector listing all the outcome probabilities for all possible choices of measurements:

$$\omega = \begin{pmatrix} p(o_1|m_1) \\ p(o_2|m_1) \\ \vdots \\ p(o_1|m_j) \\ p(o_2|m_j) \\ \vdots \end{pmatrix} \tag{1.1}$$

Here, o_k is the k-th outcome and m_j the j-th measurement setting. However, already in 2-dimensional quantum theory there would be infinitely many choices of measurements. Instead one assumes that there exists a (minimal) finite set of measurements whose outcome probabilities completely determine the outcome probabilities of all other measurements. Such a set is called *fiducial*. Then the state can be defined as the finite-dimensional vector listing the outcome probabilities for a set of fiducial measurements. The set of all states ω is called the state space Ω_A and it generalizes the set of density matrices of a given dimension. The state space is assumed to be compact: It is bounded because probabilities are. It is closed because we do not make a difference between probability distributions that can be approximated arbitrarily well and those that can be implemented perfectly.

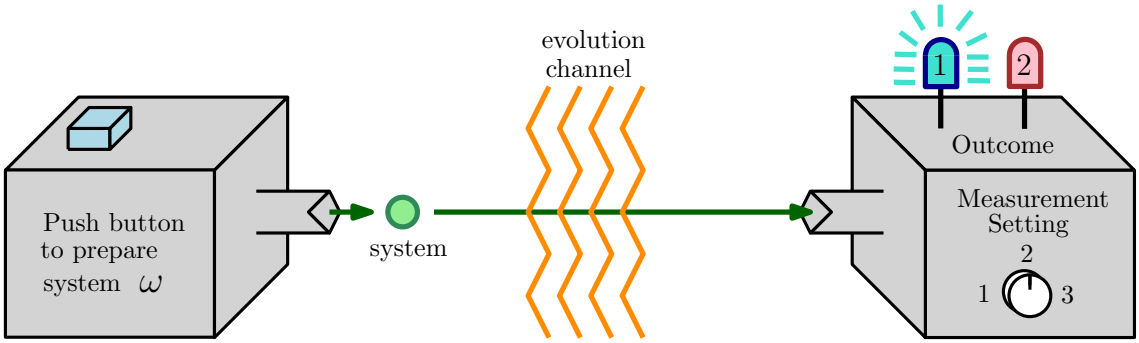


Figure 1.1: The typical operational scenario in general/operational probabilistic theories. A system preparation device emits a single physical system when a button is pushed. The state ω of the system that the device prepares is characterized as the list of outcome probabilities for all the relevant choices of measurements. The framework also allows to describe time evolution and transformations using generalizations of quantum channels and instruments.

Now that we have a concept of state, we can adapt other elementary notions of quantum information theory. The first such concept is that of a POVM. A measurement is described by a collection of maps $\{e_j\}_{j=1\ldots m}$ and these maps are called *effects*. These effects map states $\omega \in \Omega_A$ to valid probabilities, i.e. $e_j(\omega) \in [0,1]$. Furthermore, as probabilities should sum up to one, we have $\sum_{j=1}^{m} e_j(\omega) = 1$. Channels, often called *transformations*, are described by maps from a state space to a state space, i.e. $T : \Omega_A \to \Omega_B$. Reversible transformations R are those that are invertible and for which the inverse R^{-1} is also considered to be a physical transformation. If one introduces a notion of parallel composition of systems, then one can also demand

a generalization of complete positivity.

The central mathematical foundation of this framework is the field called *convex geometry* [37], because convex-linear combinations of states play an important role, compare Figure 1.2: Operationally, one assumes that one has access to randomness, i.e. a random number generator (RNG). Let us say this RNG has n possible outcomes, and outcome j occurs with probability q_j. One can imagine a preparation device that contains n smaller system preparation devices, each of which can prepare a state ω_j. If one activates the outer device, it triggers the random number generator, whose outcome selects which smaller device gets used. The resulting output ω_{full} of the outer device can once again be described as list of probabilities as in Eq. (1.1) via the law of total probability, resulting in a convex-linear mixture of the states that the smaller devices prepare:

$$\omega_{\text{full}} = \sum_{j=1}^{n} q_j \omega_j \tag{1.2}$$

Therefore, the state space is a convex set. The *pure* states are the states for which no non-trivial convex decomposition is possible, i.e. $\omega = \sum_j p_j \nu_j$ with $p_j \geq 0$, $\sum_j p_j = 1$ and states ν_j implies $p_j = 0$ or $\nu_j = \omega$. States that are not pure are called *mixed*.

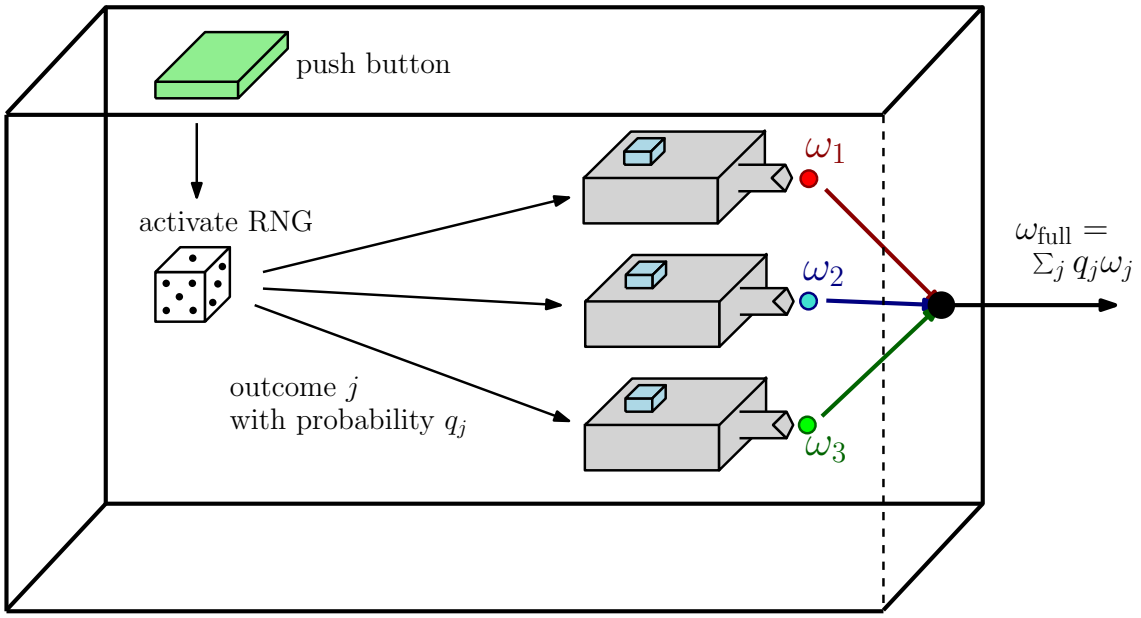


Figure 1.2: This figure shows how mixed states can be prepared in the GPT/OPT framework. One imagines a large system preparation device that contains a random number generator and several smaller state preparation devices. Activating the large device triggers the RNG and the RNG decides which of the smaller devices gets activated. The prepared state gets sent to the outside. The outside never learns the result of the RNG.

The previous thought experiment confronts us with two different descriptions of the same operational scenario. The first perspective says that our big preparation device prepares the state ω_{full} and this is the state that we use to calculate measurement probabilities $e_j(\omega_{\text{full}})$. The second perspective first conditions on the outcome

of the internal RNG. Let us say we got outcome k, i.e. the state ω_k was prepared. Now we use this state to calculate measurement probabilities, i.e. $e_j(\omega_k)$. Finally, we apply the law of total probability to average over the outcome k of the RNG, giving us outcome probabilities $\sum_k q_k e_j(\omega_k)$. As the final outcome probabilities have to agree no matter which description we adapt, we find:

$$e_j\left(\sum_k q_k \omega_k\right) = e_j(\omega_{\text{full}}) = \sum_k q_k e_j(\omega_k) \tag{1.3}$$

Therefore, measurements have to be convex-linear in the state. Similarly, transformations are convex-linear in the state. This convex-linearity can be extended to full linearity by adding one more dimension, leading to a real vector space A in which Ω_A is embedded. The extra dimension represents the normalization u_A, which is defined as 1 on the state space Ω_A. In quantum theory, it is the trace of a density matrix. So far we always assumed that if one activates the system preparation device, a system is successfully prepared. To describe the situation in which such a device may fail to output a system, one can use the normalization to model the success probability. For example, a sub-normalized state $p\omega$ with $\omega \in \Omega_A$ and $p \in [0, 1]$ would then mean that with probability p a system is prepared in the state ω. The normalization is then defined as $u_A(p\omega) = p$ and can therefore be linearly extended to A. This interpretation of sub-normalized states is consistent with our interpretation of statistical mixtures as in $\omega_{\text{full}} = \sum_j q_j \omega_j$. The individual terms $q_j\omega_j$ can be interpreted as the fact that the device associated with outcome j only outputs a state with probability q_j, but if it does the state is ω_j. The procedure that always fails outputs the zero-vector as an unnormalized state.

Mathematically, it is convenient to also include super-normalized states $\lambda\omega$ with $\lambda > 1$ and $\omega \in \Omega_A$: Together, the normalized, sub-normalized and super-normalized states form a state cone $A_+ := \mathbb{R}_{\geq 0} \cdot \Omega_A$ which is pointed (the zero-state is its tip), see Figure 1.3. As cones are commonly-studied objects in convex geometry, this allows to adapt methods and results from convex geometry, see e.g. [23].

So far, we chose to represent states as lists of outcome probabilities for fiducial measurements. However, there are many other ways to represent a state space without spoiling the (convex-) linear structure of the framework. These changes of representation can be achieved by using a linear function L that can be inverted when restricted to the state cone. One example is the density matrix formalism in quantum theory. It is standard to assume that the representation is of minimal dimension, i.e. that the linear span of the state cone A_+ is the full vector space A.

The measurement effects live in the dual cone

$$A_+^* = \{e \in A^* \mid e(\omega) \geq 0 \ \forall \omega \in \Omega_A\} \tag{1.4}$$

where $A^* \simeq A$ is the dual space of A (i.e. the set of linear functions from A to the real numbers). A common assumption is that all such linear functions that give valid probabilities on all states are physically allowed. This is called the *no-restriction hypothesis*. One motivation is that restrictions such as super-selection rules, conservation laws and practical infeasibility are properties of the physical implementation, not of the fundamental theory that seeks to replace classical or quantum theory. Similarly, one often assumes that measurements constructed from physically allowed

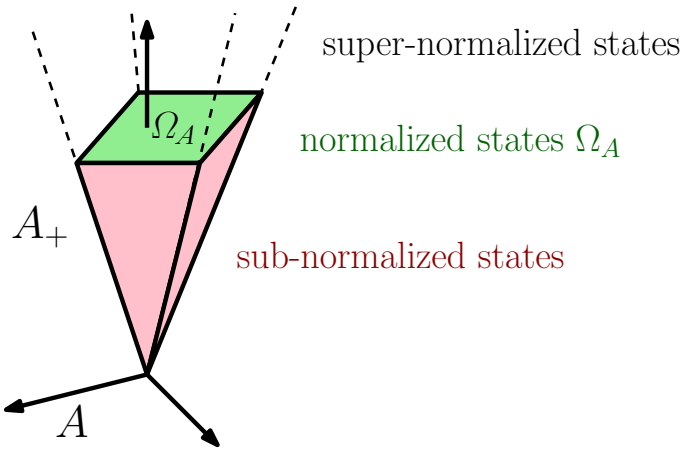


Figure 1.3: The set of normalized, sub-normalized and super-normalized states forms a state cone A_+. In the figure, the normalization u_A can be visualized as the coordinate axis along the cone. The normalized state space Ω_A is given as the subset of the state cone that satisfies $u_A(\omega) = 1$. The tip of the cone is the origin and represents definite failure to prepare a system. The vector space is called A. The standard is to assume that A is chosen to be of minimal dimension.

effects are physically allowed. If no such assumptions are made, then the set of physically allowed measurements or effects and the set of allowed transformations must be explicitly defined in addition to the state space. However, even in those cases one assumes that the set of effects is generating, i.e. of full dimension, such that all states can be distinguished via tomography, and that it is convex and closed for the same reasons as above [38].

In principle, every finite-dimensional compact and convex set represents a state space that can be investigated in this framework. Frequently studied examples of exotic state spaces include polytopes and polygons, see e.g. [39–41]. A polygonal state space of particular importance is the square bit, because it appears in the generalized no-signaling theory/box world for the case of 2 incomes and 2 outcomes [24].

Because of the generality of the framework, there is no universal notion of parallel composition of state spaces. Nonetheless, one popular and well-motivated approach is based on the framework of *tomographic locality*: First of all, one assumes that it should be possible to do protocols such as state preparations and measurements independently in parallel. For example, two separated agents Alice and Bob can prepare states $\omega_A \in \Omega_A$ and $\nu_B \in \Omega_B$ independendly. Similarly, the agents should be capable of implementing independent measurements $e_j^{(A)} \in A_+^*$ and $f_k^{(B)} \in B_+^*$ independently in parallel. The principle of *tomographic locality* says that all such parallel combinations of local measurements and their correlations already completely specify a physical state.

Let us sketch a quick way for determining the vector space AB that contains the combined state space Ω_{AB}, similar to [23]. At this point it is convenient to interpret states as functions that map measurement effects to probabilities. The principle of tomographic locality now says that we only have to consider local measurements, i.e. states can be represented as bilinear maps $\omega_{AB}(e_j^{(A)}, f_k^{(B)}) \in \mathbb{R}$. The universal

property of the tensor product implies that we can represent the same state as a linear map $\omega_{AB}(e_j^{(A)} \otimes f_k^{(B)})$. Therefore, composite states ω_{AB} live on the dual space of the tensor product $A^* \otimes B^*$, and thus by dimension counting we can see that all composite states ω_{AB} can be represented as elements of $A \otimes B$. Now parallel procedures can be represented by the tensor product, e.g. $\omega_A \otimes \nu_B$ for states. More detailed arguments can be found e.g. in [24, 27, 42, 43].

Hence, tomographic locality recovers the tensor product structure of quantum theory. Both standard quantum theory and classical probability theory satisfy this postulate. It expresses the intuition that a composite system is completely specified by its parts and their relations. In particular, parallel procedures should not lead to the emergence of "magical" extra degrees of freedom. A theory that violates this principle is *real quantum theory*, i.e. quantum theory restricted to the real numbers. Here σ_y is not a valid local observable because of its imaginary entries. However, $\sigma_y \otimes \sigma_y$ is a valid observable because the imaginary units cancel. Therefore, $\sigma_y \otimes \sigma_y$ is a new holistic degree of freedom. While tomographic locality identifies the combined vector space as the tensor product of the local spaces and allows to represent parallel procedures via the tensor product, it does not identify the combined state space itself. Indeed there are many different compatible choices, such as the minimal tensor product that only contains separable states or the maximal tensor product that contains all vectors in AB that give valid probabilities for all local, independent measurements of the form $e_j^{(A)} \otimes f_k^{(B)}$.

A main application of the GPT/OPT framework is to derive quantum theory from physical and information-theoretical postulates instead of the usual formal postulates about Hilbert spaces and projection operators. This resulted in a wave of reconstructions of quantum theory, e.g. [25, 27–29, 42]. However, the generality and operational foundation also make the framework suitable to search for theories that may one day replace quantum theory ("*post-quantum physics*"), see e.g. [44]. While quantum theory is experimentally well-confirmed, one should prepare for the possibility that quantum theory will be replaced eventually. For once, current experiments cannot access all scales. It is possible that for extreme scales quantum theory becomes inadequate. This scenario is comparable to the historical discovery of quantum mechanics: Also classical physics was well confirmed for a long time, but once the world of atoms and photons became experimentally accessible, it became clear that classical physics had to be replaced with quantum physics [1]. It is conceivable that the regime of quantum gravity will similarly show that quantum theory needs to be replaced: While most approaches to quantum gravity search for the right way to replace and quantize gravity [4–6], it is by no means clear that quantum theory will still be the adequate formalism for the right theory of quantum gravity. The framework of general/operational probabilistic theories can provide an operational starting point to develop replacements [10, 11]. Furthermore, it allows to investigate the relationship between different physical principles and phenomena by offering a description that does not rely on the abstract quantum framework, see e.g. [41, 45, 46].

An example of a quantum phenomenon is quantum Darwinism [47, 48] which Roberto D. Baldijão, Andrew J. P. Garner, Markus P. Müller, and I generalized to GPTs/OPTs [49]. Furthermore, my supervisor Markus P. Müller and I investigated theories constructed from higher dimensional Bloch balls but showed that these theories do not allow for interactions [50].

1.3 Device-independent information processing

Quantum physics allows for correlations between separated systems that are much stronger than those admitted by classical statistics. The most famous example for this crucial observation is given by Bell scenarios [21], in which non-classical correlations are called *Bell non-local* correlations. In this scenario one considers two isolated devices, typically assumed to be spacelike separated.

The devices are reduced to black boxes. The agents using those devices randomly choose an input from an abstract finite list of possible inputs. Likewise, the devices randomly return an abstract output from a finite list of possible outputs. For the first agent, Alice, the device is characterized by the probability distribution $p(a|x)$, where x is the input and a is the output, while the device of the second agent, Bob, is likewise described as $p(b|y)$. The combined statistics of the two devices is given by the probability distribution $p(a, b|x, y)$. The isolation or spacelike separation is implemented by the no-signalling assumption: The marginal statistics one agent sees should be independent of the input choice of the other agent, i.e.

$$p(a|x, y) = \sum_b p(a, b|x, y) = p(a|x) \tag{1.5}$$

$$p(b|x, y) = \sum_a p(a, b|x, y) = p(b|y) \tag{1.6}$$

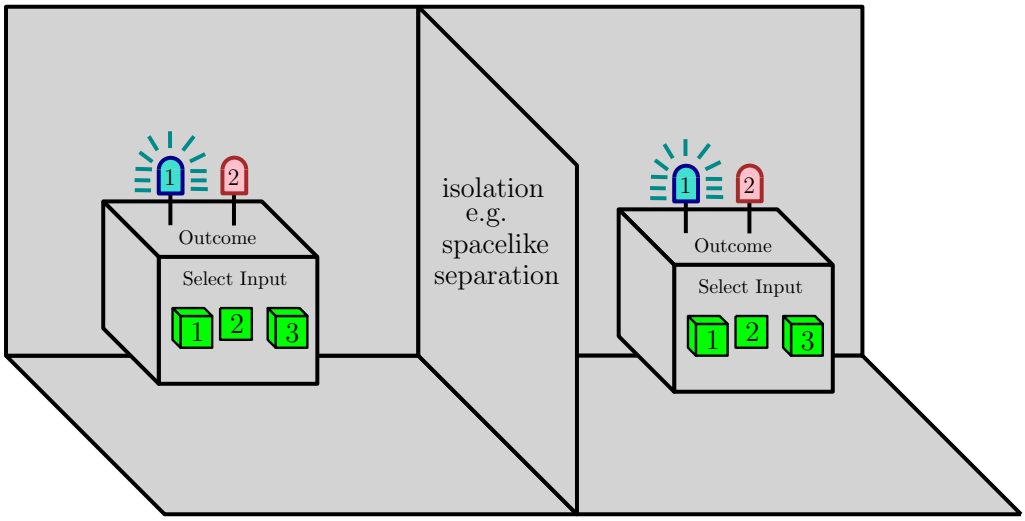


Figure 1.4: A Bell scenario modeled in the device-independent approach. The two devices are black boxes that receive an abstract input value (here the number 1, 2 or 3) and randomly produce an output value (here the number 1 or 2). In the figure, both agents chose input 2 and both got output 1. The wall represents that the devices are isolated from each other. Often times one assumes that the isolation is due to spacelike separation such that relativity does not allow the devices to signal to each other. If the devices share some particular entangled quantum states, they may have correlations that cannot be reproduced by local hidden variable models.

The notion of *local hidden variable models (LHV)* formalizes what one means by classical correlations in Bell scenarios. It is based on the idea that once one takes the entire preparation history of the two classical devices into account, they behave independently of each other while isolated. This classical preparation history may

depend on classical parameters $\lambda \in \Lambda$ that are called the *hidden variables*. The hidden variables are distributed according to a probability distribution or density $\rho(\lambda)$. If one knows the values of the hidden variables, then Alice's isolated device behaves according to $p_A(a|x,\lambda)$ while Bob's isolated device behaves according to $p_B(b|y,\lambda)$. The combined statistics of this local hidden variable model is then:

$$p(a,b|x,y) = \int \mathrm{d}\lambda \rho(\lambda) p_A(a|x,\lambda) p_B(b|y,\lambda) \tag{1.7}$$

In the simplest Bell scenario, both agents can only choose two input values $x, y \in \{1,2\}$ and only two output values $a, b \in \{-1,+1\}$ are possible. For this case of two output values, one can define a meaningful correlation function as

$$\begin{aligned} C(x,y) :=& \sum_{a,b=\pm 1} a \cdot b \cdot p(a,b|x,y) \\ =& p(+1,+1|x,y) + p(-1,-1|x,y) - p(+1,-1|x,y) - p(-1,+1|x,y) \end{aligned} \tag{1.8}$$

which can be interpreted as the probability for correlation minus the probability for anti-correlation. In this scenario, all local hidden variable models satisfy the CHSH inequality [22]

$$-2 \leq C(1,1) + C(1,2) + C(2,1) - C(2,2) \leq 2 \tag{1.9}$$

However, quantum physics allows for boxes built from entangled quantum systems that violate the CHSH inequality. Such boxes are said to have *Bell non-local* correlations. Corresponding experiments can be implemented, for example, by the use of entangled photons [51]. However, as the certification of Bell non-locality only relies on the abstract inputs and outputs and their statistics, it disregards the particular choice of physical systems.

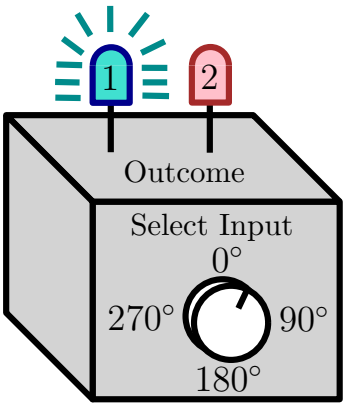


Figure 1.5: In the part of the thesis that considers device-independence, i.e. Chapter 5, we consider the following modification: Instead of the abstract inputs, the inputs are spacetime parameters. The figure shows an example in which the input is an angle. As one can imagine that the input angle can be chosen with a knob or display on the outside of the box, this modification can be argued to be compatible with the concept of device-independence.

The generality of such approaches based on non-classical correlations allowed the field of *device-independent information processing* to emerge. In its strongest form, the framework characterizes physical devices exclusively via their abstract

input-output statistics without any assumptions about the inside of the boxes, not even whether the inside is governed by quantum theory or another GPT/OPT. Despite these weak assumptions, the presence of non-classical correlations enables the use of device-independent cryptography protocols [30–32]. However, it is common to make weak extra assumptions [33, 52–55], in particular to assume that quantum theory applies. The basic intuition is that the device-independent certification of non-classical correlations allows to detect quantum properties (in particular entanglement) that allow to adapt quantum cryptography protocols.

Together with my collaborators Andrew J. P. Garner and Markus P. Müller, I developed a modification of the device-independent scenario in which the abstract inputs get replaced by space-time parameters [56], compare Figure 1.5.

1.4 Indefinite causal structure and process matrices

The process matrix framework [9] was invented to describe exotic causal structures. An essential notion in this field is that of *causal order* which refers to the ability of events to influence other events. This relation between events is usually summarized by using the words *cause* and *effect*. In physics, causality is often discussed using the light cone structure of special [2] and general relativity [7]: An event can influence events in its future lightcone, and can be influenced by events in the past light cone. However, the notion of causal structure can be investigated without the full formalism of relativity theory and it can be physically meaningful to do so: For example, spacelike separation is not the only conceivable reason why two events might not be able to influence each other: Any isolation or prevention of signals works. In statistics, the framework of *causal models* [57] is used to discuss causal structure when only probabilistic information is available.

If one combines the notion of causal structure with the principles of quantum physics, one finds exotic causal structures called *indefinite causal structure*. For example, one could imagine a superposition of an action A influencing an action B, and action B influencing action A, i.e. superpositions of cause- and effect-status. A standard example is the so-called *quantum switch* [34]: Here, a two-dimensional quantum system (qubit) controls the order of two (often unitary) transformations A and B. It is conceivable that indefinite causal structure becomes an essential phenomenon of quantum gravity, because the dynamical nature of causal structure in general relativity gets combined with quantum uncertainty [10, 11]. For example, quantum fluctuations might blur the light cone structure, or spatial superpositions of big astronomical objects might give rise to superpositions of metrics and space-times [35, 36, 58].

In the process matrix framework, causal interventions are modeled as quantum instruments. Quantum instruments generalize channels and POVMs in the sense that they have outcomes and update rules. Let $\mathcal{L}(\mathcal{H})$ be the set of linear operators on a finite-dimensional Hilbert space $\mathcal{H}$. Then a quantum instrument $\{\mathcal{T}_j\}_{j=1}^n$ is a collection of probabilistic quantum transformations $\mathcal{T}_j : \mathcal{L}(\mathcal{H}_{in}) \to \mathcal{L}(\mathcal{H}_{out})$. A probabilistic quantum transformation is linear for the same reason as the linearity of effects and transformations in GPTs/OPTs. A probabilistic quantum transformation

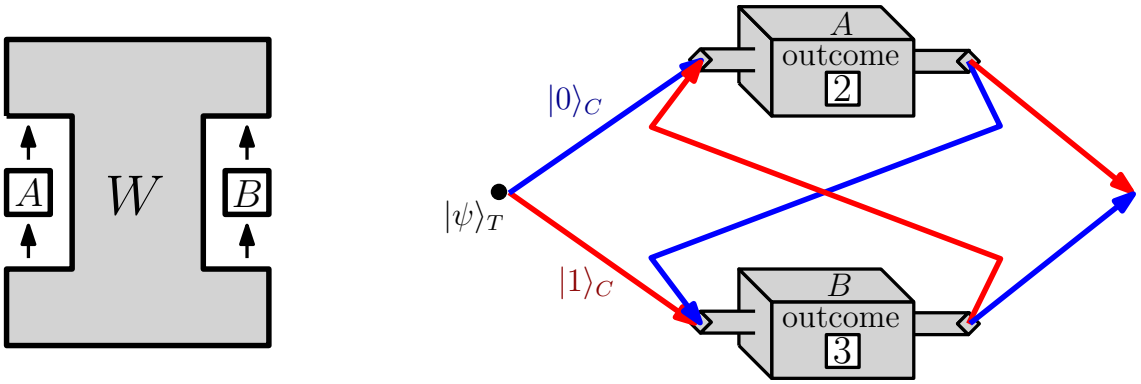


Figure 1.6: The left part of the figure shows a visualization of the basic operational scenario of process matrices that one often encounters in the literature. The process matrix W describes the environment and in particular the causal structure that connects the labs of agents A and B. The agents can freely choose their quantum instruments, visualized as blocks that get inserted into open slots of W. The right figure shows a more specific example, the quantum switch [34, 59]. Here, the environment contains a control system and a target system $|\psi\rangle_T$. If the control value is $|0\rangle_C$, the target system gets first sent to A's lab and then to B's lab, if the control value is $|1\rangle_C$ then lab B gets visited before lab A. If the control system is in a superposition state like $\frac{1}{\sqrt{2}}(|0\rangle_C + \beta|1\rangle_C)$, then both paths get coherently superimposed.

should map quantum states (described by density matrices) to quantum states, i.e. positive semi-definite operators should be mapped to positive semi-definite operators. One says that $\mathcal{T}_j$ should be *positive* or *positivity-preserving maps*. Also $\mathcal{T}_j \otimes \mathrm{Id}_{\mathcal{E}}$ should be positive maps for the identity channel of all finite-dimensional ancillary systems $\mathcal{E}$. One says that $\mathcal{T}_j$ should be *completely positive*. Just as in the GPT/OPT framework, we consider sub-normalized density matrices too: If $\rho \in \mathcal{L}(\mathcal{H}_{in})$ is a normalized density operator, i.e. $\rho \geq 0$ and $\mathrm{Tr}\rho = 1$, then $\mathrm{Tr}[\mathcal{T}_j(\rho)] \in [0, 1]$ is the probability for outcome j and $\frac{\mathcal{T}_j(\rho)}{\mathrm{Tr}[\mathcal{T}_j(\rho)]}$ is the normalized updated state if outcome j is obtained. Thus the sub-normalized operator $\mathcal{T}_j(\rho)$ describes both the updated state and the outcome probability. Therefore, $\mathcal{T}_j$ must be trace-non-increasing. Furthermore, if we average over all outcomes, the resulting map $\sum_{j=1}^{n} \mathcal{T}_j$ should be trace-preserving, i.e. a deterministic quantum transformation.

One replaces the abstract notion of events with the following operational concept, see Figure 1.6: One imagines several agents, each of them within a small lab. One assumes that inside the small lab standard quantum theory is still valid and that usual experiments are possible. However, the environment connecting the labs can be very exotic, e.g. aforementioned quantum spacetimes. During the protocol, each agent receives one physical system, acts on it with a quantum instrument (generalizing projective measurements), and then sends the system out again. The probabilities for instrument outcomes $a, b, \ldots$ given instrument choices $\{\mathcal{T}_j^{(A)}\}_j$, $\{\mathcal{T}_k^{(B)}\}_k$, $\ldots$ are then:

$$p(a, b, \ldots | \{\mathcal{T}_j^{(A)}\}_j, \{\mathcal{T}_k^{(B)}\}_k, \ldots) = \mathcal{W}(\mathcal{T}_a^{(A)}, \mathcal{T}_b^{(B)}, \ldots) \tag{1.10}$$

Here, $\mathcal{W}$ is a map that is multi-linear for the same reason most maps in GPTs/OPTs are linear: To preserve the fact that statistical mixtures are represented as convex-linear mixtures. See also Chapter 3 for how a GPT/OPT-like argument leads to multi-linearity. It is mathematically convenient to use the Choi operators associated with the (probabilistic) quantum transformations instead. For a quantum transformation $\mathcal{C} : \mathcal{L}(\mathcal{H}_{in}) \to \mathcal{L}(\mathcal{H}_{out})$, the associated Choi operator C is defined as [60, 61]

$$C := (\mathrm{Id}_{in} \otimes \mathcal{C})(|\mathrm{ent}\rangle\langle\mathrm{ent}|) \tag{1.11}$$

where $|\mathrm{ent}\rangle := \sum_{j=1}^{d_{in}} |j\rangle \otimes |j\rangle$ is an unnormalized maximally entangled state. C is positive semi-definite if and only if $\mathcal{C}$ is completely positive. The condition that $\mathcal{C}$ is trace-preserving gets replaced with $\mathrm{Tr}_{out} C = \mathbb{1}_{in}$, where the subscripts $_{in}$ and $_{out}$ refer to the input and to the output space of $\mathcal{C}$, respectively. The Choi isomorphism is linear and invertible, the inverse is

$$\mathcal{C}(\rho) = \mathrm{Tr}_{in}[(\rho^T \otimes \mathbb{1}_{out})C] \tag{1.12}$$

where ^{T} denotes transposition in the computational basis. Using the Choi operators T_j of $\mathcal{T}_j$ and the universal property of the tensor product, we find that the outcome probabilities can be expressed via a linear map $\mathcal{W}'$ as

$$p(a, b, \dots | \{T_j^{(A)}\}_j, \{T_k^{(B)}\}_k, \dots) = \mathcal{W}'(T_a^{(A)} \otimes T_b^{(B)} \otimes \dots) \tag{1.13}$$

Using the linearity of $\mathcal{W}'$, we can write this via the Hilbert-Schmidt inner product as

$$p(a, b, \dots | \{T_j^{(A)}\}_j, \{T_k^{(B)}\}_k, \dots) = \mathrm{Tr}\left[W^T \cdot (T_a^{(A)} \otimes T_b^{(B)} \otimes \dots)\right] \tag{1.14}$$

Here, W is the process matrix that gives the framework its name. It describes everything outside of the small local labs, i.e. the causal structure that connects the labs. We applied an extra transposition ^{T} in the computational basis to W. There exist different conventions concerning transpositions of process matrices and Choi operators in the literature.

The process matrix framework tries to be as general as possible. Essentially, the only condition is that the outcome probabilities are indeed probabilities, i.e. positive and sum up to 1. One also demands valid outcome probabilities for the modified scenario in which the agents share an additional entangled state. Since these conditions lead to rather technical restrictions of the valid process matrices [9, 59], we will not write them here but cite them when needed. It is important to point out that – in contrast to complete positivity of quantum transformations– the tensor product of two process matrices can fail to be a process matrix [62]. In Chapter 3 I report a result that I found with my collaborators Philippe Allard Guérin, Costantino Budroni and Časlav Brukner that there is no reasonable universal replacement to describe the parallel application of process matrices [63].

One standard example for a process matrix that is exotic but physically meaningful is the quantum switch [34], see Figure 1.6. Here, one considers two labs A and B that are part of the indefinite causal structure. The environment contains a control system and a target system in state $|\psi\rangle_T$. If the control system is in state $|0\rangle_C$, then the target system is first sent to lab A and afterwards to lab B. If the control system is in state $|1\rangle_C$, then the target system is first sent to lab B and afterwards to lab A. For a quantum system as a control system, it is possible to consider the superposition state $\frac{1}{\sqrt{2}}(|0\rangle_C + |1\rangle_C)$ that combines both paths in a coherent way.

Afterwards, both the control system and the target system get sent to a third lab. The process matrix is [59]

$$W_{\text{switch}} = |w_{\text{switch}}\rangle\langle w_{\text{switch}}| \tag{1.15}$$

$$|w_{\text{switch}}\rangle = \frac{1}{\sqrt{2}}|0\rangle_C|\psi\rangle_{A_I}|\mathbb{1}\rangle\rangle_{A_O B_I}|\mathbb{1}\rangle\rangle_{B_O C'} + \frac{1}{\sqrt{2}}|1\rangle_C|\psi\rangle_{B_I}|\mathbb{1}\rangle\rangle_{B_O A_I}|\mathbb{1}\rangle\rangle_{A_O C'} \tag{1.16}$$

Here, A_I is the input space of A, A_O the output space of A, B_I is the input space of B, B_O the output space of B. C is the space of the control system that the third lab obtains, while C' is the space of the target system that the third lab obtains. $|\mathbb{1}\rangle\rangle = \sum_j |j\rangle|j\rangle$ is an unnormalized maximally entangled state and represents the identity channel between the labs. If A applies the unitary U_A and B applies the unitary U_B, then the third lab receives the state

$$\frac{1}{\sqrt{2}}|0\rangle_C \otimes (U_B U_A|\psi\rangle)_{C'} + \frac{1}{\sqrt{2}}|1\rangle_C \otimes (U_A U_B|\psi\rangle)_{C'} \tag{1.17}$$

Several purification and dilation theorems show that quantum information processing can be described on the level of pure states and unitaries [29]. This is achieved by explicitly assigning ancillary Hilbert spaces to the environment, memories recording the outcomes, or pointer needles. To get back to the usual description, one traces out the ancillary systems or measures them. See e.g. [19, 20] for a list of such results in the finite-dimensional context of quantum information theory. Physically, these results can be interpreted as the expression of two principles: That mixed states are only due to lack of classical knowledge found in the environment. And that closed systems should have reversible evolution. Therefore, if we explicitly describe the involved environment, there should be no mixed states and time evolution should be reversible. As an example, the purification of a mixed state $\rho = \sum_k p_k|k\rangle\langle k|$ is $|\psi\rangle := \sum_k \sqrt{p_k}|k\rangle \otimes |k\rangle$. The ancillary space can be interpreted as memory of which pure state $|k\rangle$ was actually prepared. ρ is obtained by tracing out the ancillary system. As another example, the projective measurement $\{P_k\}_{k=1}^n$ performed on some state $|\phi\rangle$, using an outcome memory initialized in a state $|0\rangle$, can be described via the controlled unitary $\sum_{k=1}^n P_k \otimes U_{0\to k}$, where $U_{0\to k}$ is a unitary on the memory that maps the initialization $|0\rangle$ to the outcome $|k\rangle$.

As the process matrix framework is very general, most processes do not have a physical interpretation. Therefore, it would be helpful to have additional physical postulates for physically reasonable process matrices. One such postulate is the *purification principle* [64]. Here, one introduces a global past and a global future, and describes the agents' operations as unitaries by introducing ancillary systems. These ancillary systems are used as dilation environments and quantum memories to record measurement outcomes. Since we are explicitly describing the relevant environments, we have a closed system and closed systems should have a unitary evolution. Therefore the postulate demands that it should be possible to describe the process as a unitary from the full relevant global past to the full relevant global future, even if there is indefinite causal structure in between. As we explicitly explain the formalization in Chapter 2, we do not provide details here. However, we do point out that the process matrix formalism presented above can be recovered from this purified formulation by measuring the ancillary systems (this replaces the agents' purified unitary operations with quantum instruments), choosing an input for the global past, and a measurement for the global future.

It is known that generalizing the Page-Wootters [65, 66] framework for quantum clocks to several quantum clocks, each of them associated with an agent, can lead to indefinite causal structures [67] if relativistic time dilation is taken into account. Together with Veronika Baumann, Philippe Allard Guérin, and Časlav Brukner I developed a framework and operational protocol that systematically combines process matrices and quantum clocks [68], see Chapter 2.

1.5 Structure of the thesis

This thesis is conceptually separated into two parts. The first part is given by Chapters 2, 3, and 4. In these chapters, we assume the validity of quantum theory. However, here, we consider the formalisms that relax/generalize other aspects of physics, namely the process matrix framework for indefinite causal structure, and quantum reference frames. The second part is given by Chapters 5, 6, and 7. In these chapters, we consider the frameworks that allow us to discuss physics and information processing without assuming the validity of quantum theory, i.e. general/operational probabilistic theories and the device-independence formalism. In detail, this thesis is structured as follows:

Because of the generality of the process matrix [9] framework, many process matrices lack a physical interpretation. One important strategy to develop settings described by exotic process matrices is to combine process matrices with other approaches towards quantum gravity. In [67], Castro-Ruiz et al. consider several Page-Wootters quantum clocks [65, 66], each of them associated with an observer. They consider gravitational scenarios and model the evolution and gravitational interactions of such clocks using constraint operators. Castro-Ruiz et al. observe that the quantum nature of several clocks combined with relativistic time dilation can lead to indefinite causal structure. In particular, they find a setting implementing a gravitational quantum switch. However, as their approach is based on modeling and solving Hamiltonian constraint operators, it is not clear what process matrices can be achieved in such a setting. This motivates us to systematically investigate the probing of indefinite causal structure and extraction of process matrices in settings involving several quantum clocks with time dilation as in Castro-Ruiz et al. [67]. Therefore, in Chapter 2, we develop a systematic approach combining process matrices and discrete quantum clocks associated with observers. We formulate a standardized scenario for probing quantum causal structures: The agents start in a definite causal structure with well-synchronized quantum clocks. Then the agents and their quantum clocks enter a region of quantum or indefinite causality in which they apply their (purified) unitary operations. At last, the agents and their clocks leave this region of indefinite causality and become synchronized again. Instead of using constraint operators, we directly consider the history states and formulate physical postulates that the history states of such probing scenarios should obey. Demanding unitary time evolution and that each agent only uses their local (dilated) unitary operation once at a pre-specified local time, we show that the resulting causal structures are always described by pure process matrices. We investigate extra conditions that our setting implies for the causal reference frames [69] of the agents. We find that the unitary time evolution that the agents see is affine-linear in the operations of the other parties. This allows us to rule out a particular causal

reference frame description proposed in [69] of an exotic process known as (reversed) *Lugano process* [69–72]. We construct a detailed protocol that implements arbitrary coherently controlled causal order within our operational setting. Furthermore, we explain why such protocols cannot be directly adapted to implement the (reversed) Lugano process. I worked on this project together with Veronika Baumann, Philippe Allard Guérin and Časlav Brukner [68].

The process matrix formalism [9] for indefinite causal structure is a generalization of the usual quantum information framework of density matrices and quantum channels. An important setting in quantum information theory is given by parallel schemes in which an agent can simultaneously use several channels or states [20]. For example, two channels might get applied to one half of an entangled quantum system each. Mathematically, the simultaneous/parallel use of channels and density matrices is described by the tensor product. However, it is known that in general the tensor product of two process matrices does not result in a valid process matrix [62]. In Chapter 3, we investigate whether there exists a universal replacement for the tensor product of process matrices. To find such a replacement, we formulate postulates that a reasonable replacement should satisfy: It should be bilinear to be compatible with the interpretation of convex mixtures as statistical mixtures. Process matrices should combine to process matrices. And if the process matrices describe the same definite causal order, the replacement should reduce to the usual tensor product. We prove a no-go theorem that such a universal replacement cannot exist. Our result implies that in the context of indefinite causal structure, parallel schemes simultaneously combining several resources are not available anymore. In particular, this heavily restricts the applicability of the asymptotic regime that is usually considered in quantum information theory. This suggests that information processing in indefinite causal structures might have to focus on the single shot regime. This project is a collaboration of Philippe Allard Guérin, me, Costantino Budroni and Časlav Brukner [63].

A central long term goal of the relational approach to quantum reference frames [73–76] is to develop a completely relational replacement for quantum theory that does not rely on an outside observer. In this approach, also the observers and reference frames are described as quantum systems. Because of this, it is not always clear what this relational approach implies for near-future experiments that are always performed from the perspective of an outside experimenter. What does it mean to jump into the perspective of a particle? Is it experimentally necessary to achieve quantum coherence of actual observers? In Chapter 4, we investigate an operational scenario that clarifies the relation between relational quantum reference frames and experiments with outside observers. We consider an experimenter who performs experiments on a complex composite quantum system. However, this experimenter does not have access to a full external reference frame. Instead, the observer has to use internal relations and properties of the quantum system to make up for the insufficient external reference. We assume that the physical transformations that do not change the internal relations of the composite quantum system form a unitary group. The full Hilbert space that can only be accessed with a full external reference frame is mathematically analogous to the kinematical Hilbert space of constraint quantization [74, 77–79] and the Page-Wootters framework[65, 66], while the unitary symmetry group is analogous to gauge transformations. As a concrete example, we

consider distinguishable point particles on a ring of finitely many positions, as well as a mathematical generalization similar to [80] in which the positions get replaced by elements of a finite Abelian group. The insufficient external reference frame does not provide an absolute, global origin. The symmetry group is assumed to be given by global translations conditioned on the particle relations. That means that the choice of global translation depends on the distances the particles have to each other. We also provide an axiomatic derivation of this symmetry group based on preserving coherence, relative distances between particles and relative distances between quantum states. We perform a very elaborate mathematical characterization of this setting, relying on operator algebras of observables invariant under symmetry transformations. In particular, we show that our scenario contains (discrete analogues of) the quantum reference frames of earlier works [76, 80] as equivalence classes and the corresponding QRF transformations as symmetry transformations. We develop a systematic formalism for how to change particle number in this relational setting, based on preserving the essential properties of operator algebras. This approach shows that the appropriate relational generalization of tensor product and partial trace will in general depend on the choice of algebra of observables. We analyze the so-called *paradox of the third particle* [75], and find that there is a distinguished generalized partial trace that preserves coherence upon removing the third particle, independently of the considered reference frame. My coauthors of this project are Philipp A. Höhn and Markus P. Müller [81].

In Chapter 5, we consider the device-independence formalism [30–33, 52–55]. Usually in the device-independence formalism, the inputs are treated as abstract numbers from a finite set. However, many experiments involve input degrees of freedom related to a continuous spacetime degree of freedom. Examples include angles of polarizers, directions of magnetic fields and time durations of laser pulses. Therefore, we introduce a modification of the usual device independent framework in which the inputs are such continuous spacetime degrees of freedom. As one can imagine that a black box has an external display or knob that allows to choose the value of this spacetime parameter, this setting is compatible with the basic ideas of device independence. An important difference to the scenario with abstract inputs is that the statistics of black boxes with spacetime parameters can respond to spacetime transformations. As a weak semi-device independent assumption, we restrict our attention to finite-dimensional representations of spacetime transformations. This has the consequence that the probabilities are very well-behaved functions of the continuous parameters. In this context, we prove that the statistical response of the black boxes to spacetime transformations is described by linear representations of the transformations. To provide evidence for the conjecture that the structure of quantum theory and spacetime are intimately related to each other, we derive an exact characterization of the quantum (2,2,2)-correlations exclusively in terms of statistical responses of black boxes to rotations. As an important practical example, we consider Bell scenarios with devices that each take an angle as input. In the usual device independent formalism, local hidden variable models only have to work for finitely many input values. However, in our setting, local hidden variable models have to reproduce the statistics of continuously infinitely many possible input values. Therefore it is natural to expect that there exist shapes of correlation functions that do not allow for LHVs, even if the correlations are numerically weak. We show that this intuitive expectation is wrong. Specifically, for arbitrary weakly fluctuating

correlation functions of two angles, we explicitly construct local hidden variable models that reproduce these correlations for all choices of angles. At the other extreme, we investigate Bell non-locality for angles as local parameters. We derive a criterion that shows that strongly fluctuating relational correlation functions violate a Braunstein-Caves inequality [82]. Based on this criterion, we develop a semi-device independent Bell witness protocol in which one of the parties does not have to freely generate an input. I collaborated with Andrew J. P. Garner and Markus P. Müller on this project [56].

In Chapter 6, we analyze a particular class of post-quantum theories. Many derivations of quantum theory from physical principles prove as a first step that two-level systems must be described by a ball-shaped state space, see e.g. [42, 43, 83]. As a consequence, there exist reasonable axiomatizations of Bloch balls of unspecified dimension. This observation motivates to consider theories built from ball-shaped state spaces of arbitrary dimension. For quantum systems, the Bloch ball is three-dimensional because rotations of Bloch vectors couple to rotations in physical space. In that sense, higher dimensional Bloch balls can be seen as a particular approach to adapt quantum theory to a spacetime in which space has more than 3 dimensions [83, 84]. However, it is known that two Bloch balls of dimension other than three do not allow for bipartite interactions that are given by reversible time evolutions [43, 84]. Therefore, any reversible time evolution describing interactions of higher dimensional Bloch balls would have to be a genuine multipartite interaction that cannot be reduced to two-body interactions. Indeed, Dakić and Brukner [84] conjectured that such genuine multipartite interactions of more than two higher-dimensional Bloch balls exist. Markus P. Müller and I investigated this possibility in the context of a popular postulate called *tomographic locality* [24]. The principle says that composite systems are completely characterized by their local measurement statistics and correlations between local measurements. In [85] it was shown that theories locally described by 3-dimensional Bloch balls (in particular theories that are locally quantum) can only have reversible time evolution describing non-trivial interactions between an arbitrary amount of exchangeable 3-dimensional Bloch balls if the theory is already full quantum theory. In Chapter 6, we generalize the corresponding mathematical framework and proof strategy of de la Torre et al. [85] to arbitrary dimension. We use it to prove that for all Bloch ball dimensions different from three, no time-parametrized reversible interactions are possible, no matter how many higher dimensional Bloch balls are involved. Together with the result of [85], our result means that only quantum theory can have interactions (bipartite or multipartite) between identical ball shaped state spaces of arbitrary dimension given by reversible time evolution. My supervisor Markus P. Müller and I developed this project [50].

As explained before, GPTs/OPTs [23–29] constitute a framework that allows to develop operational theories of physics that are neither classical nor quantum. Nonetheless, every theory needs to satisfy certain requirements to be an acceptable scientific theory. One such requirement is objectivity (or, rather, inter-subjectivity) of measurement results. This means that agents who observe the same experiment should see the same outcome data. One mechanism that explains why this is the case in quantum physics is called *quantum Darwinism* [47, 48]. In the ideal case, a CNOT-like interaction between a measured system and its environment systems leads

to perfect outcome correlation of the system and its environment. In Chapter 7, we investigate generalizations of this ideal quantum Darwinism case in GPTs/OPTs. We provide both necessary and sufficient conditions for this mechanism to be applicable. As necessary conditions, we show that theories satisfying certain non-classicality conditions need to have both entangled states and entangled measurements. For the sufficiency part, we show how the ideal Darwinism mechanism relates to other important postulates and mechanisms in GPTs/OPTs. Specifically, we show that the Darwinism mechanism is enabled if the theory satisfies certain decoherence postulates [45] or symmetry postulates motivated by embedding classical computation. To provide a specific non-quantum example, we argue that Spekkens' toy theory [86] allows for a generalized CNOT gate enabling the ideal Darwinism process. I worked on this project together with Roberto D. Baldijão, Andrew J. P. Garner and Markus P. Müller [49].

Every chapter has its own list of references/bibliography.

1.6 Bibliography of Chapter 1 (*Introduction*)

[1] J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*. Cambridge University Press, 2017.

[2] M. Tsamparlis, *Special Relativity: An Introduction with 200 Problems and Solutions*. Springer, 2010.

[3] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[4] C. Rovelli, *Quantum Gravity*. Cambridge University Press, Cambridge, 2004.

[5] D. Oriti, ed., *Approaches to quantum gravity*. Cambridge University Press, Cambridge, 2009.

[6] C. Kiefer, *Quantum Gravity*. Oxford University Press, Oxford, 3rd ed., 2012.

[7] R. M. Wald, *General relativity*. Chicago Univ. Press, Chicago, IL, 1984.

[8] G. Tóth and I. Apellaniz, "Quantum metrology from a quantum information science perspective," *Journal of Physics A: Mathematical and Theoretical* **47**, 424006 (2014).

[9] O. Oreshkov, F. Costa, and C. Brukner, "Quantum correlations with no causal order," *Nature Commun.* **3**, 1092 (2012), `arXiv:1105.4464 [quant-ph]`.

[10] L. Hardy, "Probability Theories with Dynamic Causal Structure: A New Framework for Quantum Gravity." arXiv:gr-qc/0509120, 2005.

[11] L. Hardy, "Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure," *J. Phys. A: Math. Theor.* **40**, 3081–3099 (2007).

[12] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Theoretical framework for quantum networks," *Physical Review A* **80**, 022339 (2009).

[13] A. Kent, "Summoning, No-Signalling and Relativistic Bit Commitments," *Entropy* **21**, 534 (2019).

[14] A. Valentini, "Non-local correlations in quantum electrodynamics," *Physics Letters A* **153**, 321–325 (1991).

[15] A. R. H. Smith, "The Unruh-DeWitt Detector and Entanglement Harvesting," in *Detectors, Reference Frames, and Time*, pp. 17–40.
Springer, 2019.

[16] I. Fuentes, "Lecture series on relativistic quantum information," in *Diversities in Quantum Computation and Quantum Information*, pp. 107–147.
World Scientific, 2013.

[17] S. Milz and K. Modi, "Quantum stochastic processes and quantum non-Markovian phenomena." arXiv:2012.01894, 2020.

[18] S. Milz, F. A. Pollock, and K. Modi, "An introduction to operational quantum dynamics," *Open Systems & Information Dynamics* **24**, 1740016 (2017).

[19] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*.
Cambridge University Press, 2010.

[20] M. M. Wilde, *Quantum Information Theory*.
Cambridge University Press, 2 ed., 2017.

[21] J. S. Bell, "On the Einstein Podolsky Rosen paradox," *Physics Physique Fizika* **1**, 195–200 (1964).

[22] J. Clauser, M. Horne, A. Shimony, and R. Holt, "Proposed Experiment to Test Local Hidden-Variable Theories," *Physical Review Letters* **23**, 880–884 (1969).

[23] M. Plávala, "General probabilistic theories: An introduction." arXiv:2103.07469, 2021.

[24] J. Barrett, "Information processing in generalized probabilistic theories," *Physical Review A* **75**, 032304 (2007).

[25] L. Hardy, "Quantum Theory from Five Reasonable Axioms." arXiv:quant-ph/0101012, 2001.

[26] B. Dakic and C. Brukner, "Quantum Theory and Beyond: Is Entanglement Special?." In "Deep Beauty: Understanding the Quantum World through Mathematical Innovation", Ed. H. Halvorson, Cambridge University Press, Cambridge, 2011.

[27] L. Masanes and M. P. Müller, "A derivation of quantum theory from physical requirements," *New Journal of Physics* **13**, 063001 (2011).

[28] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Informational derivation of quantum theory," *Physical Review A* **84**, 012311 (2011).

[29] G. Chiribella, G. M. D' Ariano, and P. Perinotti, "Quantum theory, namely the pure and reversible theory of information," *Entropy* **14**, 1877–1893 (2012).

[30] J. Barrett, L. Hardy, and A. Kent, "No Signaling and Quantum Key Distribution," *Physical Review Letters* **95**, 010503 (2005).

[31] A. Acin, S. Massar, and S. Pironio, "Efficient quantum key distribution secure against no-signalling eavesdroppers," *New Journal of Physics* **8**, 126 (2006).

[32] A. Acín, N. Gisin, and L. Masanes, "From Bell's Theorem to Secure Quantum Key Distribution," *Physical Review Letters* **97**, 120405 (2006).

[33] U. Vazirani and T. Vidick, "Fully Device-Independent Quantum Key Distribution," *Physical Review Letters* **113**, 140501 (2014).

[34] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, "Quantum computations without definite causal structure," *Phys. Rev. A* **88**, 022318 (2013).

[35] M. Zych, *Quantum Systems under Gravitational Time Dilation.* Springer Theses. Springer, 2017.

[36] M. Zych, F. Costa, I. Pikovski, and Č. Brukner, "Bell's theorem for temporal order," *Nature Communications* **10**, 3772 (2019).

[37] R. Webster, *Convexity.* Oxford University Press, Oxford, 1994.

[38] P. Janotta and R. Lal, "Generalized probabilistic theories without the no-restriction hypothesis," *Phys. Rev. A* **87**, 052131 (2013).

[39] S. Massar and M. K. Patra, "Information and communication in polygon theories," *Phys. Rev. A* **89**, 052124 (2014).

[40] C. Pfister and S. Wehner, "An information-theoretic principle implies that any discrete physical theory is classical," *Nature Communications* **4**, 1851 (2013).

[41] T. Heinosaari, L. Leppäjärvi, and M. Plávala, "No-free-information principle in general probabilistic theories," *Quantum* **3**, 157 (2019).

[42] L. Masanes, M. P. Müller, R. Augusiak, and D. Pérez-García, "Existence of an information unit as a postulate of quantum theory," *Proceedings of the National Academy of Sciences* **110**, 16373–16377 (2013).

[43] L. Masanes, M. P. Müller, D. Pérez-García, and R. Augusiak, "Entanglement and the three-dimensionality of the Bloch ball," *Journal of Mathematical Physics* **55**, 122203 (2014).

[44] M. D. Mazurek, M. F. Pusey, K. J. Resch, and R. W. Spekkens, "Experimentally Bounding Deviations From Quantum Theory in the Landscape of Generalized Probabilistic Theories," *PRX Quantum* **2**, 020302 (2021).

[45] J. G. Richens, J. H. Selby, and S. W. Al-Safi, "Entanglement is Necessary for Emergent Classicality in All Physical Theories," *Phys. Rev. Lett.* **119**, 080503 (2017).

[46] M. P. Müller and C. Ududec, "Structure of Reversible Computation Determines the Self-Duality of Quantum Theory," *Phys. Rev. Lett.* **108**, 130401 (2012).

[47] W. H. Zurek, "Relative States and the Environment: Einselection, Envariance, Quantum Darwinism, and the Existential Interpretation." Los Alamos preprint LAUR 07-4568, arXiv:0707.2832, 2007.

[48] W. H. Zurek, "Quantum Darwinism," *Nature Physics* **5**, 181–188 (2009).

[49] R. D. Baldijão, M. Krumm, A. J. P. Garner, and M. P. Müller, "Quantum Darwinism and the spreading of classical information in non-classical theories." arXiv:2012.06559, 2020.

[50] M. Krumm and M. P. Müller, "Quantum computation is the unique reversible circuit model for which bits are balls," *npj Quantum Information* **5**, 7 (2019).

[51] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-A. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger, "Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons," *Phys. Rev. Lett.* **115**, 250401 (2015).

[52] M. Pawłowski and N. Brunner, "Semi-device-independent security of one-way quantum key distribution," *Physical Review A* **84**, 010302 (2011).

[53] D. Rusca, T. Van Himbeeck, A. Martin, J. B. Brask, W. Shi, S. Pironio, N. Brunner, and H. Zbinden, "Self-testing quantum random-number generator based on an energy bound," *Physical Review A* **100**, 062338 (2019).

[54] T. van Himbeeck, E. Woodhead, N. J. Cerf, R. García-Patrón, and S. Pironio, "Semi-device-independent framework based on natural physical assumptions," *Quantum* **1**, 33 (2017).

[55] T. Van Himbeeck and S. Pironio, "Correlations and randomness generation based on energy constraints." arXiv:1905.09117, 2019.

[56] A. J. P. Garner, M. Krumm, and M. P. Müller, "Semi-device-independent information processing with spatiotemporal degrees of freedom," *Physical Review Research* **2**, 013112 (2020).

[57] J. Pearl, *Causality: Models, reasoning, and inference*. Cambridge University Press, 2000.

[58] A. Feix, M. Araújo, and Č. Brukner, "Quantum superposition of the order of parties as a communication resource," *Physical Review A* **92**, 052326 (2015).

[59] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, "Witnessing causal nonseparability," *New Journal of Physics* **17**, 102001 (2015).

[60] A. Jamiołkowski, "Linear transformations which preserve trace and positive semidefiniteness of operators," *Reports on Mathematical Physics* **3**, 275–278 (1972).

[61] M.-D. Choi, "Completely positive linear maps on complex matrices," *Linear algebra and its applications* **10**, 285–290 (1975).

[62] D. Jia and N. Sakharwade, "Tensor products of process matrices with indefinite causal structure," *Phys. Rev. A* **97**, 032110 (2018).

[63] P. A. Guérin, M. Krumm, C. Budroni, and Č. Brukner, "Composition rules for quantum processes: a no-go theorem," *New Journal of Physics* **21**, 012001 (2019).

[64] M. Araújo, A. Feix, M. Navascués, and Č. Brukner, "A purification postulate for quantum mechanics with indefinite causal order," *Quantum* **1**, 10 (2017).

[65] D. N. Page and W. K. Wootters, "Evolution without evolution: Dynamics described by stationary observables," *Physical Review D* **27**, 2885–2892 (1983).

[66] W. K. Wootters, ""Time" Replaced by Quantum Correlations," *Int. J. Theor. Phys.* **23**, 701 (1984).

[67] E. Castro-Ruiz, F. Giacomini, A. Belenchia, and C. Brukner, "Quantum clocks and the temporal localisability of events in the presence of gravitating quantum systems," *Nature Commun.* **11**, 2672 (2020), `arXiv:1908.10165 [quant-ph]`.

[68] V. Baumann, M. Krumm, P. A. Guérin, and Č. Brukner, "Page-Wootters formulation of indefinite causal order." arXiv:2105.02304, 2021.

[69] P. A. Guérin and Č. Brukner, "Observer-dependent locality of quantum events," *New Journal of Physics* **20**, 103031 (2018).

[70] Ä. Baumeler and S. Wolf, "Perfect signaling among three parties violating predefined causal order," in *2014 IEEE International Symposium on Information Theory (ISIT)*, pp. 526–530.
IEEE, Piscataway, NJ, June, 2014.

[71] Ä. Baumeler, A. Feix, and S. Wolf, "Maximal incompatibility of locally classical behavior and global causal order in multi-party scenarios," *Phys. Rev. A* **90**, 042106 (2014).

[72] Ä. Baumeler and S. Wolf, "The space of logically consistent classical processes without causal order," *New J. Phys.* **18**, 013036 (2016).

[73] Y. Aharonov and T. Kaufherr, "Quantum frames of reference," *Physical Review D* **30**, 368 (1984).

[74] C. Rovelli, "Quantum reference systems," *Classical and Quantum Gravity* **8**, 317 (1991).

[75] R. M. Angelo, N. Brunner, S. Popescu, A. J. Short, and P. Skrzypczyk, "Physics within a quantum reference frame," *Journal of Physics A: Mathematical and Theoretical* **44**, 145304 (2011).

[76] F. Giacomini, E. Castro-Ruiz, and Č. Brukner, "Quantum mechanics and the covariance of physical laws in quantum reference frames," *Nature Communications* **10**, 494 (2019).

[77] M. Henneaux and C. Teitelboim, *Quantization of gauge systems*. Princeton University Press, 1992.

[78] P. A. Dirac, *Lectures on Quantum Mechanics*. Yeshiva University Press, 1964.

[79] A. Vanrietvelde, P. A. Hoehn, F. Giacomini, and E. Castro-Ruiz, "A change of perspective: switching quantum reference frames via a perspective-neutral framework," *Quantum* **4**, 225 (2020).

[80] A.-C. de la Hamette and T. D. Galley, "Quantum reference frames for general symmetry groups," *Quantum* **4**, 367 (2020).

[81] M. Krumm, P. A. Höhn, and M. P. Müller, "Quantum reference frame transformations as symmetries and the paradox of the third particle." arXiv:2011.01951, 2020.

[82] S. L. Braunstein and C. M. Caves, "Wringing out better Bell inequalities," *Annals of Physics* **202**, 22–56 (1990).

[83] M. P. Mueller and L. Masanes, "Three-dimensionality of space and the quantum bit: an information-theoretic approach," *New Journal of Physics* **15**, 053040 (2013).

[84] B. Dakić and Č. Brukner, "The classical limit of a physical theory and the dimensionality of space." in "Quantum Theory: Informational Foundations and Foils", eds. G. Chiribella and R. W. Spekkens, Springer, 2016.

[85] G. de la Torre, L. Masanes, A. J. Short, and M. P. Müller, "Deriving quantum theory from its local structure and reversibility," *Physical Review Letters* **109**, 090403 (2012).

[86] R. W. Spekkens, "Evidence for the epistemic view of quantum states: A toy theory," *Physical Review A* **75**, 032110 (2007).

Chapter 2

Page-Wootters formulation of indefinite causal order

Authors: Veronika Baumann*, Marius Krumm*, Philippe Allard Guérin, Časlav Brukner
Submitted to *Physical Review Research* for review (18th of May, 2021).
Preprint available as: arXiv:2105.02304
https://arxiv.org/abs/2105.02304

Contribution: Veronika Baumann and I share first authorship. I made essential contributions to all parts of the research. I provided the leading input for developing the protocol to implement arbitrary coherently controlled causal order, and for the result that the causal past of an agent has to be affine-linear in the operations of the other agents.

Abstract: One of the most fundamental open problems in physics is the unification of general relativity and quantum theory to a theory of quantum gravity. An aspect that might become relevant in such a theory is that the dynamical nature of causal structure present in general relativity displays quantum uncertainty. This may lead to a phenomenon known as indefinite or quantum causal structure, as captured by the process matrix framework. Due to the generality of that framework, however, for many process matrices there is no clear physical interpretation. A popular approach towards a quantum theory of gravity is the Page-Wootters formalism, which associates to time a Hilbert space structure similar to spatial position. By explicitly introducing a quantum clock, it allows to describe time-evolution of systems via correlations between this clock and said systems encoded in history states. In this paper we combine the process matrix framework with a generalization of the Page-Wootters formalism in which one considers several observers, each with their own discrete quantum clock. We describe how to extract process matrices from scenarios involving such observers with quantum clocks, and analyze their properties. The description via a history state with multiple clocks imposes constraints on the physical implementation of process matrices and on the perspectives of the observers as described via causal reference frames. While it allows for describing scenarios where different definite causal orders are coherently controlled, we explain why certain non-causal processes might not be implementable within this setting.

Note on changes: Compared to the preprint-version arXiv:2105.02304v1 of this paper, I applied several larger changes to the version of the paper presented in this thesis. These changes do not change the physical results, but they add clarifications and more detailed

explanations.

- Replaced a motivational paragraph in the introduction (Section I).
- Towards the end of the introduction in Section I (both in the preprint and the thesis), a summary of the approach of our paper is given. This summary explains that we do not start by solving constraint operators, but directly use history states. And that we consider finite-dimensional instead of continuous clocks. In this summary, added an explicit reference to the paper of Castro-Ruiz et al. (2020) to make clear that these are some of the crucial points that make our approach different from their approach.
- In Section III, for the relation between unitary time evolution and the projection operator for circuits, added an explicit assumption that $t_2 > t_1$.
- Made the discussion of the motivation for normalization operators in the main text more detailed. (Section IV A).
- Explicitly pointed out the assumption that our normalization operators and time evolution operators continue to be valid if the ancillary systems are initialized to other states (Section IV A).
- Added an explicit calculation that shows that the process is indeed a multi-linear unitary, as claimed in the preprint version (Section IV B).
- For the result that the perspectival time evolutions must be affine-linear in the unitaries of the other agents, added explicit calculations, detailed arguments, and clarifications (Section IV C).
- In Section IV D, changed a formulation to make clear that we do not know whether history states are orthogonal if their input states are.
- In the context of coherent control of causal order (Section V C and Appendix C), gave more detailed explanations of quantum combs and their purification. In the main text (Section V C), added an explicit example for how comb dimensions can be compatible, although the protocol works independently of how the combs are made compatible. In the appendix, added a detailed discussion about how to treat controlled superpositions of combs whose dimensions are not directly compatible. Rearranged Figure 7 such that the individuals parts can be larger, while fitting the page width. Added more explanations for the formal details of the protocol. Added more steps with just synchronized clock steps in between the protocol phases to make it more obvious that the clock freezes do not overlap with the application of the combs.
- Expanded the discussion in the conclusions (Section VI).
- Clarified Appendix B, i.e. the discussion about the physical projector and the perspectival unitaries: Pointed out that the normalization operators do not play an important role for the discussion, and emphasized that the specific choice of counter example depends on the continuation of the operators T_i'.
- Added a reference to a paper by Bavaresco et al. (arXiv:2105.13369) which also considers quantum controlled combs.
- Fixed several typos, grammatical mistakes and clarified a few smaller formulations.

I. INTRODUCTION

Indefinite causal structure is an extension of the usual notion of causal structure that is expected to become relevant in quantum gravity: In general relativity, causal structure is dynamical instead of fixed and attributing quantum properties [1–8] would imply the existence of exotic causal structures, like superpositions of space-times and superpositions of the order of events. The process matrix framework [9, 10] was invented to systematically describe such indefinite or quantum causal structures. However, many processes that arise in this framework have no clear physical interpretation and it is not known which of them are realizable in nature. It has, therefore, been suggested that only processes, which reversibly map a well defined casual past to a well defined casual future with possibly indefinite causal order in between, are physical [10]. For such processes, it has been shown that one can always find a causal reference frame that represents the perspective of an agent or observer within the causal structure [11]. While the observer's event is local in their causal frame of reference, the events of other observers may be "smeared" over the causal past and future of the event. Still the question which processes are realizable in nature remains open.

A crucial obstacle in finding a complete theory of quantum gravity is caused by the different role of time in general relativity and quantum theory. As an approach to bridge this conceptual gap, one can use a timeless formalism [12–22], which we refer to as the Page-Wootters formalism in this paper. In this formalism, one also associates a Hilbert space with time, which can be interpreted as a quantum clock. One describes the physics of the extended system including the clock by using history states which are obtained via a Wheeler-DeWitt-like equation using a constraint operator. These history states encode dynamics as correlations between the main system and the quantum clock.

In Ref. [22], the authors considered a generalized Page-Wootters approach using several clocks. The authors found that history states arising from solving a Hamiltonian constraint for gravitationally interacting clocks can give rise to indefinite causal order and studied the time evolution according to the perspectives of different clocks. In particular, they showed how the Page-Wootters formalism can recover the so-called gravitational quantum switch [5]. Their approach works for important examples, but it is not clear in general which process (if any) is implemented by a given history state, or what is the set of non-causal processes that can be implemented within such a framework.

Some non-causal processes can violate device-independent causal inequalities [23–25], although no physical interpretation for such processes are known. Other processes, for example the so-called quantum switch [26–28], where the order of operations is controlled by a quantum system, cannot violate causal inequalities but exhibit indefinite causal order that can be identified by causal witnesses [29]. Moreover, it is possible to experimentally implement such coherent quantum control of causal order [27, 28, 30–35].

The Page-Wootters formalism can be regarded as an independent formulation of quantum theory, similar to the path-integral formulation. It is known that any quantum circuit (i.e. a definite causal order) can be implemented within the Page-Wootters formalism as a Feynman's quantum computer and a single (i.e. global) quantum clock [36–39]. As the Page-Wootters formalism is used as an approach to solve the problem of time in quantum gravity, and the process matrix formalism was invented to describe indefinite causal structures that may arise in quantum gravity, it is natural to combine these two approaches towards quantum gravity. What process matrices are compatible with a local perception of time modeled via Page-Wootters quantum clocks?

To approach this question, in the present work, we present a general definition of what it means for a history state to implement a pure process matrix [10], for the case of finite dimensional systems and several clocks. We describe how to extract the agents' perspectives from the history states, which corresponds to a refinement of the concept of causal reference frames [11] that explicitly includes the quantum clocks. We show that arbitrary coherently controlled causal order can be extracted from our framework when different clocks tick at different rates (for example, due to time dilation effects). Moreover, we analyze the additional restrictions that the history states impose on the extracted process matrices and propose that these restrictions might be regarded as reasons why some processes cannot be implemented in nature. Thus, while the Page-Wootters formalism with several clocks can enable the extraction of processes with definite causal order and quantum controlled causal order, it additionally provides insights into why some processes might not be realizable within the framework. Contrary to Castro-Ruiz et al. [22], our approach does not start by defining a constraint operator and solving it; instead we work directly at the level of history states. We consider discrete instead of continuous clocks because this allows us to express the perspectives of the agents using circuits. We develop a systematic framework that combines process matrices and Page-Wootters history states with several discrete clocks. We describe how to model scenarios where these clocks are associated with agents or observers that are initially all part of a definite space-time causal structure. Then the agents might enter a "region" of quantum causal structure, where the global order of events is no longer well defined. At the end, however, all agents return to

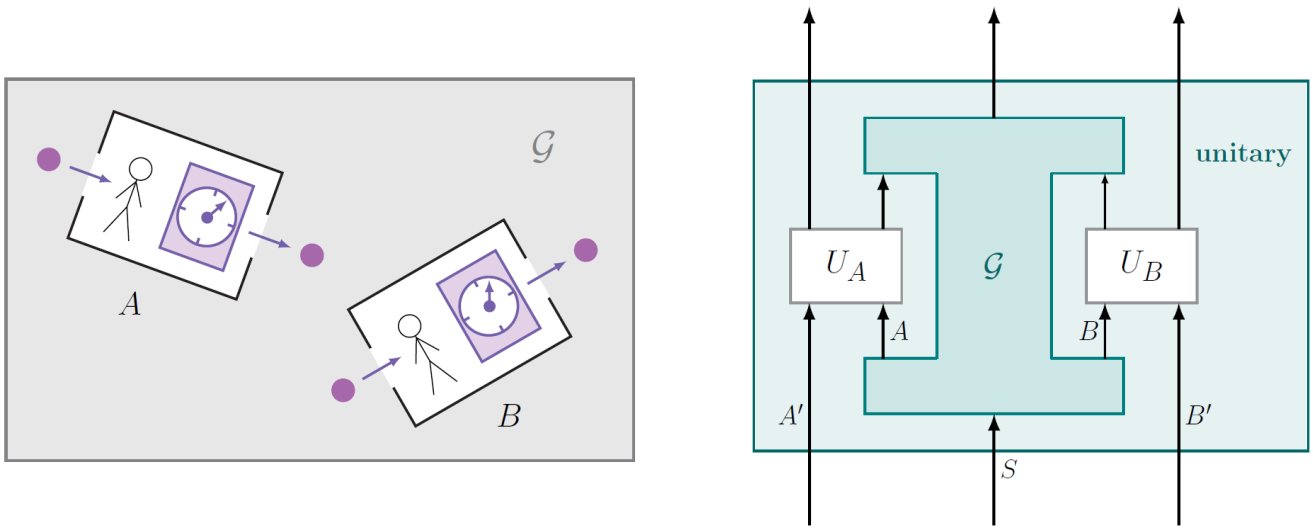
a definite causal structure.

The paper is structured as follows: We first recapitulate important aspects of the process matrix formalism (including causal reference frames) and the Page-Wootters formalism in Sections II and III before we motivate and introduce our framework, which combines these two approaches, in detail in Section IV. In that context, we derive several mathematical properties of our setting, in particular restrictions on process matrices. In Section V, we first construct examples involving varying clock speeds and indefinite causal structure before we explain how to implement arbitrary quantum-controlled causal order in our setting. At the end of Section V we discuss why another well-known, non-causal process might not be implementable in our framework. Finally, we discuss our findings in Section VI.

II. PROCESS MATRICES AND CAUSAL REFERENCE FRAMES

In this section we give a short introduction to the operational setting of the process matrix formalism and explain the parts of the framework that are important for the rest of the paper.

The basic operational setting of the process matrix formalism concerns several agents (here N of them), labeled $A_1 \dots A_N$, each of them inside their own (small) lab where the usual rules of quantum theory are valid. The outside "environment", which relates the various agents, is not assumed to be causally definite, for example it could be a superposition of space-time structures. During the protocol, each agent receives a quantum system from the "environment", applies a quantum instrument i.e. a probabilistic quantum channel (for example a measurement or a pure quantum channel) to that system and then sends it out again. This well defined local time ordering inside the lab can be thought of as being tracked by a clock associated with each agent, the bipartite case is depicted in Figure 1a. The process matrix $\mathcal{G}$ is the mathematical object that encodes the observed outcome probabilities for any choice of local quantum instruments. Process matrices describing definite causal order, i.e. the global order of operations performed by different agents is well defined, are equivalent to higher order quantum maps or quantum combs [40–44]. In general, however, they allow scenarios with indefinite causal order, where no such global order exists, and are in that sense generalizations of quantum combs.



(a) Basic setting

(b) A pure process

FIG. 1: Example of bipartite processes: The two agents (here called A and B) are each situated in their own lab. Each agent obtains a system from the environment, acts on it with a quantum instrument and then sends it out again. While inside the labs the order of events is well defined, there need not be a well defined global ordering imposed by the environment. The outcome statistics of the operations performed by A and B is described by a process matrix $\mathcal{G}$, see (1a). These quantum instruments of A and B can be represented as unitaries U_A, U_B by introducing ancillary systems A', B'. A pure process $\mathcal{G}$ is a (multilinear) supermap that gives an induced unitary transformation on $S \otimes A' \otimes B'$ when the agents are applying unitary operations U_A, U_B, see (1b).

In [10], processes (called quantum superchannels in [44]) are formalized as maps from a global past to a global future, that depend on the agents' operations. In addition to the agent's systems, whose Hilbert spaces are labelled $A_1 \ldots, A_N$, we introduce ancillary systems $A'_1, \ldots, A'_N$. The ancillary system can be used, for example, as a quantum memory recording a measurement outcome. Each agent is allowed to act with a quantum channel on their system and their ancilla, i.e. a completely positive trace-preserving map (CPTP) $\mathcal{C}_{A_j} : \mathcal{L}(A_j A'_j) \to \mathcal{L}(A_j A'_j)$. One assumes that the ancillas have trivial evolution except when the respective operations of the agents are applied. Then, a *process* (or *quantum superchannel*) is a multilinear map $\mathcal{G}$ that maps the agents' quantum channels to a quantum channel, while acting as the identity on all the ancillary systems (just as in Figure 1b, but with the unitaries replaced by quantum channels). This map encodes the causal structure given by the environment.

In this work, we only consider pure processes. Using Stinespring's dilation theorem [45, 46] one can represent the quantum operations of the agents as unitaries $U_{A_1} \ldots U_{A_N}$ acting on the respective system and ancilla that the agents obtain from the environment. The ancillas serve as both purifying systems and as memories recording the outcomes. We say that a process $\mathcal{G}$ is a pure process if it is a unitary preserving map, i.e. $\mathcal{G}(U_{A_1} \ldots U_{A_N})$ is unitary for any unitaries $U_{A_1} \ldots U_{A_N}$, while acting as the identity on the ancillary systems. The basic mathematical structure of bipartite pure process is depicted in Figure 1b.

Since many non-causal process matrices lack a clear interpretation and it is not clear whether they are compatible with the known physical laws, it has been suggested that only purifiable processes, which means they can be obtained from pure processes, are physical [10]. Such processes can be regarded as reversible transformations from a well defined casual past to a well defined casual future, with indefinite causal order in between. We note that the quantum-switch is an example of a pure process, as shown explicitly in [10], and it is physically realizable either in gravitational [5, 22] or optical setups [27, 28, 30–35].

The notion of causal reference frames [11] was introduced as an equivalent description of the pure process matrix formalism. The causal reference frame represents the perspective of an agent inside a (possibly indefinite) causal structure. More concretely, one imagines the perspective of an agent, say A_1, as follows: The crucial moment for agent A_1 is when he or she applies unitary U_{A_1}. The evolution starting from the beginning of the protocol up to that moment is described by a unitary $\Pi_{A_1}(U_{A_2} \ldots U_{A_N})$, which is called the causal past of A_1 and can depend on the instruments of all other agents. Then A_1 enforces time evolution via U_{A_1} on the input to his or her lab and the ancilla A'_1, while all other degrees of freedom evolve in an uncorrelated way. The evolution of these other degrees of freedom can be absorbed into $\Pi_{A_1}(U_{A_2} \ldots U_{A_N})$ such that without loss of generality we can assume that during A_1's time of action, evolution is given by $U_{A_1} \otimes \mathbb{1}$. Afterwards the evolution up to the end of the protocol is described by a unitary $\Phi_{A_1}(U_{A_2} \ldots U_{A_N})$, which is called the causal future of A_1. It can again depend on the instruments of all other agents. As shown in Ref. [11] all pure processes admit a decomposition in causal reference frames, i.e. if $\mathcal{G}$ is a pure process, then $\mathcal{G}$ can be written as

$$\mathcal{G}(U_{A_1} \ldots U_{A_N}) = \Phi_{A_1}(U_{A_2} \ldots U_{A_N}) \, (U_{A_1} \otimes \mathbb{1}) \, \Pi_{A_1}(U_{A_2} \ldots U_{A_N}), \tag{1}$$

where $\Phi_{A_1}(U_{A_2} \ldots U_{A_N}), \Pi_{A_1}(U_{A_2} \ldots U_{A_N})$ are unitaries that depend on all unitaries other than U_{A_1} and that describe the time-evolution according to A_1's point of view. A similar decomposition exists from the point of view of all other agents. In the present work we take a similar approach, but we make the addition of a localized quantum clock associated to each observer, and explain how the perspectives of various agents can arise from a perspective neutral history state as given by the Page-Wootters formalism.

III. THE PAGE-WOOTTERS FORMALISM

In this section we give a brief general overview of the Page-Wootters formalism for continuous as well as discrete quantum clocks. A possible justification for considering quantum clocks with discrete Hilbert spaces comes from arguments involving the Bekenstein bound [47] that Hilbert space is fundamentally finite-dimensional [48, 49]. Also, it can be argued that all information that can ever be acquired via measurements is finite and that therefore on the fundamental level physics should be discrete as well and indeed, finite [50]. Furthermore, we get significant technical simplifications due to the fact that for finite-dimensional Hilbert spaces, the physical Hilbert space is a subspace of the kinematical Hilbert space, while this is not the case in the infinite-dimensional case [20, 51]. Most importantly for our purpose, the assumption of finite dimensional clocks allows us to get a physical picture of indefinite causal structure in form of generalizations of quantum circuits.

In addition to the usual system Hilbert space $\mathcal{H}_S$ the Page-Wootters formalism introduces an additional Hilbert space $\mathcal{H}_c$ associated with time that can be interpreted as an ideal quantum clock. In analogy to position in non-relativistic quantum mechanics $\mathcal{H}_c$ can be chosen to be spanned by square integrable functions on the real line; informally it is common in physics to imagine this Hilbert space as $\mathcal{H}_c = \text{span}\{|t\rangle \mid t \in \mathbb{R}\}$. In analogy to the usual momentum operator, one can define an operator $\hat{p}_t$ as the generator of translations on $\mathcal{H}_c$. In the time representation, i.e. $\langle t|\psi\rangle$, it is given by $\hat{p}_t = -i\frac{\partial}{\partial t}$. Let $\hat{H}_S$ be the Hamiltonian of the system and consider the constraint operator $\hat{C} := \hat{p}_t + \hat{H}_S$. Let $|\Psi\rangle\rangle$ be a state on $\mathcal{H}_c \otimes \mathcal{H}_S$ that satisfies the Wheeler-DeWitt-like constraint equation $\hat{C}|\Psi\rangle\rangle = 0$. Such states $|\Psi\rangle\rangle$ are often called physical states. Without worrying about normalizability, one can formally expand $|\Psi\rangle\rangle$ by using the time basis as

$$|\Psi\rangle\rangle = \int \mathrm{d}t\ |t\rangle \otimes |\psi(t)\rangle. \tag{2}$$

With this expansion it becomes clear why states $|\Psi\rangle\rangle$ are also called history states: For each time t, they encode a system state $|\psi(t)\rangle$ and an ordered time sequence $t_0 < t_1 < t_2$ corresponds to the history of the state given by $|\psi(t_0)\rangle, |\psi(t_1)\rangle$ and $|\psi(t_2)\rangle$. Plugging the expansion Eq. (2) into the constraint equation, one can show that the system state satisfies

$$i\frac{\partial}{\partial t}|\psi(t)\rangle = H_S|\psi(t)\rangle, \tag{3}$$

which is the standard Schrödinger equation. Therefore, this approach recovers the usual quantum formalism. In general, solutions to the constraint equation can be obtained via an operator

$$\hat{P} := \int_{\mathbb{R}} ds\, e^{-is\hat{C}}, \tag{4}$$

which gives a valid physical (or history) state, i.e. solution to the constraint equation $\hat{C}(\hat{P}|\phi\rangle) = 0$, when applied to arbitrary states $|\phi\rangle \in \mathcal{H}_c \otimes \mathcal{H}_S$. For this reason the operator $\hat{P}$ is sometimes called the physical projector [52], although it is not a projector in the strict mathematical sense. Moreover, $\langle t_2|\hat{P}|t_1\rangle = \mathcal{U}(t_2, t_1)$ is a unitary operator on $\mathcal{H}_S$ and in case of there being no interaction term between clock and system, i.e. $\hat{C} = \hat{H}_S + \hat{p}_t$, it can be shown that it gives the time evolution according to the Schrödinger equation, i.e. $\langle t_2|\hat{P}|t_1\rangle = e^{-i(t_2-t_1)H_S}$.

The Page-Wootters formalism has been adapted to regular (i.e. causal) quantum circuits, see for example [36–39]. It uses one finite dimensional quantum clock and is described by the constraint equation

$$\hat{C}|\Psi\rangle\rangle = \sum_t \hat{H}_t|\Psi\rangle\rangle = 0, \tag{5}$$

where the Hamiltonians

$$\hat{H}_t = -\frac{1}{2}\left(|t\rangle\langle t-1| \otimes U_t + |t-1\rangle\langle t| \otimes U_t^\dagger - |t-1\rangle\langle t-1| - |t\rangle\langle t|\right), \tag{6}$$

can be understood as making the clock tick once and applying some unitary U_t to the system. In other words, at time t the circuit applies gate U_t. Solutions to Eq. (5) are history states of this quantum circuit in the form

$$|\Psi\rangle\rangle = \frac{1}{\sqrt{T+1}}\sum_{t=0}^{T} |t\rangle_C \otimes U_t \ldots U_1|\phi\rangle_S = \sum_{t=0}^{T} |t\rangle_C \otimes |\psi(t)\rangle_S, \tag{7}$$

with $|\phi\rangle \in \mathcal{H}_S$ being the circuit's input, see Fig. 2. When projecting the clock onto the final time the system is in the state $|\psi\rangle = U_T \ldots U_1|\phi\rangle$, which corresponds to the output of the circuit under consideration. While it is not straightforward to write a physical projector analogous to the continuous case in Eq. (4), we can define a projection operator onto the space of solutions to the constraint equation by

$$\hat{P} := \sum_i |\Psi_i\rangle\rangle\langle\langle\Psi_i|, \tag{8}$$

where the $|\Psi_i\,\rangle\rangle$ are given according to Eq. (7) with initial states $|\phi_i\rangle$ taken from an orthonormal basis for $\mathcal{H}_S$. $\hat{P}$ is the projector onto the space of physical states, which contrarily to the continuous case is now a proper subspace of

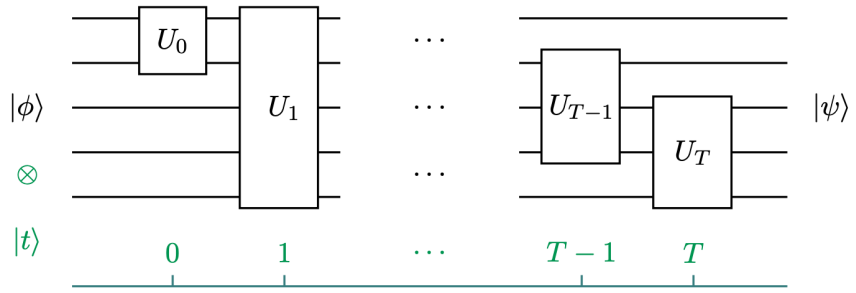


FIG. 2: The main idea of a Page-Wootters formulation of a quantum circuit. One considers a quantum clock that keeps track of the number of computational steps that have happened so far. At computational step t, the circuit applies the gate U_t. The input to the quantum circuit is $|\phi\rangle$ and the output of the circuit is $|\psi\rangle = U_T \cdots U_0 |\phi\rangle$.

$\mathcal{H}_c \otimes \mathcal{H}_S$. Note that similarly to the continuous case we can relate the physical projector to the unitary evolution of the circuit between the respective times by $(t_2 > t_1)$

$$\langle t_2 | \hat{P} | t_1 \rangle = \frac{1}{T+1} U_{t_2} \cdots U_{t_1+1}. \tag{9}$$

In what follows we will associate a discrete clock c_X with each agent $X \in \{A_1 \dots A_N\}$ which gives rise to history states of the form

$$|\Psi\rangle\rangle = \sum_{t_{A_1}=0,\dots t_{A_N}=0}^{T_{A_1}\dots T_{A_N}} |t_{A_1}, \dots t_{A_N}\rangle \otimes |\psi(t_{A_1} \dots t_{A_N})\rangle_S = \sum_{t_{A_1}=0,\dots t_{A_N}=0}^{T_{A_1}\dots T_{A_N}} |t_{A_1}\rangle \dots |t_{A_N}\rangle \otimes M_{t_{A_1}\dots t_{A_N}} |\phi\rangle, \tag{10}$$

where $|\phi\rangle$ is the initial state of the system. Intuitively, the matrices $M_{t_{A_1},\dots t_{A_N}}$ encode what happens to the system between the initial time and the time when the collection of clocks shows the respective values. By projecting onto a certain clock state $\langle t_X | \Psi \rangle\rangle$ we will obtain conditional or perspectival states that correspond to the state agent X assigns to everything other than their own clock at time t_X. In the next section we present what we consider reasonable physical assumptions the conditional states and hence the history state have to fulfill. We will almost exclusively consider the history states $|\Psi\rangle\rangle$ as they explicitly represent the perspectives of the agents and the systems and we can directly impose physical requirements on them. The constraint operator $\hat{C}$ can then be implicitly defined afterwards as an operator that annihilates this family of history states. Whether this constraint operator has a simple form, or has desirable properties such as locality, is an interesting question that is nevertheless not pursued in this work.

IV. PROCESS MATRICES WITHIN A TIMELESS FORMALISM

A. The operational setting and postulates

In this section, we develop our framework that allows to model experiments described by pure process matrices within a generalized Page-Wootters approach.

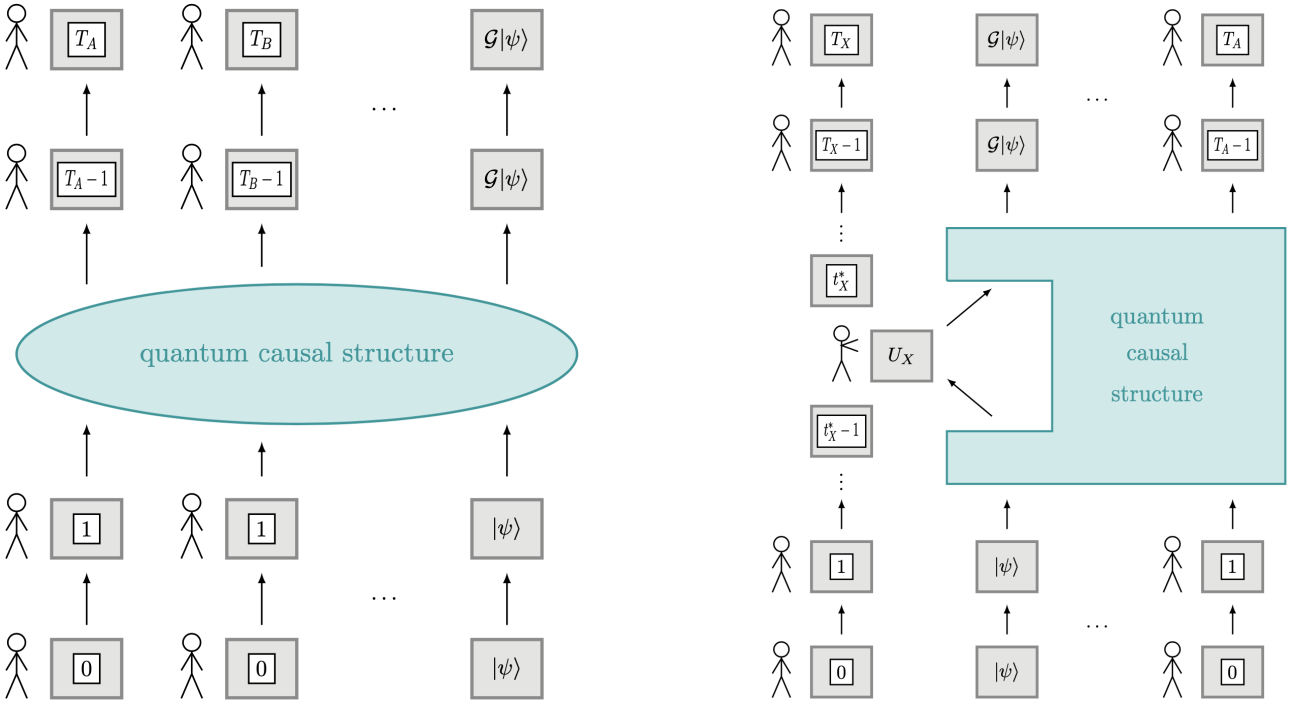
As in the pure process matrix formalism we will describe scenarios with multiple agents being parts of a well-defined standard global causal past and future. In addition to the usual process matrix approach, the progress of time within the agents' labs is described by quantum clocks with Hilbert spaces $\mathcal{H}_{c_{A_1}} \dots \mathcal{H}_{c_{A_N}}$. We refer to the set of all clock variables collectively as $\mathcal{H}_c$. The idea of a well defined global causal past and future common to all agents is formalized by the assumption that at the beginning as well as at the end of the protocol all the clocks experience at least one well-synchronized time step, see Figure 3a. During the protocol, each observer or agent applies his or her quantum instrument on a part of a system which is common to all agents. As done in [10] we will assume that each agent has access to an ancillary degree of freedom, denoted by Hilbert spaces $\mathcal{H}_{A'_1} \dots \mathcal{H}_{A'_N}$ of unspecified dimension, to implement their quantum instrument. This ancillary system allows to represent the quantum instrument as a unitary within our pure history state approach. The ancilla acts as the environment for a dilation and as memory recording measurement outcomes. We assume that the ancilla systems are initialized to $|0\rangle$, and that they have trivial time

evolution, except at the moment when the corresponding quantum instrument is applied. We collectively label the ancillas as $\mathcal{H}_{S'} := \mathcal{H}_{A'_1} \otimes \cdots \otimes \mathcal{H}_{A'_N}$.

In addition to the agents, their ancillas and quantum clocks, we also consider another quantum system that participates in the protocol, described by a Hilbert space $\mathcal{H}_S$. This quantum system represents the degrees of freedom that play an active role in the protocol, but are not directly associated with the agents or their labs. As in the formalism for pure processes, we assume that this quantum system is an input to the causal structure from the global past. We will often call it the *main system* and we denote its initial state by $|\psi\rangle_S$.

Hence our history states live on $\mathcal{H}_c \otimes \mathcal{H}_S \otimes \mathcal{H}_{S'}$. We assume that the clocks are initialized to time 0 at the start of the protocol and show times $T_{A_1}, T_{A_2}, \ldots T_{A_N}$ at the end. Then our history states can be expanded in the form:

$$|\Psi\rangle\rangle = \sum_{t_{A_1}=0,\ldots t_{A_N}=0}^{T_{A_1}\ldots T_{A_N}} |t_{A_1}, \ldots t_{A_N}\rangle_c \otimes |\psi(t_{A_1} \ldots t_{A_N})\rangle_{SS'}. \tag{11}$$



(a) Global perspective

(b) X's perspective

FIG. 3: The protocol of an experiment involving quantum causal structure from a global (3a) and a local (3b) point of view. At the beginning and end of the experiment, the agents are assumed to be in a standard definite causal structure. This is expressed by having their clocks tick in synchronization. However, in between the agents and the main system enter a possibly indefinite causal structure in which the clocks, the main system and the labs might get entangled with each other. The ancillary systems for the laboratories are not shown. Inside the labs standard quantum theory is valid and therefore each agent X only sees the other agents and the main system as part of a quantum causal structure. At some time $t_X^* - 1$, X receives the part of the main system described by $\mathcal{H}_X$ from the environment. X applies unitary operation U_X to this part of the main system and potential ancillas. Afterwards, X sends that part of the main system back into the environment at time t_X^*. The actions of all agents together lead to the process $\mathcal{G}$ being applied to the main system at the end of the protocol.

Now we can formalize our requirements for the timeless state describing the protocol depicted in Figure 3. As mentioned before all clocks and ancillas are initialized to the states $|0\rangle$ and, therefore, we can write:

S.1 $|\psi(0,0,\ldots)\rangle = |\psi\rangle_S|0\rangle_{S'}$, where $|0\rangle_{S'} = |0\rangle_{A'_1} \otimes |0\rangle_{A'_2} \otimes \cdots \otimes |0\rangle_{A'_N}$ is a fixed ancillary state and $|\psi\rangle_S$ is an arbitrary state of the system.

At the beginning and end of the experiment, physics should be given by a standard space-time causal structure. Hence, at the beginning and in the end, we assume the clocks of the agents are well-synchronized. In particular the clocks perform at least one synchronized step before and after they are part of any exotic causal structure. We further assume that during these initial and final well-synchronized time-steps nothing happens to the main system and formulate this in terms of agent A for the sake of readability. Note that this does not conceptually single out agent A but can equally be written analogously for any of the agents.

S.2 $|\psi(0,\dots,t_X,\dots)\rangle \neq 0$ only for $t_X = 0\ \forall X \neq A_1$ and $|\psi(T_{A_1},\dots,t_X,\dots)\rangle \neq 0$ only for $t_X = T_X\ \forall X \neq A_1$ and furthermore $|\psi(1,1\dots1)\rangle = |\psi(0,0,\dots,0)\rangle$ and $|\psi(T_{A_1}-1,T_{A_2}-1,\dots,T_{A_N}-1)\rangle = |\psi(T_{A_1},T_{A_2},\dots,T_{A_N})\rangle$.

Analogous to the pure process matrices formalism described in Section II, we model the input from the environment as parts of the main system, i.e. we assume that the input to agent X lives on a subspace $\mathcal{H}_X \subseteq \mathcal{H}_S$, and X's quantum instrument is described by a unitary U_X which acts on the received part of the main system and X's ancilla, i.e. U_X acts on $\mathcal{H}_X \otimes \mathcal{H}_{X'}$. Note that different $\mathcal{H}_X$ do not need to be different or orthogonal, in fact all of them might even be the full main system Hilbert space $\mathcal{H}_S$.

Next, we consider the perspective of the agents. As in Ref. [22] we condition the history state $|\Psi\rangle\rangle$ on X's clock showing time t_X , i.e. ${}_{c_X}\langle t_X|\Psi\rangle\rangle$, to describe what agent X sees at time t_X. In principle, the inner product in the kinematical Hilbert space is not necessarily the same as the inner product for the Hilbert space associated to the perspective of agent X [20]. Indeed in the usual Page-Wooters formalism with infinite dimensional systems, the physical Hilbert space is not a proper subspace of the kinematical Hilbert space; this necessitates to define a new inner product for the perspectival states. Moreover, even in the finite-dimensional setting that we study here, in scenarios involving clocks with varying relative ticking speeds, one runs into normalization issues if one simply uses the kinematical inner product for the perspectival states. To see this, consider the example of the history state $|\Psi\rangle\rangle = \int \mathrm{d}t_A |t_A\rangle_{c_A} \otimes |2t_A\rangle_{c_B}$ in which one clock runs twice as fast as the other [53]. For A's perspective we find ${}_{c_A}\langle t_A|\Psi\rangle\rangle = |2t_A\rangle$. However, for B's perspective we find

$$ {}_{c_B}\langle t_B|\Psi\rangle\rangle = \int \mathrm{d}t_A |t_A\rangle\langle t_B|2t_A\rangle = \frac{1}{2}\int \mathrm{d}t'_B |\frac{1}{2}\ t'_B\rangle\langle t_B|t'_B\rangle = \frac{1}{2}|1/2\ t_B\rangle, $$

where the prefactor $\frac{1}{2}$ comes from the measure via the change of the integration variable. Here, the clock ticking rates are constant. However, in general the rates might change dynamically and the corresponding prefactor will depend on time. In this example, if one demands that the states of the agents are normalized such that the agents each see a state $|\tau\rangle$ for some time τ, the different agents will need different renormalizations, which can be accounted for in the definition of the inner products for the perspectival states. For the finite dimensional case with discrete clocks this motivates the introduction of normalization operators $N_{t_X}^{(X)}$ in order to relate the normalization of the multipartite history state with the normalization of the time-dependent perspectival states.

Another motivation for the introduction of normalization operators for discrete clocks is given by the process of discretization itself. A typical discretization procedure ("time-binning") would map similar continuous times t, t', i.e. $t \approx t'$ but $t \neq t'$, to the same discrete time t_{discrete}. However, while $|t\rangle$ and $|t'\rangle$ are orthogonal states, they get approximated by the same discrete state $|t_{\text{discrete}}\rangle$. This procedure may lead to discrete-time states that are not properly normalized, in particular for the perspectives of the agents. But this can be fixed via the introduction of normalization operators. Details of the issues related to the process of discretization are discussed in Appendix A.

We assume that the state X sees at time t_X is

$$ |\psi_X(t_X)\rangle = N_{t_X}^{(X)} \langle t_X|\Psi\rangle\rangle = \langle t_X|_{c_X} \otimes N_{t_X}^{(X)} |\Psi\,\rangle\rangle, \tag{12} $$

where $N_{t_X}^{(X)} \in \mathcal{L}(\mathcal{H}_{c_{\setminus X}} \otimes \mathcal{H}_S \otimes \mathcal{H}_{S'})$ is the normalization operator that relates the perspective-neutral description to the perspective of agent X at time t_X. Here, $\mathcal{H}_{c_{\setminus X}}$ is the Hilbert space formed by all clocks except the clock of agent X.

A priori, the normalization operators make this approach extremely general. In principle, they could give us any state $|\psi_X(t_X)\rangle$ that we want. Therefore it is important that we impose some extra conditions. First of all, as the normalization operators generalize normalization constants, they should be linear, positive and invertible. Moreover, we wish that all the relevant physics concerning the initial system state $|\psi\rangle_S$ and the agents' operations is encoded in the history state, not the normalization operators. The normalization operator should just correct the normalization depending on the clocks. Therefore, we demand that the operators $N_{t_X}^{(X)}$ are independent of the initial system state $|\psi\rangle$ and the choice of quantum instruments by the agents.

N.1 $N_{t_X}^{(X)}$ is an invertible, linear, positive operator. It is independent of the input state $|\psi\rangle_S$ and the local operations $U_{A_1}\dots U_{A_N}$.

Without the latter restriction, one could use $N_{t_X}^{(X)}$ to introduce copies of the initial state $|\psi\rangle_S$ or apply copies of the agents' instruments to violate the no-cloning principle. We further assume that the normalization operator does not perturb how one agent sees the clocks of the other agents:

N.2 The normalization operator has the form

$$N_{t_X}^{(X)} = \sum_{t_{A_1},\dots,\widehat{t_X},\dots t_{A_N}} |t_{A_1},\dots \widehat{t}_X,\dots t_{A_N}\rangle\langle t_{A_1},\dots \widehat{t}_X,\dots t_{A_N}| \otimes n^{(X)}_{t_{A_1},\dots \widehat{t}_X,\dots t_{A_N}} \otimes \mathbb{1}_{S'} \tag{13}$$

where the sum is taken over all clocks except the clock of agent X, which is omitted, as indicated by $\widehat{t_X}$. The operator $n^{(X)}_{t_{A_1},\dots \widehat{t}_X,\dots t_{A_N}}$ is a linear, invertible and positive operator acting on $\mathcal{H}_S$ (but not on the ancillas $\mathcal{H}_{S'}$).

This assumption is motivated by the requirement that the history state should represent the relevant physics and relations of the clocks. As their name implies, the normalization operators should adjust the normalization, but not introduce new clock physics.

Our previous requirement of well-synchronized clocks at the beginning and end of the experiment additionally implies that the respective normalization operators should just be identity operators.

N.3 $N_1^{(X)} = N_0^{(X)} = \mathbb{1}$ as well as $N_{T_X-1}^{(X)} = N_{T_X}^{(X)} = \mathbb{1}\ \forall X$.

Finally, we have to explain how the perspectival states $N_{t_X}^{(X)}\langle t_X|\Psi\rangle\rangle$ are related to each other. We will assume that each agent X sees a unitary time evolution as dictated by quantum theory in a pure state framework. This means we assume that for all t_X, t'_X there exists a unitary operator $\mathcal{U}_X(t_X, t'_X)$ such that

$$|\psi_X(t_X)\rangle\rangle = \mathcal{U}_X(t_X, t'_X)|\psi_X(t'_X)\rangle\rangle. \tag{14}$$

Furthermore, just like in usual quantum theory, $\mathcal{U}_X(t,t')$ should not depend on the initial system state $|\psi\rangle_S$.

U.1 $\mathcal{U}_X(t,t')$ is a unitary operator, independent of the initial state $|\psi\rangle_S$.

Moreover, time-evolution from t'' to t' to t is the same as time-evolution from t'' to t.

U.2 $\mathcal{U}_X(t,t')\mathcal{U}_X(t',t'') = \mathcal{U}_X(t,t'')$, $\forall t,t',t''$.

Next we discuss the crucial assumption that connects our framework to the formalism of pure processes. In the process matrix framework, one assumes that during the protocol each agent eventually receives a quantum system from the environment. We will assume that each agent is promised that they will receive this quantum system at a specific time $t_X^* - 1$. As explained before, we model that quantum system to be a part of the main system, described by subspaces $\mathcal{H}_{A_1}\dots\mathcal{H}_{A_N}$ of the main system space $\mathcal{H}_S$. Each agent X acts with their local operation U_X on that system and their ancilla (i.e. $U_X \in \mathcal{L}(\mathcal{H}_X \otimes \mathcal{H}_{X'})$) and then sends out the system at t_X^*. In particular, this is the only time the agents use their quantum instrument. While the agents enforce evolution via their instrument, the remaining degrees of freedom should evolve in an uncorrelated way.

U.3 X's quantum instrument is used at the so called *time of action* t_X^*, i.e.

$$\mathcal{U}_X(t_X^*, t_X^* - 1) = U_X \otimes \mathrm{Rest}^{(X)}. \tag{15}$$

Furthermore at other times $t \neq t_X^*$ the evolution operator $\mathcal{U}_X(t, t-1)$ is independent of U_X and only acts as the identity on the ancilla of X, i.e. on $\mathcal{H}_{X'}$.

Our assumptions introduce a transformation that maps the initial state $|\psi(0,\dots,0)\rangle$ to the final state $|\psi(T_{A_1},\dots,T_{A_N})\rangle$. This transformation depends on the agents' actions U_X and is visualized in Fig. 3. Our next assumption is that this transformation can be extended to a full process [10]/quantum superchannel [44]. This means that it must be possible to interpret the quantum causal structure as a process, even if we describe the agents' operations as channels instead of (purified) unitaries.

We make the implicit assumption that our Postulates **S.2**, **N.1**, **N.2**, **N.3**, **U.1**, **U.2** and **U.3** continue to be satisfied if the ancillary systems are initialized to states other than $|0\rangle_{A'_j}$, and that we can continue to use the same

normalization operators $N_{t_X}^{(X)}$ and perspectival time evolutions $\mathcal{U}_X(t_X+1,t_X)$ as for the initialization $|0,\dots,0\rangle_{S'}$. This is no substantial conceptual restriction, because none of these postulates explicitly refers to any particular initial ancillary system state. Postulate **S.1** just defines the particular choice of initialization for the protocol.

This concludes the description of the operational setting and of our assumptions. We will subsequently investigate the mathematical and physical implications of our setting and postulates.

B. History states lead to pure processes

First, we show that the evolution of the main system and the ancillas must be given by a pure process. For that purpose we have to analyze the relation between the initial and the final state, in particular with respect to the operations of the agents. This can be done by taking the perspective of an agent, for example A_1. We apply our unitary time evolution postulates **U.1**, **U.2** and **U.3** to decompose:

$$\mathcal{U}_{A_1}(T_{A_1},0) = \mathcal{U}_{A_1}(T_{A_1},t^*_{A_1})(U_{A_1}\otimes \mathrm{Rest}^{(A_1)})\mathcal{U}_{A_1}(t^*_{A_1}-1,0) \tag{16}$$

Postulate **N.3** tells us that at times 0 and T_{A_1} the normalization operators are just identity operators. Furthermore, our Postulate **S.2** assumes that at the beginning all clocks show time 0, at the end all clocks show time T_{A_j}, and that in the beginning and at the end there is at least one synchronized time step. Furthermore, we assume that all clock states $|t\rangle_{cX}$ are normalized, i.e. ${}_{cX}\langle t|t\rangle_{cX}=1$. This shows us that:

$$|\psi_{A_1}(0)\rangle = {}_{c_{A_1}}\langle 0| \sum_{t_{A_1}=0,\dots t_{A_N}=0}^{T_{A_1}\dots T_{A_N}} |t_{A_1},\dots t_{A_N}\rangle_c \otimes |\psi(t_{A_1}\dots t_{A_N})\rangle_{SS'} = |0,\dots 0\rangle_{c_{\setminus A_1}}\otimes|\psi(0,0,\dots 0)\rangle \tag{17}$$

$$|\psi_{A_1}(T_{A_1})\rangle = {}_{c_{A_1}}\langle T_{A_1}| \sum_{t_{A_1}=0,\dots t_{A_N}=0}^{T_{A_1}\dots T_{A_N}} |t_{A_1},\dots t_{A_N}\rangle_c \otimes |\psi(t_{A_1}\dots t_{A_N})\rangle_{SS'} = |T_{A_2},\dots,T_{A_N}\rangle_{c_{\setminus A_1}}\otimes|\psi(T_{A_1},T_{A_2},\dots,T_{A_N})\rangle$$

Together with the information that $\mathcal{U}_{A_1}(T_{A_1},0)$ maps $|\psi_{A_1}(0)\rangle$ to $|\psi_{A_1}(T_{A_1})\rangle$, we find

$$|T_{A_2}\dots T_{A_N}\rangle_{c_{\setminus A_1}}\otimes|\psi(T_{A_1},T_{A_2},\dots T_{A_N})\rangle = \mathcal{U}_{A_1}(T_{A_1},0)\big(|0,\dots 0\rangle_{c_{\setminus A_1}}\otimes|\psi(0,0,\dots 0)\rangle\big) \tag{18}$$

$$=: |T_{A_2},\dots T_{A_N}\rangle\otimes\mathcal{G}(U_{A_1},U_{A_2}\dots)|\psi(0,0,\dots 0)\rangle. \tag{19}$$

Here, we defined a map $\mathcal{G}$ that describes how the final main system and ancilla state is related to the initial state:

$$\begin{aligned}|\psi(T_{A_1}\dots T_{A_N})\rangle &= \Big(\langle T_{A_2}\dots T_{A_N}|_{c_{\setminus A_1}}\,\mathcal{U}_{A_1}(T_{A_1},0)|0,\dots,0\rangle_{c_{A_1}}\Big)\,|\psi(0,0,\dots,0)\rangle\\ &=: \mathcal{G}(U_{A_1}\dots U_{A_N})|\psi(0,0,\dots 0)\rangle.\end{aligned} \tag{20}$$

At the end of Section IV A, we assumed that all our assumptions continue to be satisfied if the ancillary systems are initialized to another state than $|0,\dots,0\rangle_{S'}$. In particular, we assumed that we can continue to use the same perspectival unitaries $\mathcal{U}_X(t_X,t_X-1)$ and normalization operators $N_{t_X}^{(X)}$, even if the ancillary systems are initialized in a state different from $|0,\dots,0\rangle_{S'}$. Therefore, the states $|\psi(0,0,\dots 0)\rangle$ span the entire input space of $\mathcal{G}(U_{A_1}\dots U_{A_N})$. In particular, while we took the perspective of agent A_1 to derive $\mathcal{G}(U_{A_1}\dots U_{A_N})$, Equation (20) also holds for all other agents, because $|\psi(0,\dots,0)\rangle$ and $|\psi(T_{A_1},\dots,T_{A_N})\rangle$ are pieces of the perspective-independent history state.

Equation (18) shows that $\mathcal{G}(U_{A_1}\dots U_{A_N})$ is a unitary that maps the initial main system and ancilla state to the final state and that it is multilinear in the local operations (at least, as long as the linear combinations result in another unitary). To see linearity, one can infer from Eqs. (16) and (18) that:

$$\begin{aligned}&|\psi(T_{A_1},T_{A_2},\dots T_{A_N})\rangle =\\ &\Big(\langle T_{A_2}\dots T_{A_N}|_{c_{\setminus A_1}}\,\mathcal{U}_{A_1}(T_{A_1},t^*_{A_1})(U_{A_1}\otimes\mathrm{Rest}^{(A_1)})\mathcal{U}_{A_1}(t^*_{A_1}-1,0)|0,\dots 0\rangle_{c_{\setminus A_1}}\Big)\,|\psi(0,0,\dots 0)\rangle\end{aligned} \tag{21}$$

and use that the agent's unitary U_{A_1} is only applied at their time of action. Unitarity can be seen, for example, in the following way:

$$\langle\psi_j(T_{A_1},T_{A_2},\dots T_{A_N})|\psi_k(T_{A_1},T_{A_2},\dots T_{A_N})\rangle$$

$$
\begin{aligned}
&=\langle T_{A_2}\dots T_{A_N}|T_{A_2}\dots T_{A_N}\rangle\langle\psi_j(T_{A_1},T_{A_2},\dots T_{A_N})|\psi_k(T_{A_1},T_{A_2},\dots T_{A_N})\rangle\\
&=\Big(\langle 0,\dots,0|_{c_{\setminus A_1}}\otimes\langle\psi_j(0,0,\dots,0)|\Big)\,\mathcal{U}_{A_1}(T_{A_1},0)^\dagger\,\mathcal{U}_{A_1}(T_{A_1},0)\,\Big(|0,\dots,0\rangle_{c_{\setminus A_1}}\otimes|\psi_k(0,0,\dots,0)\rangle\Big)\\
&=\langle\psi_j(0,0,\dots,0)|\psi_k(0,0,\dots,0)\rangle
\end{aligned}
$$

Therefore, inner products (and in particular orthonormality) are preserved. Alternatively, one can use that the unitaries in Eq. (18) imply that $|\psi(T_{A_1},T_{A_2},\dots T_{A_N})\rangle$ and $|\psi(0,0,\dots,0)\rangle$ always have the same norm.

Furthermore, at the end of Postulate **U.3** we assumed that $\mathcal{U}_X(t_X,t_X-1)$ acts as the identity on the ancillary system of agent X, except for the time of action of that agent, where that agent's operation U_X is applied. Now, Equation (21) shows that the only change in the state of the ancilla of A_j is caused by A_j's local operation. We assumed that the map $\mathcal{G}(U_{A_1},\dots,U_{A_N})$ from $|\psi(0,\dots,0)\rangle$ to $|\psi(T_{A_1},\dots,T_{A_N})\rangle$ can be extended to a full process. We conclude that $\mathcal{G}$ is a pure process as defined in [10, 44] with the difference that Equation (21) represents a refined causal reference frame decomposition that explicitly includes the quantum clocks, compare to Equation (1).

The fact that we obtain pure processes has important consequences: According to [44, 54], in the bipartite case our setting implies that no violation of device-independent causal inequalities can occur: The bipartite pure process can only be causally ordered or quantum-controlled causal order.

C. Additional restrictions for the local perspectives of the agents

Let us further investigate the relation between our framework and the original causal reference frame framework of [11], in particular the relation between Equations (1) and (21), in further detail. Both frameworks work with purifications, in particular the actions of the agents are described by purified unitaries $U_{A_1}\dots U_{A_N}$ and the relevant process matrices turn out to be the pure processes. The crucial objects of the causal reference frame framework are the unitaries that describe the past before and the future after an agent's action. More specifically, from the point of view of agent X, the evolution from the beginning of the protocol up to the time of X's action is described by the unitary Π_X. In our framework, this unitary corresponds to $\mathcal{U}_X(t_X^*-1,0)$. The evolution directly after X's action up to the end of the protocol is described by Φ_X, which in our framework corresponds to $\mathcal{U}_X(T_X,t_X^*)$.

The crucial difference between our framework and that of causal reference frames is that we explicitly model the quantum clocks and explain how the agents' perspectives arise from a perspective neutral history state. This gives us a refined description of the agents' perspectives because we explicitly model individual time steps $t_X\to t_X+1$ in between the beginning of the experiment, the time of action and the end of the protocol. In Ref. [11] the causal future and past unitaries Φ_X and Π_X are allowed to be arbitrary as long as they combine to the pure process $\mathcal{G}$ via Equation (1). However, in our setting the history state induces further compatibility constraints on the perspectives of the agents.

We will now present one such constraint that is particularly restrictive: Affine-linearity in the operations of the other agents. Consider a history state as in Eq. (11). We can write

$$|\psi(t_{A_1}\dots t_{A_N})\rangle=M_{t_{A_1}\dots t_{A_N}}|\psi(0,0,\dots 0)\rangle, \tag{22}$$

with

$$M_{t_{A_1}\dots t_{A_N}}={}_{c_{\setminus A_1}}\langle t_{A_2},\dots t_{A_N}|(N^{(A_1)}_{t_{A_1}})^{-1}\,\mathcal{U}_{A_1}(t_{A_1},0)|0,\dots 0\rangle_{c_{\setminus A_1}}. \tag{23}$$

To see this, use Eq. (17), i.e. $|\psi_{A_1}(0)\rangle=|0,\dots,0\rangle_{c_{\setminus A_1}}\otimes|\psi(0,\dots,0)\rangle$. Remember, we got this equation from assuming that in the beginning, the normalization operator is the identity, i.e. Postulate **N.3**, and that in the beginning all the clocks are initialized to 0 and tick synchronized, i.e. Postulate **S.2**. Then we find:

$$
\begin{aligned}
M_{t_{A_1},\dots,t_{A_N}}|\psi(0,0,\dots,0)\rangle=&{}_{c_{\setminus A_1}}\langle t_{A_2},\dots t_{A_N}|(N^{(A_1)}_{t_{A_1}})^{-1}\,\mathcal{U}_{A_1}(t_{A_1},0)|0,\dots 0\rangle_{c_{\setminus A_1}}\otimes|\psi(0,0,\dots,0)\rangle_{SS'}\\
=&{}_{c_{\setminus A_1}}\langle t_{A_2},\dots t_{A_N}|(N^{(A_1)}_{t_{A_1}})^{-1}\,\mathcal{U}_{A_1}(t_{A_1},0)|\psi_{A_1}(0)\rangle\\
=&{}_{c_{\setminus A_1}}\langle t_{A_2},\dots t_{A_N}|(N^{(A_1)}_{t_{A_1}})^{-1}|\psi_{A_1}(t_{A_1})\rangle={}_{c_{\setminus A_1}}\langle t_{A_2},\dots t_{A_N}|\otimes\langle t_{A_1}|_{c_{A_1}}|\Psi\rangle\rangle\\
=&|\psi(t_{A_1},\dots t_{A_N})\rangle
\end{aligned} \tag{24}
$$

We already mentioned that all assumptions, perspectival unitaries $\mathcal{U}_X(t_X, t_X - 1)$, and normalization operators $N_{t_X}^{(X)}$ continue to work if the ancillary systems are initialized to states other than $|0, \ldots, 0\rangle_{S'}$. Therefore, the states $|\psi(0, \ldots, 0)\rangle\rangle$ span the entire input space of $M_{t_{A_1}, \ldots, t_{A_N}}$. Furthermore, the states $|\psi(t_{A_1}, \ldots, t_{A_N})\rangle\rangle$ are pieces of the perspective-neutral history state. Therefore Eq. (23) is also true for the agents exchanged.

Consider Equation (23). Postulate **U.2** allows us to decompose $\mathcal{U}_{A_1}(t_{A_1}, 0)$ into individual time steps. By Postulate **U.3**, the unitary action U_{A_1} of the agent A_1 only appears at the time of action, and there it appears linearly. By Postulate **N.1**, the normalization operators are independent of the agents' actions. Therefore, we can see that $M_{t_{A_1} \ldots t_{A_N}}$ is constant in U_{A_1} for $t_{A_1} < t_{A_1}^*$ and linear in U_{A_1} for $t_{A_1} \geq t_{A_1}^*$ (if the linear combination results in another unitary), because the same is true for $\mathcal{U}_{A_1}(t_{A_1}, 0)$.

We pointed out that Eq. (23) is also true with exchanged agents. We can relate the time evolutions of two different agents (here A_1 and A_2) via

$$\mathcal{U}_{A_2}(t_{A_2}, 0)|0, 0, \ldots 0\rangle_{c_{\setminus A_2}} = \sum_{t_{A_1}, t_{A_3}, \ldots t_{A_N}} N_{t_{A_2}}^{(A_2)} |t_{A_1}, t_{A_3}, \ldots t_{A_N}\rangle_{c_{\setminus A_2}} M_{t_{A_1}, \ldots t_{A_N}} \tag{25}$$

To see this, exchange the agents in Eq. (23) to find

$$\sum_{t_{A_1}, t_{A_3}, \ldots t_{A_N}} N_{t_{A_2}}^{(A_2)} |t_{A_1}, t_{A_3}, \ldots t_{A_N}\rangle_{c_{\setminus A_2}} M_{t_{A_1}, \ldots t_{A_N}} \tag{26}$$

$$= \sum_{t_{A_1}, t_{A_3}, \ldots t_{A_N}} N_{t_{A_2}}^{(A_2)} |t_{A_1}, t_{A_3}, \ldots t_{A_N}\rangle_{c_{\setminus A_2}} \langle t_{A_1}, t_{A_3}, \ldots, t_{A_N}| (N_{t_{A_2}}^{(A_2)})^{-1} \mathcal{U}_{A_2}(t_{A_2}, 0)|0, \ldots, 0\rangle_{c_{\setminus A_2}} \tag{27}$$

and use that the time states $|t\rangle_{c_{A_j}}$ form an orthonormal basis of the clock spaces.

As the normalization operators are independent of the agents' actions, the dependence of $M_{t_{A_1}, \ldots t_{A_N}}$ on U_{A_1} shows that $\mathcal{U}_{A_2}(t_{A_2}, 0)|0, 0, \ldots\rangle$ is a sum of functions linear in U_{A_1} or constant in U_{A_1}, i.e. $\mathcal{U}_{A_2}(t_{A_2}, 0)|0, 0, \ldots\rangle$ is affine-linear in U_{A_1} (if the linear combination gives another unitary). The same argument can be made for all other agents. Hence we get that any time evolution $\mathcal{U}_X(t_X, 0)|0, \ldots 0\rangle_{c_{\setminus A_X}}$ as seen by agent X with all clocks initialized to time 0 has to be an affine-linear function of the operations of all other agents (at least, if the linear combinations give unitaries). This affine-linearity is a severe restriction and a potential obstacle for implementing some non-causal processes in this framework. In Section V D we will apply this insight to an example involving an exotic tripartite process [10, 55] to see that a causal reference frame decomposition for this process in Ref. [11] is incompatible with our framework.

D. Discrete constraint operators and physical projectors

Finally, we will briefly discuss constraint operators and physical projectors in our framework since they are among the main objects of interest in the Page-Wootters formalism presented in Section III. By construction, our history states form a subspace $\mathcal{H}_H \subset \mathcal{H}_c \otimes \mathcal{H}_S \otimes \mathcal{H}_{S'}$ and by linearity, $\alpha|\Psi\rangle\rangle + \beta|\Psi'\rangle\rangle$ is the history state associated with the input state $\alpha|\psi\rangle_S + \beta|\psi'\rangle_S$, as one can see e.g. from Equation (22). Therefore, we can define a constraint operator $\hat{C}$ as $\hat{C} = \mathbb{1} - \hat{P}_H$ where $\hat{P}_H$ is the orthogonal projector onto $\mathcal{H}_H$. Then the kernel of $\hat{C}$ is given by $\mathcal{H}_H$.

We note that in general, it is unclear whether $\hat{P}_H$ in our framework can be written analogous to the case of a standard quantum circuit with one clock, see Equation (8). More specifically, for an orthonormal basis $|\psi_j\rangle_S$, the corresponding history states

$$|\Psi_j\rangle\rangle = \sum_{t_{A_1} \ldots t_{A_N}} |t_{A_1} \ldots t_{A_N}\rangle \otimes |\psi_j(t_{A_1} \ldots t_{A_N})\rangle_S = \sum_{t_{A_1}=0}^{T_{A_1}} |t_{A_1}\rangle (N_{t_{A_1}}^{(A_1)})^{-1} |\psi_{A_1,j}(t_{A_1})\rangle\rangle$$

$$= \sum_{t_{A_1}=0}^{T_{A_1}} |t_{A_1}\rangle (N_{t_{A_1}}^{(A_1)})^{-1} \mathcal{U}_{A_1}(t_{A_1}, 0)|0, 0, \ldots\rangle \otimes |\psi_j\rangle_S$$

may fail to be orthogonal due to the normalization operators:

$$\langle\langle\Psi_k|\Psi_j\rangle\rangle = \sum_{t_{A_1}=0}^{T_{A_1}} \langle\psi_k|_S \otimes \langle 0,0,\dots 0|\,\mathcal{U}_{A_1}(t_{A_1},0)^\dagger [(N_{t_{A_1}}^{(A_1)})^{-1}]^\dagger (N_{t_{A_1}}^{(A_1)})^{-1}\,\mathcal{U}_{A_1}(t_{A_1},0)|0,0,\dots 0\rangle \otimes |\psi_j\rangle_S. \tag{28}$$

In that sense, the map from initial states to history states is not necessarily unitary, in contrast to the unitary evolution of the main system and ancilla state, see Eq. (20).

We can, however, write $\hat{P}_H$ in a form more reminiscent of the original Page-Wootters framework, compare Eq. (4), as

$$\hat{P}_H = \frac{1}{T}\sum_{k=0}^{T-1} \exp\left(-2\pi i \hat{C}\frac{k}{T}\right), \tag{29}$$

where T is an integer (we could take $T = T_{A_1}$). This can be seen by noting that $\hat{C}$ is a hermitian matrix with only eigenvalues 0 or 1. If $|\phi_0\rangle$ is an eigenvector of $\hat{C}$ with $\hat{C}|\phi_0\rangle = 0$, we have $\hat{P}_H|\phi_0\rangle = |\phi_0\rangle$, while if $\hat{C}|\phi_1\rangle = |\phi_1\rangle$ we have $\hat{P}_H|\phi_1\rangle = \frac{1}{T}\sum_{k=0}^{T-1} e^{-2\pi i \frac{k}{T}}|\phi_1\rangle = 0$, showing $\hat{P}_H = \mathbb{1} - \hat{C}$. As discussed in Appendix B it is not clear whether $\hat{P}_H$ in Equation (29) can be linked to the perspectival unitaries $\mathcal{U}_X(t'_X, t_X)$ similar to Equation (9).

V. CAUSAL AND NON-CAUSAL PAGE-WOOTTERS CIRCUITS

In this section we now apply our framework to give several examples of physical scenarios that go beyond the standard setting of circuits with well-synchronized clocks. First, in Sec. V A, we consider a setup inspired by the famous twin paradox in which the clocks of two agents are still in a well-defined relation to each other, but tick at different rates. Afterwards, in Section V B, we consider a scenario for the bipartite quantum switch as a prototypical example of a known class of non-causal processes. There, a control quantum system determines the tick rates of the agents' clocks and more importantly the order of the agents' operations. In Section V C, we go beyond the example of the bipartite switch and show that arbitrary coherently controlled causal orders can be realized in our framework. Finally, in Section V D, we consider an interesting pure, non-causal process that is further known to violate causal inequalities [10, 24, 55]. We will argue that this process cannot be implemented as a superposition of classical histories and that the causal reference frame decomposition from Ref. [11] cannot be adapted to our setting.

A. A history state for a scenario with varying clock ticking rates

Our first example of an interesting process that fits into our setting but is not a standard circuit is inspired by the famous twin paradox. Specifically we consider a scenario that features varying clock speeds of two agents A and B where during the protocol the clock of one ticks slower than the clock of the other, reminiscent of the one twin that leaves earth traveling at relativistic speed and returns to find his or her sibling older than they are themselves.

Here the two agents act on subsystems S_A and S_B of the input quantum state $|\phi\rangle \in \mathcal{H}_S$ with unitary operations U_A and U_B respectively. The casual order in this example is well defined and we consider the case where A acts before B. Moreover, between A's and B's time of action some global evolution V of the system happens, which is independent of the two agents. In the beginning and at the end the agents' clocks tick at the same speed, but in between the clock of A ticks more slowly than that of B. The scenario is depicted in Figure 4.

It is captured by the following history state $|\Psi\rangle\rangle \in \mathcal{H}_{c_A} \otimes \mathcal{H}_{c_B} \otimes \mathcal{H}_{S_A} \otimes \mathcal{H}_{S_B}$:

$$\begin{aligned}|\Psi\rangle\rangle =& |0_A,0_B\rangle_c \otimes |\phi\rangle + |1_A,1_B\rangle_c \otimes |\phi\rangle + |2_A,2_B\rangle_c \otimes (U_A\otimes\mathbb{1})|\phi\rangle + |2_A,3_B\rangle_c \otimes (U_A\otimes\mathbb{1})|\phi\rangle + |3_A,4_B\rangle_c \otimes V(U_A\otimes\mathbb{1})|\phi\rangle \\ &+ |3_A,5_B\rangle_c \otimes V(U_A\otimes\mathbb{1})|\phi\rangle + |4_A,6_B\rangle_c \otimes (\mathbb{1}\otimes U_B)V(U_A\otimes\mathbb{1})|\phi\rangle + |4_A,7_B\rangle_c \otimes \mathcal{G}(U_A,U_B)|\phi\rangle \\ &+ |5_A,8_B\rangle_c \otimes \mathcal{G}(U_A,U_B)|\phi\rangle + |6_A,9_B\rangle_c \otimes \mathcal{G}(U_A,U_B)|\phi\rangle\end{aligned} \tag{30}$$

where $\mathcal{G}(U_A,U_B) = (\mathbb{1}\otimes U_B)V(U_A\otimes\mathbb{1})$. The perspectival states for the two agents including the non-trivial normalization operators are

$$|\psi_A(0)\rangle = |0_B\rangle_{c_B}\otimes|\phi\rangle, \qquad |\psi_A(1)\rangle = |1_B\rangle_{c_B}\otimes|\phi\rangle, \qquad |\psi_B(0)\rangle = |0_A\rangle_{c_A}\otimes|\phi\rangle, \qquad |\psi_B(1)\rangle = |1_A\rangle_{c_A}\otimes|\phi\rangle,$$

$$
\begin{aligned}
|\psi_A(2)\rangle &= \frac{1}{\sqrt{2}}(|2_B\rangle + |3_B\rangle)_{c_B} \otimes (U_A \otimes \mathbb{1})|\phi\rangle, & |\psi_B(2)\rangle &= |2_A\rangle_{c_A} \otimes (U_A \otimes \mathbb{1})|\phi\rangle \\
&\text{with } N_2^{(A)} = \frac{1}{\sqrt{2}}\mathbb{1}_S, & &= |\psi_B(3)\rangle, \\
|\psi_A(3)\rangle &= \frac{1}{\sqrt{2}}(|4_B\rangle + |5_B\rangle)_{c_B} \otimes V(U_A \otimes \mathbb{1})|\phi\rangle & |\psi_B(4)\rangle &= |3_A\rangle_{c_A} \otimes V(U_A \otimes \mathbb{1})|\phi\rangle \\
&\text{with } N_3^{(A)} = \frac{1}{\sqrt{2}}\mathbb{1}_S, & &= |\psi_B(5)\rangle, \\
|\psi_A(4)\rangle &= \frac{1}{\sqrt{2}}(|6_B\rangle + |7_B\rangle)_{c_B} \otimes (\mathbb{1} \otimes U_B)V(U_A \otimes \mathbb{1})|\phi\rangle & |\psi_B(6)\rangle &= |4_A\rangle_{c_A} \otimes (\mathbb{1} \otimes U_B)V(U_A \otimes \mathbb{1})|\phi\rangle \\
&\text{with } N_4^{(A)} = \frac{1}{\sqrt{2}}\mathbb{1}_S, & &= |\psi_B(7)\rangle, \\
|\psi_A(5)\rangle &= |8_B\rangle_{c_B} \otimes \mathcal{G}(U_A, U_B)|\phi\rangle, & |\psi_B(8)\rangle &= |5_A\rangle_{c_A} \otimes \mathcal{G}(U_A, U_B)|\phi\rangle, \\
|\psi_A(6)\rangle &= |9_B\rangle_{c_B} \otimes \mathcal{G}(U_A, U_B)|\phi\rangle, & |\psi_B(9)\rangle &= |6_A\rangle_{c_A} \otimes \mathcal{G}(U_A, U_B)|\phi\rangle.
\end{aligned}
\tag{31}
$$

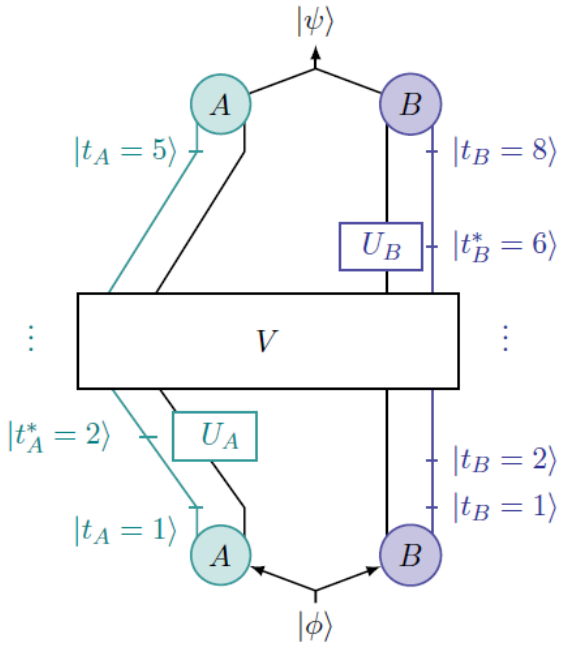


FIG. 4: Example of a setting involving a clock with changing ticking rate. The two agents A and B each receive a part of the input system and experience one synchronized time step. After that the clock of A starts ticking slower and A applies their unitary U_A to their subsystem of state $|\phi\rangle$. This is followed by some unitary evolution V of the full system, which is independent of the two agents. Then B applies their unitary U_B to his or her subsystem and finally, at the end of the protocol, the two clocks tick in synchronization once more.

Note that the normalization operators are non trivial for precisely those times where the clock of A ticks slower than the clock of B. The states in Equations (31) can be reproduced by the following unitary evolutions with respect to the two agents

$$
\begin{aligned}
\mathcal{U}_A(1,0) &= T_{c_B} \otimes \mathbb{1}_\$, & \mathcal{U}_B(1,0) &= T_{c_A} \otimes \mathbb{1}_S, \\
\mathcal{U}_A(2,1) &= (T_2')_{c_B} \otimes (U_A \otimes \mathbb{1})_S, & \mathcal{U}_B(2,1) &= T_{c_A} \otimes (U_A \otimes \mathbb{1})_S, \\
\mathcal{U}_A(3,2) &= (T^2)_{c_B} \otimes V_S, & \mathcal{U}_B(3,2) &= \mathbb{1}, \\
\mathcal{U}_A(4,3) &= (T^2)_{c_B} \otimes (\mathbb{1} \otimes U_B)_S, & \mathcal{U}_B(4,3) &= T_{c_A} \otimes V_S, \\
\mathcal{U}_A(5,4) &= (T_6')_{c_B} \otimes \mathbb{1}_S, & \mathcal{U}_B(5,4) &= \mathbb{1}, \\
& & \mathcal{U}_B(6,5) &= T_{c_A} \otimes (\mathbb{1} \otimes U_B)_S,
\end{aligned}
\tag{32}
$$

$$\mathcal{U}_A(6,5) = T_{c_B} \otimes \mathbb{1}_S, \qquad \mathcal{U}_B(7,6) = \mathbb{1}$$
$$\mathcal{U}_B(8,7) = T_{c_A} \otimes \mathbb{1}_S = \mathcal{U}_B(9,8),$$

where T is the unitary that makes the clock of the other agent tick, i.e. $T : |t\rangle \mapsto |t+1\rangle$. Similarly, T_i' is any unitary that acts as $|i-1\rangle \mapsto 1/\sqrt{2}(|i\rangle + |i+1\rangle)$, $1/\sqrt{2}(|i\rangle + |i+1\rangle) \mapsto |i+2\rangle$. We see that our axioms are fulfilled and the times of action are $t_A^\star = 2$ and $t_B^\star = 6$ respectively. As one can see in Equations (32), from A's perspective B's clock seems to tick at double the rate in the middle of the process, while from the point of view of B, A's clock seems partially frozen in time.

B. The quantum switch

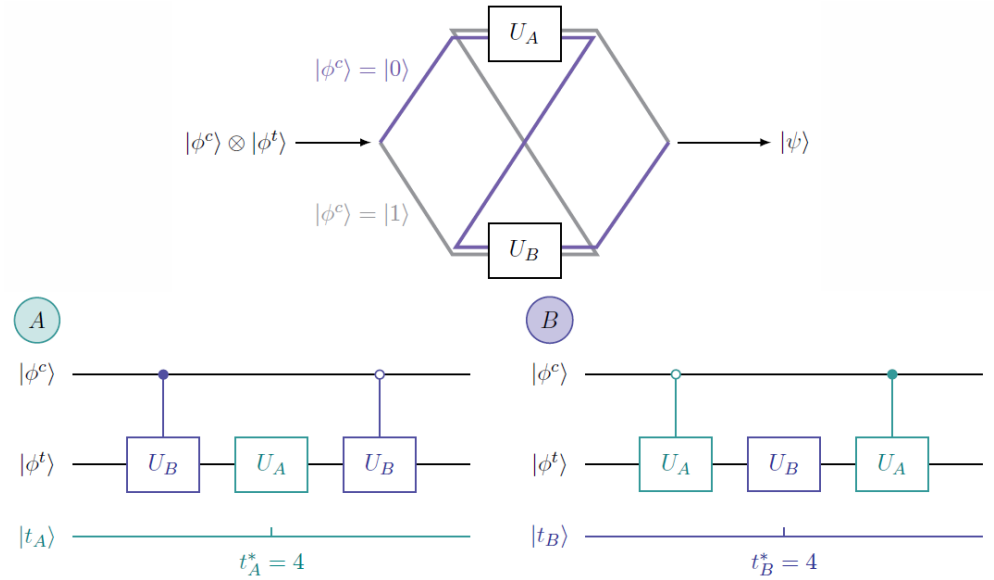


FIG. 5: The bipartite quantum switch: Depending on the value of a control qubit the two unitaries U_A, U_B are applied to the target system in different order (top). According to the perspectives of the two agents, A or B apply their own unitary to the target system at time t_A^* or t_B^* respectively, while the other agent's unitary is applied either before or after that depending on the value of the control system (bottom). The perspectival circuits equal the causal reference frames for the quantum switch given in Ref. [11].

Our second example describes the probably best known non-causal process, namely the bipartite quantum switch [26]. A schematic picture as well as the two perspectival circuits analogous to the causal reference frame decomposition given in Ref. [11] are shown in Figure 5. The bipartite quantum switch can be modeled by a history state complying with our axioms starting with an initial state $|\phi\rangle_S \in \mathcal{H}_{S_c} \otimes \mathcal{H}_{S_t}$ consisting of a control system and a target system. Both agents are acting on the target system Hilbert space; $\mathcal{H}_A = \mathcal{H}_B = \mathcal{H}_{S_t}$.

A history state of the quantum switch is given by

$$\begin{aligned}|\Psi\rangle\rangle = {} & |0_A, 0_B\rangle_c \otimes |\phi\rangle + |1_A, 1_B\rangle_c \otimes |\phi\rangle + |2_A, 2_B\rangle_c \otimes |\phi\rangle + |3_A, 2_B\rangle_c \otimes (|0\rangle\langle 0| \otimes \mathbb{1})|\phi\rangle + |2_A, 3_B\rangle_c \otimes (|1\rangle\langle 1| \otimes \mathbb{1})|\phi\rangle \\ & + |4_A, 3_B\rangle_c \otimes (|0\rangle\langle 0| \otimes U_A)|\phi\rangle + |3_A, 4_B\rangle_c (|1\rangle\langle 1| \otimes U_B)|\phi\rangle + |5_A, 4_B\rangle_c \otimes (|0\rangle\langle 0| \otimes U_B U_A)|\phi\rangle \\ & + |4_A, 5_B\rangle_c \otimes (|1\rangle\langle 1| \otimes U_A U_B)|\phi\rangle + |5_A, 5_B\rangle_c \otimes (|0\rangle\langle 0| \otimes U_B U_A + |1\rangle\langle 1| \otimes U_A U_B)|\phi\rangle \\ & + |6_A, 6_B\rangle_c \otimes \mathcal{G}(U_A, U_B)|\phi\rangle + |7_A, 7_B\rangle_c \otimes \mathcal{G}(U_A, U_B)|\phi\rangle,\end{aligned} \tag{33}$$

where $\mathcal{G}(U_A, U_B) = |0\rangle\langle 0| \otimes U_B U_A + |1\rangle\langle 1| \otimes U_A U_B$ is the (pure) process matrix. Intuitively the history state in Equation (33) describes the scenario where, depending on the value of the control, different time orderings (A's clock ticks at a faster rate than B's or vice versa) are initiated by de-synchronizing initially synchronized clocks. For the two time orderings different orders of the agents' operations (either U_A or U_B first) are applied. Finally the clocks are re-synchronized, again making use of the control degree of freedom, such that they can tick together at the end of the protocol. We obtain the following perspectival states

$$
\begin{aligned}
|\psi_A(0)\rangle &= |0_B\rangle_{c_B} \otimes |\phi\rangle \\
|\psi_A(1)\rangle &= |1_B\rangle_{c_B} \otimes |\phi\rangle \\
|\psi_A(2)\rangle &= |2_B\rangle_{c_B} \otimes (|0\rangle\langle 0| \otimes \mathbb{1})|\phi\rangle \\
&\quad + \frac{1}{\sqrt{2}}(|2_B\rangle + |3_B\rangle)_{c_B} \otimes (|1\rangle\langle 1| \otimes \mathbb{1})|\phi\rangle \\
\text{with } N_2^{(A)} &= |0\rangle\langle 0|_{Sc} + \frac{1}{\sqrt{2}}|1\rangle\langle 1|_{Sc} \\
|\psi_A(3)\rangle &= |2_B\rangle_{c_B} \otimes (|0\rangle\langle 0| \otimes \mathbb{1})|\phi\rangle \\
&\quad + |4_B\rangle_{c_B} \otimes (|1\rangle\langle 1| \otimes U_B)|\phi\rangle
\end{aligned}
\qquad
\begin{aligned}
|\psi_B(0)\rangle &= |0_A\rangle_{c_A} \otimes |\phi\rangle \\
|\psi_B(1)\rangle &= |1_A\rangle_{c_A} \otimes |\phi\rangle \\
|\psi_B(2)\rangle &= |2_A\rangle_{c_A} \otimes (|1\rangle\langle 1| \otimes \mathbb{1})|\phi\rangle \\
&\quad + \frac{1}{\sqrt{2}}(|2_A\rangle + |3_A\rangle)_{c_A} \otimes (|0\rangle\langle 0| \otimes \mathbb{1})|\phi\rangle \\
\text{with } N_2^{(B)} &= \frac{1}{\sqrt{2}}|0\rangle\langle 0|_{Sc} + |1\rangle\langle 1|_{Sc} \\
|\psi_B(3)\rangle &= |2_A\rangle_{c_A} \otimes (|1\rangle\langle 1| \otimes \mathbb{1})|\phi\rangle \\
&\quad + |4_A\rangle_{c_A} \otimes (|0\rangle\langle 0| \otimes U_A)|\phi\rangle
\end{aligned}
\tag{34}
$$

$$
\begin{aligned}
|\psi_A(4)\rangle &= |3_B\rangle_{c_B} \otimes (|0\rangle\langle 0| \otimes U_A)|\phi\rangle \\
&\quad + |5_B\rangle_{c_B} \otimes (|1\rangle\langle 1| \otimes U_A U_B)|\phi\rangle \\
|\psi_A(5)\rangle &= \frac{1}{\sqrt{2}}(|4_B\rangle + |5_B\rangle)_{c_B} \otimes (|0\rangle\langle 0| \otimes U_B U_A)|\phi\rangle \\
&\quad + |5_B\rangle_{c_B} \otimes (|1\rangle\langle 1| \otimes U_A U_B)|\phi\rangle \\
\text{with } N_5^{(A)} &= \frac{1}{\sqrt{2}}|0\rangle\langle 0|_{Sc} + |1\rangle\langle 1|_{Sc} \\
|\psi_A(6)\rangle &= |6_B\rangle_{c_B} \otimes \mathcal{G}(U_A, U_B)|\phi\rangle \\
|\psi_A(7)\rangle &= |7_B\rangle_{c_B} \otimes \mathcal{G}(U_A, U_B)|\phi\rangle
\end{aligned}
\qquad
\begin{aligned}
|\psi_B(4)\rangle &= |3_A\rangle_{c_A} \otimes (|1\rangle\langle 1| \otimes U_B)|\phi\rangle \\
&\quad + |5_A\rangle_{c_A} \otimes (|0\rangle\langle 0| \otimes U_B U_A)|\phi\rangle \\
|\psi_B(5)\rangle &= \frac{1}{\sqrt{2}}(|4_A\rangle + |5_A\rangle)_{c_A} \otimes (|1\rangle\langle 1| \otimes U_A U_B)|\phi\rangle \\
&\quad + |5_A\rangle_{c_A} \otimes (|0\rangle\langle 0| \otimes U_B U_A)|\phi\rangle \\
\text{with } N_5^{(B)} &= |0\rangle\langle 0|_{Sc} + \frac{1}{\sqrt{2}}|1\rangle\langle 1|_{Sc} \\
|\psi_B(6)\rangle &= |6_A\rangle_{c_A} \otimes \mathcal{G}(U_A, U_B)|\phi\rangle \\
|\psi_B(7)\rangle &= |7_A\rangle_{c_A} \otimes \mathcal{G}(U_A, U_B)|\phi\rangle
\end{aligned}
$$

which can be related to each other by unitaries

$$
\begin{aligned}
\mathcal{U}_A(1,0) &= T_{c_B} \otimes \mathbb{1}_S \\
\mathcal{U}_A(2,1) &= T_{c_B} \otimes (|0\rangle\langle 0| \otimes \mathbb{1})_S + (T_2')_{c_B} \otimes (|1\rangle\langle 1| \otimes \mathbb{1})_S \\
\mathcal{U}_A(3,2) &= \mathbb{1}_{c_B} \otimes (|0\rangle\langle 0| \otimes \mathbb{1})_S + (T_2')_{c_B} \otimes (|1\rangle\langle 1| \otimes U_B)_S \\
\mathcal{U}_A(4,3) &= T_{c_B} \otimes (\mathbb{1} \otimes U_A)_S \\
\mathcal{U}_A(5,4) &= (T_4')_{c_B} \otimes (|0\rangle\langle 0| \otimes U_B)_S + \mathbb{1}_{c_B} \otimes (|1\rangle\langle 1| \otimes \mathbb{1})_S \\
\mathcal{U}_A(6,5) &= (T_4')_{c_B} \otimes (|0\rangle\langle 0| \otimes \mathbb{1})_S + T_{c_B} \otimes (|1\rangle\langle 1| \otimes \mathbb{1})_S \\
\mathcal{U}_A(7,6) &= T_{c_B} \otimes \mathbb{1}_S
\end{aligned}
\qquad
\begin{aligned}
\mathcal{U}_B(1,0) &= T_{c_A} \otimes \mathbb{1}_S \\
\mathcal{U}_B(2,1) &= T_{c_A} \otimes (|1\rangle\langle 1| \otimes \mathbb{1})_S + (T_2')_{c_A} \otimes (|0\rangle\langle 0| \otimes \mathbb{1})_S \\
\mathcal{U}_B(3,2) &= \mathbb{1}_{c_A} \otimes (|1\rangle\langle 1| \otimes \mathbb{1}) + (T_2')_{c_A} \otimes (|0\rangle\langle 0| \otimes U_A) \\
\mathcal{U}_B(4,3) &= T_{c_A} \otimes (\mathbb{1} \otimes U_B)_S \\
\mathcal{U}_B(5,4) &= (T_4')_{c_A} \otimes (|1\rangle\langle 1| \otimes U_A)_S + \mathbb{1}_{c_A} \otimes (|0\rangle\langle 0| \otimes \mathbb{1})_S \\
\mathcal{U}_B(6,5) &= (T_4')_{c_A} \otimes (|1\rangle\langle 1| \otimes \mathbb{1})_S + T_{c_A} \otimes (|0\rangle\langle 0| \otimes \mathbb{1})_S \\
\mathcal{U}_B(7,6) &= T_{c_A} \otimes \mathbb{1}_S.
\end{aligned}
\tag{35}
$$

where T and T_i' are the same as in the previous example. It is straightforward to see that all our axioms are fulfilled. Note that the unitaries in Equations (35) are not unique but were chosen such that the perspectives of the agents resemble the causal reference frames of the two agents presented in Ref. [11]. For both A and B the time of action is $t_A^\star = 4 = t_B^\star$ and depending on the value of the control the other agent applies their unitary either before or after $t_A^\star$ or $t_B^\star$ respectively, compare to Figure 5.

C. General coherent control of causal order

The quantum switch from the previous section is a famous example for an important class of processes with indefinite causal structure, namely processes with coherently controlled causal order [56–58]. There, for each value of the control system $|k\rangle \in \mathcal{H}_{S_C}$ one associates a process with definite causal order or quantum comb $\widetilde{\mathcal{G}}_k$ [42, 43] and the definite causal order is different for at least two different k [59]. We will now present the general idea for implementing such processes in our framework, while details are given in Appendix C.

Quantum combs describe processes in which agents $A_1, \ldots, A_N$ have a definite causal order. They get their name from the figures that are used to visualize them, see Figure 6. Combs have a global past to receive an input state, and a global future to which they output a state. In between, there are N open slots into which agents $A_1, \ldots, A_N$ can insert quantum operations. Each slot belongs to one agent, fixing a definite causal order.

For our purpose, quantum combs are sequences of channels with memory [42–44], see Figure 6 for a tripartite example. Each of these channels comes with open ends such that the agents can insert their operations. More

precisely, in between two consecutive comb channels one agent operation can be inserted. Meanwhile, the memory directly connects the channels. The ancillary systems of the agents run parallel to the combs, i.e. the comb channels do not act on the ancillas the agents use. Similarly, the agents do not act on the memory of the comb.

The channels of the combs can be purified to unitaries [45] V_j by extending the memories, having an extra pure fixed input to the comb, and discarding part of the output of the comb. As we work in a purified formulation of quantum theory and quantum processes in particular, and as our goal is to describe a coherently controlled process, we will only consider pure combs that are given as a sequence of unitaries with memory.

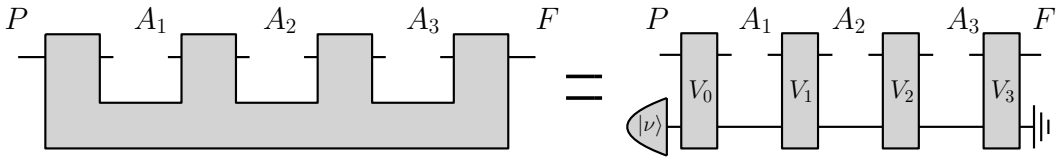


FIG. 6: A tripartite quantum comb: A general process with fixed causal order is a map on the actions of three agents A_1, A_2 and A_3 (left). Time passes from left to right, where P stands for past and F for future, and, hence, A_1 acts before A_2 and agent A_3 acts last. The agents' actions are quantum instruments which must be inserted into the slots of the comb. Combs can be implemented as sequence of unitary channels with memory when adding an additional environment system with input $|\nu\rangle$ and discarding part of the output (right) [42–45].

Let us formalize our considerations from above. We consider N agents $A_1, \ldots, A_N$ and consider M pure combs $\widetilde{\mathcal{G}}_k$, $1 \leq k \leq M$, given by

$$\widetilde{\mathcal{G}}_k(U_1, \ldots, U_N) = V_N^{(k)} U_{\pi_k(N)} V_{N-1}^{(k)} U_{\pi_k(N-1)} \ldots V_1^{(k)} U_{\pi_k(1)} V_0^{(k)}. \tag{36}$$

Here, $V_j^{(k)}$ are the unitary comb channels with memory. Identity operators on the comb memories and ancillary systems are left implicit. $U_1, \ldots, U_N$ are the unitary operations of the agents $A_1, \ldots A_N$. The order of the agents in comb $\widetilde{\mathcal{G}}_k$ is determined by a permutation π_k, with the meaning that agent $A_{\pi_k(j)}$ will be the j-th agent to act in comb $\widetilde{\mathcal{G}}_k$.

The combs $\widetilde{\mathcal{G}}_k$ have to satisfy some compatibility conditions such that we can combine them into a coherently controlled superposition. To demonstrate our protocol, it will be enough to assume that all the open ends of all the slots of all the combs have the same dimension, and that all the memories of all the combs have the same dimension. The first assumption means that there is a dimension d_{agents} that determines the dimension of the input that the agents receive from the comb, and that this dimension is independent of agent A_j and the comb index k. It is also the dimension of the output that the agents provide to the combs. This assumption implies that the dimensions of the ancillary systems do not change after the agents have applied their operations. The second assumption means that there is a memory dimension d_{memory} that is independent of the comb index k and the index of the slot. This assumption will allow us to let all combs use the same memory system. As we will see, our protocol does not crucially depend on these assumptions. The only purpose of these assumptions is to make the natural definition of coherently controlled causal order in Equation (38) well-defined. However, the protocol will not care about the precise meaning of Eq. (38). In Appendix C we will discuss relaxations, generalizations and alternatives to these assumptions.

Now, we introduce a M-dimensional control system $\mathcal{H}_{S_C}$. Control value $|k\rangle_{S_C}$ means that comb $\widetilde{\mathcal{G}}_k$ is implemented. Thus, we can define processes with coherently controlled causal order as

$$\mathcal{G}(U_1, \ldots, U_N) = \sum_{k=1}^{M} |k\rangle\langle k|_{S_C} \otimes \widetilde{\mathcal{G}}_k(U_1, \ldots, U_N) \tag{37}$$

$$= \sum_{k=1}^{M} |k\rangle\langle k|_{S_C} \otimes V_N^{(k)} U_{\pi_k(N)} V_{N-1}^{(k)} U_{\pi_k(N-1)} \ldots V_1^{(k)} U_{\pi_k(1)} V_0^{(k)}, \tag{38}$$

Again, identity operators on the agents' ancillary systems and comb memories are left implicit. As one can see, $\mathcal{G}$ is unitary and multilinear in the operations of the agents. Such processes were also considered in [56–58].

Now we describe the history state implementing $\mathcal{G}(U_1, \ldots, U_N)$ given by Equation (38). The input state $|\psi\rangle_S \in \mathcal{H}_{S_C} \otimes \mathcal{H}_{S_P}$ comprises a control system ($\in \mathcal{H}_{S_C}$) and another system ($\in \mathcal{H}_{S_P}$) which represent the input to the combs from the global past. We decompose the protocol and, hence, the history state into three parts as

$$|\Psi\rangle\rangle = |\Psi_{\text{desync}}\rangle\rangle + |\Psi_{\text{combs}}\rangle\rangle + |\Psi_{\text{resync}}\rangle\rangle, \tag{39}$$

where $|\Psi_{\text{desync}}\rangle\rangle$ describes the beginning of the protocol, where we use the control degree of freedom to desynchronize the clocks such that the agents are put into the right order. Afterwards the different combs are applied depending on the value of the control in $|\Psi_{\text{combs}}\rangle\rangle$. At last, the resynchronization of the agents' clocks is described by $|\Psi_{\text{resync}}\rangle\rangle$. The strategy is depicted in Figure 7.

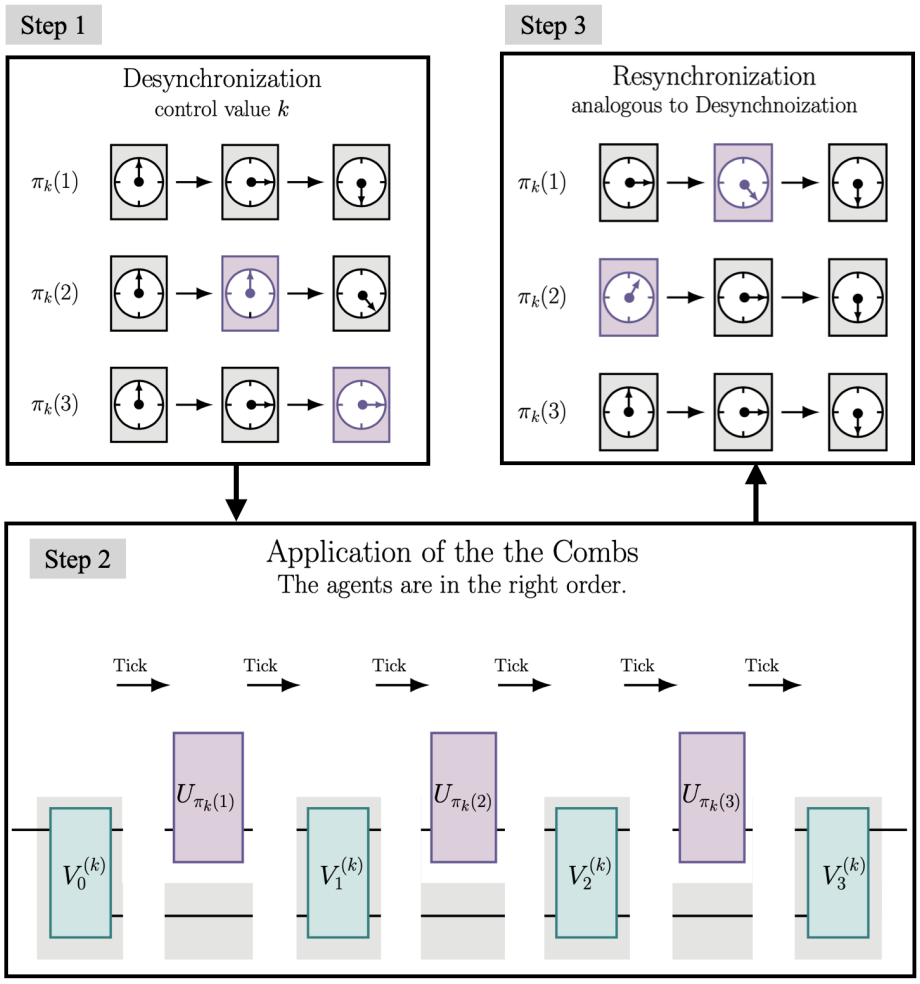


FIG. 7: The strategy for implementing coherently controlled causal order: The protocol consists of three steps which are all conditioned on the control value, namely desynchronization, application of the combs and resynchronization. The time of action t^* is chosen the same for all agents. To be able to implement the comb of a given k, the clocks of the agents get desynchronized such that the agents will act in the right order. This is achieved by first freezing the clocks of all but the agent $A_{\pi_k(1)}$ one after the other. First the clock of the agent acting second gets frozen followed by that of the agent acting third etc. The clock freezes are indicated in purple. The duration of the freeze depends on when the agent will act. After the desired orderings have been implemented, the combs get applied while all clocks tick in synchronization. First one applies the first unitary with memory of the comb, then the first agent acts (their clock shows t^*). Then the second comb unitary is applied followed by the second agent's action etc. At last the clocks get resynchronized again by using the desynchronization protocol, but with the role of the agents reversed.

During the desynchronization of the clocks nothing happens to the input to the combs and we can write

$$\begin{aligned}|\Psi_{\text{desync}}\rangle\rangle =&|0,\dots,0\rangle_c\otimes|\psi\rangle_S+|1,\dots,1\rangle_c\otimes|\psi\rangle_S+|2,\dots,2\rangle_c\otimes|\psi\rangle_S\\&+\sum_{k=1}^{M}\sum_{j=3}^{T_0}|t_1^{(k)}(j),\dots,t_N^{(k)}(j)\rangle_c\otimes(|k\rangle\langle k|_{S_C}\otimes\mathbb{1}_{S_P})|\psi\rangle_S,\end{aligned}\tag{40}$$

where, as we will see, $T_0 := t^* - 2$, with t^* being the time of action for all agents. The $t_i^{(k)}(j)$ give different time orderings by freezing different clocks for different amounts of time steps during which the other clocks keep ticking. More precisely, if an agent will act as the m-th agent in the comb with the control value $|k\rangle_{S_C}$, then the clock of that agent gets frozen for $2(m-1)$ time steps. This ensures that two consecutive agents are two time steps apart from each other when they enter $|\Psi_{\text{combs}}\rangle\rangle$. While one agent's clock is frozen, the clocks of the other agents march on. See Appendix C for the detailed clock freezing and desynchronization protocol. Afterwards, we include additional synchronized ticks at the end of the desynchronization step to ensure that for all k the clock freezes are far away from the application of the combs. This together with the fact that the agents' operations U_j have not been used, yet, gives perspectival controlled unitaries that act non-trivially only on the clocks of the other agents, i.e. $\mathcal{U}_{A_j}(t,t-1)=\sum_{k=1}^{M}u_{c,k}^{A_j}(t,t-1)\otimes|k\rangle\langle k|_{Sc}\otimes\mathbb{1}_{Sp}$. Here, $u_{c,k}^{A_j}(t,t-1)$ are the unitaries that only act on the clocks of the other agents.

In $|\Psi_{\text{combs}}\rangle\rangle$ the clocks will continue to tick in synchronization by means of the unitary T introduced in Section V A while, given a control value k, the unitaries of the comb $\widetilde{\mathcal{G}}_k$ are applied one after the other. All the agents, for a given k, see the following sequence of unitaries at the respective time steps:

$$V_0^{(k)}\otimes T^{\otimes(N-1)},\ U_{\pi_k(1)}\otimes T^{\otimes(N-1)},\ V_1^{(k)}\otimes T^{\otimes(N-1)},\ U_{\pi_k(2)}\otimes T^{\otimes(N-1)},\ \dots,\ U_{\pi_k(N)}\otimes T^{\otimes(N-1)},\ V_N^{(k)}\otimes T^{\otimes(N-1)}\tag{41}$$

Further details are again in Appendix C. The time differences caused by freezing the clocks ensure that the time of action t^* satisfies $t^* = T_0 + 2$ for all agents.

For the resynchronization in $|\Psi_{\text{resync}}\rangle\rangle$ we repeat the procedure from $|\Psi_{\text{desync}}\rangle\rangle$, but with the role of the agents inverted, i.e. $t_{\pi_k(m)}^{(k)}\mapsto t_{\pi_k(N+1-m)}^{(k)}$. In the end, all the clocks tick in synchronization and show the same time. Like the desynchronization this last part of the protocol is independent of the agents' operations U_j and our axioms are fulfilled. Hence, any coherent control of causal order as described by Equation (38) can be implemented in our framework.

In Appendix C we will discuss how to relax or modify the consistency conditions for the combs $\widetilde{\mathcal{G}}_k$ that allows them to appear in a controlled superposition. We will also provide details about the protocol itself. Furthermore, we will discuss how to take the purification process of the combs into account.

D. About an exotic process

A notorious example of a tripartite pure process with indefinite causal order from [10, 55] is known to violate causal inequalities. Said process is *not* an example of coherent control of causal order. It is often referred to as the *Lugano process*. The time reversed version of the Lugano process was discussed in Ref. [11] and can be written as

$$\mathcal{G}(U_A,U_B,U_C)|jjj\rangle = U_A\otimes U_B\otimes U_C|jjj\rangle\tag{42}$$
$$\mathcal{G}(U_A,U_B,U_C)|j01\rangle = XU_A\otimes U_B\otimes U_C|j01\rangle\tag{43}$$
$$\mathcal{G}(U_A,U_B,U_C)|1j0\rangle = U_A\otimes XU_B\otimes U_C|1j0\rangle\tag{44}$$
$$\mathcal{G}(U_A,U_B,U_C)|01j\rangle = U_A\otimes U_B\otimes XU_C|01j\rangle\tag{45}$$

where $j\in\{0,1\}$ and $X=\sigma_X$ is the Pauli-X matrix. Defining projectors $P_A=\sum_j|j01\rangle\langle j01|$, $P_B=\sum_j|1j0\rangle\langle 1j0|$, $P_C=\sum_j|01j\rangle\langle 01j|$ and $P_\perp=\sum_j|jjj\rangle\langle jjj|$ one gets

$$\mathcal{G}(U_A,U_B,U_C)|\phi\rangle=(U_A\otimes U_B\otimes U_CP_\perp+XU_A\otimes U_B\otimes U_CP_A+U_A\otimes XU_B\otimes U_CP_B+U_A\otimes U_B\otimes XU_CP_C)|\phi\rangle.\tag{46}$$

One crucial difference between the reversed Lugano process and the non-casual processes discussed in Section V C is the lack of a control degree of freedom. Therefore, it is not possible to directly adapt the history state procedure that we used for coherently controlled causal order to the reversed Lugano process. Instead, the main system itself has to control the desynchronization process. One can try to, similarly to the quantum switch, use the projectors P_A,P_B,P_C and $P_\perp$ to define a controlled operation that de-synchronizes the clocks. Afterwards, one can use the clocks as a control system to define another controlled operation that applies the unitary operations (42)- (45) for the different control values. However, the re-synchronization cannot be done independently of the unitaries U_A, U_B and U_C. More specifically, the described procedure will lead to a term in the history state of the form

$$\begin{aligned}|\Psi\rangle\rangle = \cdots + |\gamma_\perp\rangle_c \otimes (U_A \otimes U_B \otimes U_C P_\perp)|\phi\rangle_S + |\gamma_A\rangle_c \otimes (XU_A \otimes U_B \otimes U_C P_A)|\phi\rangle_S \\ + |\gamma_B\rangle_c \otimes (U_A \otimes XU_B \otimes U_C P_B)|\phi\rangle_S + |\gamma_C\rangle_c \otimes (U_A \otimes U_B \otimes XU_C P_C)|\phi\rangle_S + \dots \end{aligned} \tag{47}$$

with some clock states $|\gamma_\perp\rangle_c$, $|\gamma_A\rangle_c$, $|\gamma_B\rangle_c$ and $|\gamma_C\rangle_c$, which represent the different time orderings. The question is how to complete the history state, i.e. how to resynchronize the clocks. We are only allowed to use the agents' operations once and this has already happened. The states $U_A \otimes U_B \otimes U_C P_\perp|\phi\rangle_S$, $XU_A \otimes U_B \otimes U_C P_A|\phi\rangle_S$, $U_A \otimes XU_B \otimes U_C P_B|\phi\rangle_S$ and $U_A \otimes U_B \otimes XU_C P_C|\phi\rangle_S$ all depend on U_A, U_B, U_C in different, non-trivial ways. This means any overall map using them to "resynchronize" the clocks will non-trivially depend on U_A, U_B and U_C as well. This in turn leads to a non-trivial dependence of $\mathcal{U}_X(t_X, t_X - 1)$ on U_X for all $X \in \{A, B, C\}$ during the resynchronization part towards the end of the protocol, i.e. for $t_X > t_X^*$, which is a violation of Assumption **U.3**.

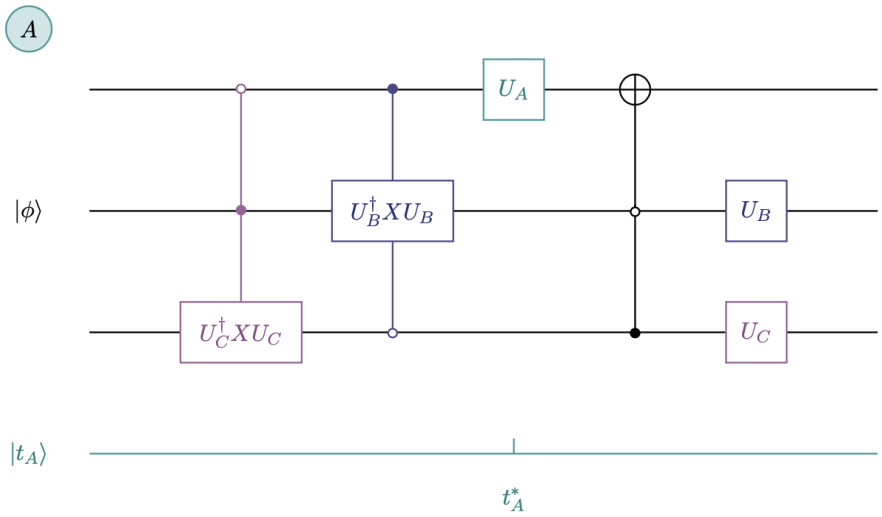


FIG. 8: The causal reference frame of agent A inside the time reversed Lugano process as given in [11] (B's and C's perspectival circuits look analogous). There is no control degree of freedom. All three agents act on different subsystems of the input system $|\phi\rangle$, but in a way that depends on the subsystems the other agents act on. Because of the gates that are not affine-linear in U_B and U_C this causal reference frame decomposition is incompatible with our setting.

Ref. [11] presented a causal reference frame decomposition of the reverse Lugano process, which for agent A is shown in Figure 8. It uses perspectival circuits with gates that are not affine-linear in the respective unitaries of the other agents, $U_B^\dagger X U_B$ and $U_C^\dagger X U_C$ for A's perspective. However, the corresponding perspectival states are forbidden in our framework due to the requirement of affine-linearity for $M_{t_A,t_B,t_C}(U_A, U_B, U_C)$ discussed in Section IV C.

Note, however, that the two impossible implementations of the process $\mathcal{G}(U_A, U_B, U_C)$ discussed above, namely the causal reference frame decomposition of [11] according to Figure 8 and the desynchronization-resynchronization-protocol, are not necessarily the only strategies for how to describe the reverse Lugano process within our non-causal Page-Wootters framework. Determining whether this process can be realized in this framework remains an open problem left for future work.

VI. CONCLUSION

In this work we showed how the Page-Wootters approach and the process matrix formalism may be combined to give a history state description of non-causal processes. We considered an operational setting that allows for probing indefinite causal structure. In this setting we explicitly modeled the passage of time as perceived by different agents using discrete quantum clocks. This allowed us to use a history state approach to which we added a set of well-motivated axioms about the protocol and the perspectives of the agents. As a consequence of these axioms, the causal structures arising in our setting are described by pure process matrices. A well-known result from previous literature about pure process matrices implies that in the bipartite case, no violation of device-independent causal inequalities can occur in our setting. Nonetheless, we could show that important physical scenarios beyond causal circuits and beyond non-relativistic clocks fit into our framework. More specifically, we showed how to describe a scenario inspired by the twin paradox involving varying clock ticking speeds with our approach. But most importantly, we proved that all processes representing coherent control of causal order (e.g. the quantum switch) can be implemented using our description.

We showed how to extract the time-evolution corresponding to the perspective of any given agent. This lead us to a refinement of the *causal reference frame* picture of Ref. [11] in which also the quantum clocks are explicitly modeled. The presence of these clocks and a perspective-neutral history state impose extra conditions on the causal reference frames. As an example, we showed that the evolution described by the causal past needs to be affine-linear in the operations of the other agents. We applied this extra condition to rule out a specific causal reference decomposition of the so-called time-reversed Lugano process provided in [11].

We conclude by pointing out a few directions for future research. While we focused on discrete clocks, this framework can be adapted to continuous clocks, extending the approach of Ref. [22] to a systematic operational protocol that allows for the extraction of process matrices.

In order to model the protocol for probing causal structure, we worked directly with history states instead of starting with a constraint operator or physical projector. As a consequence, the relation between the physical projector and the perspectival unitaries $\mathcal{U}_X(t', t)$ is an open question. Resolving this question might reveal further constraints on the history states, possibly restricting the set of process matrices that can be considered physical.

Also, one could consider more restrictions on the normalization operators. In our investigations, Postulate **N.2** did not play an important role. So it might be fruitful to investigate its consequences, or be even more restrictive about the specific form of the normalization operators.

An important class of non-causal processes that lack a physical interpretation are those that violate causal inequalities (e.g. the aforementioned Lugano process). If one could show that these processes do not fit into our setting, this would be important evidence that such processes should not be considered physical.

Furthermore, for notational convenience, we described our operational scenario only for the case that the unitaries of the agents do not change the dimensions of the ancillary systems and of the parts of the main system that the agents receive. Therefore, it is important to develop an extension of our formalism in which the dimensions are allowed to change.

At last, we pointed out that our assumption about the initialization of the clocks means that our protocol does not probe the full input space of the perspectival unitaries. Thus, it might be fruitful to investigate extensions and modifications of the operational scenario in which the agents enter the causal structure at other times than expected by the protocol. Doing this in a way that is mathematically and operationally self-consistent might reveal further mathematical restrictions. For example, if the initialization time of an agent is wrong, should the time of action be shifted too? If so, if one of the agents starts the protocol in a superposition of clock states, is the agent allowed to have a superposition of times of actions too?

VII. ACKNOWLEDGEMENTS

We thank Esteban Castro-Ruiz, Marco Túlio Quintino and David Trillo Fernandez for interesting discussions. We acknowledge the support of the Vienna Doctoral School in Physics (VDSP) and the support of the Austrian Science Fund (FWF) through the Doctoral Programme CoQuS. Č.B. acknowledges financial support from the Austrian Science Fund (FWF) through BeyondC (F7103-N48) and the project no. I-2906, from the European Commission via Testing the Large-Scale Limit of Quantum Mechanics (TEQ) (No. 766900) project, and from Foundational Questions Institute (FQXi). This research was supported by FQXi FFF Grant number FQXi-RFP-1815 from the Foundational Questions Institute and Fetzer Franklin Fund, a donor advised fund of Silicon Valley Community Foundation. We acknowledge a grant from the John Templeton Foundation (ID# 61466) as part of the The Quantum Information Structure of

Spacetime (QISS) Project (qiss.fr). Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Colleges and Universities.

[1] L. Hardy, Probability theories with dynamic causal structure: A new framework for quantum gravity (2005), arXiv:gr-qc/0509120 [gr-qc].
[2] L. Hardy, Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure, Journal of Physics A: Mathematical and Theoretical **40**, 3081 (2007).
[3] L. Hardy, Quantum gravity computers: On the theory of computation with indefinite causal structure, in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: The Western Ontario Series in Philosophy of Science* (Springer Netherlands, Dordrecht, 2009) pp. 379–401, arXiv:quant-ph/0701019 [quant-ph].
[4] R. P. Feynman, The necessity of gravitational quantization, in *The Role of Gravitation in Physics: Report from the 1957 Chapel Hill Conference*, edited by D. Rickles and C. M. DeWitt (Edition Open Sources, 2011).
[5] M. Zych, F. Costa, I. Pikovski, and Č. Brukner, Bell's theorem for temporal order, Nature Communications **10**, 3772 (2019).
[6] C. Rovelli, *Quantum Gravity*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, 2004).
[7] C. Kiefer, *Quantum Gravity*, 3rd ed., International Series of Monographs on Physics (Oxford University Press, 2012).
[8] J. Butterfield and C. Isham, Spacetime and the philosophical challenge of quantum gravity, in *Physics Meets Philosophy at the Planck Scale: Contemporary Theories in Quantum Gravity*, edited by C. Callender and N. Huggett (Cambridge University Press, 2001) 1st ed.
[9] O. Oreshkov, F. Costa, and C. Brukner, Quantum correlations with no causal order, Nature Commun. **3**, 1092 (2012), arXiv:1105.4464 [quant-ph].
[10] M. Araújo, A. Feix, M. Navascués, and Č. Brukner, A purification postulate for quantum mechanics with indefinite causal order, Quantum **1**, 10 (2017).
[11] P. A. Guérin and Č. Brukner, Observer-dependent locality of quantum events, New Journal of Physics **20**, 103031 (2018).
[12] D. N. Page and W. K. Wootters, Evolution without evolution: Dynamics described by stationary observables, Physical Review D **27**, 2885 (1983).
[13] W. G. Unruh and R. M. Wald, Time and the interpretation of canonical quantum gravity, Physical Review D **40**, 2598 (1989).
[14] C. Rovelli, Quantum mechanics without time: A model, Physical Review D **42**, 2638 (1990).
[15] M. Reisenberger and C. Rovelli, Spacetime states and covariant quantum theory, Physical Review D **65**, 125016 (2002), arXiv:gr-qc/0111016 [gr-qc].
[16] F. Hellmann, M. Mondragon, A. Perez, and C. Rovelli, Multiple-event probability in general-relativistic quantum mechanics, Physical Review D **75**, 084033 (2007), arXiv:gr-qc/0610140 [gr-qc].
[17] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum time, Physical Review D **92**, 045033 (2015), arXiv:1504.04215.
[18] P. A. Höhn and A. Vanrietvelde, How to switch between relational quantum clocks, New Journal of Physics **22**, 123048 (2020).
[19] P. A. Höhn, Switching Internal Times and a New Perspective on the 'Wave Function of the Universe', Universe **5**, 116 (2019), arXiv:1811.00611 [gr-qc].
[20] P. A. Hoehn, A. R. H. Smith, and M. P. E. Lock, The trinity of relational quantum dynamics (2019), arXiv:1912.00033 [quant-ph].
[21] P. A. Hoehn, A. R. H. Smith, and M. P. E. Lock, Equivalence of approaches to relational quantum dynamics in relativistic settings (2020), arXiv:2007.00580 [gr-qc].
[22] E. Castro-Ruiz, F. Giacomini, A. Belenchia, and C. Brukner, Quantum clocks and the temporal localisability of events in the presence of gravitating quantum systems, Nature Commun. **11**, 2672 (2020), arXiv:1908.10165 [quant-ph].
[23] C. Branciard, M. Araújo, A. Feix, F. Costa, and Č. Brukner, The simplest causal inequalities and their violation, New Journal of Physics **18**, 013008 (2016).
[24] Ä. Baumeler, A. Feix, and S. Wolf, Maximal incompatibility of locally classical behavior and global causal order in multiparty scenarios, Physical Review A **90**, 042106 (2014).
[25] Č. Brukner, Bounding quantum correlations with indefinite causal order, New Journal of Physics **17**, 083034 (2015).
[26] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Physical Review A **88**, 022318 (2013).
[27] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. A. Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, Experimental superposition of orders of quantum gates, Nature Communications **6**, 7913 (2015).
[28] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, Experimental verification of an indefinite causal order, Science Advances **3**, e1602589 (2017).
[29] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, Witnessing causal nonseparability, New Journal of Physics **17**, 102001 (2015).
[30] K. Wei, N. Tischler, S.-R. Zhao, Y.-H. Li, J. M. Arrazola, Y. Liu, W. Zhang, H. Li, L. You, Z. Wang, Y.-A. Chen, B. C. Sanders, Q. Zhang, G. J. Pryde, F. Xu, and J.-W. Pan, Experimental quantum switching for exponentially superior

quantum communication complexity, Phys. Rev. Lett. **122**, 120504 (2019).
[31] K. Goswami, Y. Cao, G. A. Paz-Silva, J. Romero, and A. G. White, Increasing communication capacity via superposition of order, Phys. Rev. Research **2**, 033292 (2020).
[32] M. M. Taddei, J. Cariñe, D. Martínez, T. García, N. Guerrero, A. A. Abbott, M. Araújo, C. Branciard, E. S. Gómez, S. P. Walborn, L. Aolita, and G. Lima, Computational advantage from the quantum superposition of multiple temporal orders of photonic gates, PRX Quantum **2**, 010320 (2021).
[33] G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M. Zych, Č. Brukner, and P. Walther, Experimental entanglement of temporal orders (2017), arXiv:1712.06884.
[34] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, Indefinite Causal Order in a Quantum Switch, Physical Review Letters **121**, 090503 (2018), arXiv:1803.04302.
[35] Y. Guo, X.-M. Hu, Z.-B. Hou, H. Cao, J.-M. Cui, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, and G. Chiribella, Experimental Transmission of Quantum Information Using a Superposition of Causal Orders, Physical Review Letters **124**, 030502 (2020).
[36] R. P. Feynman, Quantum mechanical computers, Optics news **11**, 11 (1985).
[37] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation* (American Mathematical Society, USA, 2002).
[38] N. P. Breuckmann and B. M. Terhal, Space-time circuit-to-Hamiltonian construction and its applications, Journal of Physics A: Mathematical and Theoretical **47**, 195304 (2014), arXiv:1311.6101.
[39] L. Caha, Z. Landau, and D. Nagaj, Clocks in Feynman's computer and Kitaev's local Hamiltonian: Bias, gaps, idling, and pulse tuning, Physical Review A **97**, 062306 (2018).
[40] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Transforming quantum operations: Quantum supermaps, EPL (Europhysics Letters) **83**, 30004 (2008).
[41] P. Perinotti, Causal structures and the classification of higher order quantum computations, in *Time in Physics*, edited by R. Renner and S. Stupar (Springer International Publishing, Cham, 2017) pp. 103–127.
[42] A. Bisio, G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum networks: General theory and applications, Acta Physica Slovaca **61**, 273 (2011), arXiv:1601.04864 [quant-ph].
[43] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009), arXiv:0904.4483 [quant-ph].
[44] W. Yokojima, M. T. Quintino, A. Soeda, and M. Murao, Consequences of preserving reversibility in quantum superchannels, Quantum **5**, 441 (2021), arXiv:2003.05682 [quant-ph].
[45] W. F. Stinespring, Positive functions on C*-algebras, Proceedings of the American Mathematical Society **6**, 211 (1955).
[46] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
[47] J. D. Bekenstein, Black Holes and Entropy, Physical Review D **7**, 2333 (1973).
[48] L. J. Garay, Quantum gravity and minimum length, International Journal of Modern Physics A **10**, 145 (1995), arXiv:gr-qc/9403008 [gr-qc].
[49] N. Bao, S. M. Carroll, and A. Singh, The Hilbert space of quantum gravity is locally finite-dimensional, International Journal of Modern Physics D **26**, 1743013 (2017), arXiv:1704.00066v2.
[50] N. Gisin, Indeterminism in Physics, Classical Chaos and Bohmian Mechanics: Are Real Numbers Really Real?, Erkenntnis (2019), arXiv:1803.06824.
[51] D. Marolf, Group averaging and refined algebraic quantization: Where are we now?, in *The Ninth Marcel Grossmann Meeting* (2002) pp. 1348–1349, arXiv:gr-qc/0011112.
[52] C. E. Dolby, The conditional probability interpretation of the hamiltonian constraint, arXiv preprint gr-qc/0406034 (2004).
[53] We will not worry here about our use of non-normalisable states in the infinite dimensional case.
[54] J. Barrett, R. Lorenz, and O. Oreshkov, Cyclic quantum causal models, Nature Communications **12**, 885 (2021), arXiv:2002.12157 [quant-ph].
[55] Ä. Baumeler and S. Wolf, The space of logically consistent classical processes without causal order, New Journal of Physics **18**, 013036 (2016).
[56] T. Purves and A. J. Short, Quantum theory cannot violate a causal inequality (2021), arXiv:2101.09107 [quant-ph].
[57] J. Wechs, H. Dourdent, A. A. Abbott, and C. Branciard, Quantum circuits with classical versus quantum control of causal order (2021), arXiv:2101.08796 [quant-ph].
[58] J. Bavaresco, M. Murao, and M. T. Quintino, Unitary channel discrimination beyond group structures: Advantages of sequential and indefinite-causal-order strategies (2021), arXiv:2105.13369 [quant-ph].
[59] We do not consider the more general case in which one agent can control the order of the other agents.

Appendix A: Discretization and normalization operators

An important difference between continuous and discrete clocks is the fact that integrals $\int \mathrm{d}t$ pick up prefactors when changing the integration variable while sums do not pick up such a prefactor under change of summation index. Consider again the example from the main text of the history state $|\Psi\rangle\rangle = \int \mathrm{d}t_A |t_A\rangle_{c_A} \otimes |2t_A\rangle_{c_B}$ with perspectival states ${}_{c_A}\langle t_A|\Psi\rangle\rangle = |2t_A\rangle$ and ${}_{c_B}\langle t_B|\Psi\rangle\rangle = \int \mathrm{d}t_A |t_A\rangle\langle t_B|2t_A\rangle = \frac{1}{2}|1/2\ t_B\rangle$. If we consider a naive discretization of the above example of the form $|\Psi\rangle\rangle = \sum_k |k\rangle_{c_A} \otimes |2k\rangle_{c_B}$ we find ${}_{c_A}\langle t_A|\Psi\rangle\rangle = |2t_A\rangle$ and for even values of t_B we find ${}_{c_B}\langle t_B|\Psi\rangle\rangle = |1/2\ t_B\rangle$. Hence, the continuous and discrete version of ${}_{c_B}\langle t_B|\Psi\rangle\rangle$ differ by a factor $\frac{1}{2}$.

The second important issue arises from the use of approximations like time-binning, i.e. to assign every continuous time state $|t\rangle$ to the closest discrete time state, for discretization. Such procedures are not injective: Continuous times $|t+\delta t\rangle$ and $|t\rangle$ with very small δt will in general get mapped to the same discrete time state. This means that the discretization procedure itself can change the normalization and inner product of states. Such artifacts of the discretization procedure can be countered by the introduction of normalization operators.

For well-synchronized clocks with constant and same ticking speed, the aforementioned discretization artifacts can usually be avoided. However, we will show now issues one encounters in the context of different or varying clock ticking speeds. As a warm-up, let us consider again a history state of a clock that ticks twice as fast as another clock: $\int \mathrm{d}t |t\rangle_{c_A} \otimes |2t\rangle_{c_B}$. A first guess for a discretization might be something of the form $|\Psi\,\rangle\rangle = \sum_k |k\rangle_{c_A} \otimes |2k\rangle_{c_B}$ with k taking integer values. This would not be an acceptable discretization in our approach: Our postulates demand that agent B sees a state for each time value of their clock, and that this state evolves via unitary time evolution. However, ${}_{c_B}\langle t|\Psi\,\rangle\rangle = 0$ for t odd makes this impossible.

One possible approach to fix this issue might be to instead use $|\Psi\,\rangle\rangle = \sum_k |k\rangle_{c_A} \otimes |k\rangle_{c_B}$ and keep a note that says that the times on B's clock must be multiplied by another factor of 2 to obtain the "real" time of Bob. Such a fix is very unappealing and goes against the idea that the clock states directly reflect the time of the agents, up to rounding error. Also, in the context of varying ticking rates the implementation of such a fixing strategy can become very complicated. The situation becomes even worse in the context of superpositions of histories: Here, the mapping of $|k\rangle_{c_B}$ to the actual value of Bob's clock might depend on the branch of the superposition and the same $|k\rangle_{c_B}$ might correspond to vastly different times on B's clock. An example might be a history state that is a superposition of A's clock being twice as fast and B's clock being twice as fast, i.e. $\int \mathrm{d}t(\alpha|2t\rangle \otimes |t\rangle + \beta|t\rangle \otimes |2t\rangle)$. Obviously, now a "fix" like $\sum_k(\alpha|k\rangle \otimes |k\rangle + \beta|k\rangle \otimes |k\rangle)$ cannot work.

Let us look for a good history state that can describe discrete clocks of different ticking rates. We would like the clock states to directly tell us the time of the clock. Also, as we argued before, our discretization of clocks is not allowed to leave out any times. Then, to describe clocks of different speeds, one is left with the option to instead repeat times: A discretization of $\int \mathrm{d}t|2t\rangle_{c_A} \otimes |t\rangle_{c_B}$ might be $|\Psi\,\rangle\rangle = \sum_k |k\rangle_{c_A} \otimes |\lfloor \frac{k}{2} \rfloor\rangle_{c_B}$, with $\lfloor \bullet \rfloor$ meaning "rounded down". As a specific example, let us consider

$$|\Psi\rangle\rangle = |0\rangle \otimes |0\rangle + |1\rangle \otimes |0\rangle + |2\rangle \otimes |1\rangle + |3\rangle \otimes |1\rangle + |4\rangle \otimes |2\rangle + |5\rangle \otimes |2\rangle \ldots \tag{A1}$$

This procedure of repeating times does have a nice interpretation: One can interpret the clock states $|k\rangle$ as the number of ticks the agent has heard so far. As A's clock is twice as fast, B hears the first tick when A already hears the second.

Let us see what the states for the different perspectives look like. We have

$${}_{c_A}\langle 0|\Psi\,\rangle\rangle = |0\rangle_{c_B}, \quad {}_{c_A}\langle 1|\Psi\,\rangle\rangle = |0\rangle_{c_B}, \quad {}_{c_A}\langle 2|\Psi\,\rangle\rangle = |1\rangle_{c_B}, \quad {}_{c_A}\langle 3|\Psi\,\rangle\rangle = |1\rangle_{c_B}, \quad {}_{c_A}\langle 4|\Psi\,\rangle\rangle = |2\rangle_{c_B}, \quad \ldots$$

This fits to the interpretation that whenever B hears one tick, A already hears the second tick. Note that the states are properly normalized. For B's perspective we find

$${}_{c_B}\langle 0|\Psi\,\rangle\rangle = |0\rangle_{c_A} + |1\rangle_{c_A}, \qquad {}_{c_B}\langle 1|\Psi\,\rangle\rangle = |2\rangle_{c_A} + |3\rangle_{c_A}, \qquad {}_{c_B}\langle 2|\Psi\,\rangle\rangle = |4\rangle_{c_A} + |5\rangle_{c_A}, \qquad \ldots$$

First we note that these states are not properly normalized. But this can be easily fixed with a normalization operator $N_{t_B}^{(B)} = \frac{1}{\sqrt{2}}\mathbb{1}$. Indeed, this normalization factor arises because we map continuous times $|k+\Delta t\rangle_{c_B}$,with k an integer, $0 \leq \Delta t < 1$, to the same discrete state $|k\rangle_{c_B}$, as mentioned previously. Furthermore we note that B "coherently interpolates" between the two times of A that are consistent with B's time. Also this is reasonable: As B cannot have "which-time"-information about A's clock without a measurement (in analogy to which-path-information), B puts the two possible times in superposition.

Appendix B: About the physical projector and its relation to unitary time evolution

It is unclear whether the unitaries $\mathcal{U}_X(t'_X, t_X)$ can always be chosen such that they satisfy a nice relationship, similar to Equation (9), with $\hat{P}_H$. The examples considered in Ref. [22] would seem to suggest an equation of a form like

$$\langle t'_X|\hat{P}_H|t_X\rangle \stackrel{?}{=} (N^{(X)}_{t'_X})^{-1}\,\mathcal{U}_X(t'_X, t_X)N^{(X)}_{t_X}. \tag{B1}$$

However, we will show now that there exist choices of history states $|\Psi\rangle\rangle$ and time evolutions $\mathcal{U}_X(t'_X, t_X)$ that are compatible with our framework as presented in Section IV A, but do not satisfy Equation (B1).

If Equation (B1) was true we could alternatively write

$$\hat{P}_H = \sum_{t,t'} |t'\rangle\langle t|_{cA} \otimes (N^{(A)}_{t'})^{-1}\,\mathcal{U}_A(t', t)N^{(A)}_t, \tag{B2}$$

and a similar decomposition held for all other agents. Looking at two different ways to write out $\langle t'_A|\langle t'_B|P_H|t_A\rangle|t_B\rangle$, we have (using the notation $N_X(t) := N^{(X)}_t$)

$$\langle t'_B|N_A^{-1}(t'_A)\,\mathcal{U}_A(t'_A, t_A)N_A(t_A)|t_B\rangle = \langle t'_A|N_B^{-1}(t'_B)\,\mathcal{U}_B(t'_B, t_B)N_B(t_B)|t_A\rangle. \tag{B3}$$

By explicitly plugging in normalization operators N_X and unitaries $\mathcal{U}_X$ from the example in Section V B we can see that Equation (B3) does not hold for this representation of the quantum switch. More specifically, taking $t'_B = 3$, $t_B = 2$, $t'_A = 5$, $t_A = 4$, we obtain

$$\langle t_B = 3|N_A^{-1}(5)\,\mathcal{U}_A(5, 4)N_A(4)|t_B = 2\rangle = \sqrt{2}(\langle 3|T'_4|2\rangle)|0\rangle\langle 0| \otimes U_B = \sqrt{2}|0\rangle\langle 0| \otimes U_B, \tag{B4}$$

which is not equal to

$$\langle t_A = 5|N_B^{-1}(3)\,\mathcal{U}_B(3, 2)N_B(2)|t_A = 4\rangle = \frac{1}{\sqrt{2}}(\langle 5|T'_2|4\rangle)|0\rangle\langle 0| \otimes U_A = \frac{1}{\sqrt{2}}|0\rangle\langle 0| \otimes U_A. \tag{B5}$$

Here, we extended T'_i from the main text to act like the clock ticking operator T, when applied to $|j\rangle$ with $j \neq i-1, i, i+1$.

Therefore, Equation (B1) does not hold. If one exchanged the normalization operators in Equation (B1) with their inverses or adjoints, the $\sqrt{2}$ factors would change, but we would still have the U_A vs U_B-mismatch. This mismatch is caused by having both agents consider a time step that does not correspond to their own time of action $t^* - 1 \mapsto t^*$, with $t^* = 4$, but compatible with a time step at which one agent sees the other agent act. However, our extension of T'_2 and T'_4 to the full clock space was quite arbitrary, and one can make other choices for which Equations (B5) and (B4) simple evaluate to zero. Therefore, for this particular choice of extension and agent times, the contradiction disappears.

In general, it is unclear whether it is impossible to satisfy Equation (B1) or a similar equation. In our operational setting, we assumed that the clocks are initialized to the states $|0\rangle$. Therefore, the states that emerge during the protocol do not probe the full input space of the $\mathcal{U}_X(t'_X, t_X)$. In other words, there are several choices for $\mathcal{U}_X(t'_X, t_X)$ that are compatible with our assumptions in Section IV A. We leave for future work the question of whether there exists a choice of $\mathcal{U}_X(t'_X, t_X)$ that simultaneously satisfies our postulates and a relation similar to Equation (B1). It is conceivable that further postulates that determine how the $\mathcal{U}_X(t'_X, t_X)$ act if not all clocks were initialized to time 0 at the beginning of the protocol further restrict the set of allowed process matrices and causal reference frames.

Appendix C: Coherently controlled causal order for an arbitrary number of agents

1. Consistency requirements

In this appendix we show explicitly how to implement coherently controlled causal order for an arbitrary number of agents. We start with some general considerations concerning the coherent superposition of quantum combs. Afterwards we present each of the three conceptual steps of the implementation, i.e. desynchronization, application of the combs and resynchronization, in detail.

As mentioned in the main text, we work in a purified description of quantum theory and quantum processes in particular. Furthermore, our goal is to describe a coherently controlled superposition of different causal orders. Therefore, we will focus on definite causal orders that are described as sequences of unitary channels with memory.

That means we will consider combs $\widetilde{\mathcal{G}}_k$, $1 \leq k \leq M$, that are given by sequences of unitaries $V_j^{(k)}$ with memory. Each of the combs $\widetilde{\mathcal{G}}_k$ has a definite order of the agents that we describe via a permutation π_k. More specifically, the j-th agent in comb k is given by $A_{\pi_k(j)}$. The $V_j^{(k)}$ act trivially on the agents' ancillas, while the agents' unitaries U_j do not act on the memory of the comb, see Figure 6 for a tripartite example. Leaving identity operations on the ancillary systems of the agents and the comb memories implicit, our combs are thus

$$\widetilde{\mathcal{G}}_k(U_1, U_2, \dots U_N) = V_N^{(k)} U_{\pi_k(N)} V_{N-1}^{(k)} \dots V_1^{(k)} U_{\pi_k(1)} V_0^{(k)}, \tag{C1}$$

as explained in the main text.

Introducing an M-dimensional control system $\mathcal{H}_{S_C}$, we would like to consider a controlled superposition of these M combs in the form [56–58]

$$\mathcal{G}(U_1, \dots U_N) = \sum_{k=1}^{M} |k\rangle\langle k|_{S_C} \otimes \widetilde{\mathcal{G}}_k(U_1, U_2, \dots U_N) = \sum_{k=1}^{M} |k\rangle\langle k|_{S_C} \otimes V_N^{(k)} U_{\pi_k(N)} V_{N-1}^{(k)} \dots V_1^{(k)} U_{\pi_k(1)} V_0^{(k)}. \tag{C2}$$

While this definition looks intuitive, it leaves many details about the memories and ancillary systems implicit. Operationally, it is not directly clear what it means to put different combs into a controlled superposition. What if the dimensions do not match? How do we treat the different memory systems of the different combs? In the main text, we made the assumption that all the relevant dimensions agree and that the combs can use the same memory. While this is true for many examples of interest, e.g. optical experiments in which all agents act on the polarization of a photon and no memory is involved at all, here we will discuss relaxations of these assumptions. After all, as we pointed out in the main text, these assumptions are just necessary to make the definition of a controlled superposition of combs well-defined, but the protocol does not care about the implicit details in Equation (C2).

Therefore, we will now consider the compatibility conditions of the combs and the dimensions in further detail. First of all, all combs need the same dimension for the input from the causal past, otherwise they cannot receive the same input from the causal past. Similarly, they need the same dimension for the output to the causal future. Furthermore, the dimension of the input that an agent receives from the combs has to be the same for all combs. Similarly, the dimension of the output that the agent provides to the combs has to be independent of the comb. Otherwise, the agent could use the input or output dimension to find out that some of the combs are not used.

For notational convenience, in Section IV we assumed that the dimension of the input to the causal structure (i.e. P in the left half of Figure 6) and the dimension of its output to the causal future (i.e. F in the left half of Figure 6) are the same, and we called the corresponding physical system the main system $\mathcal{H}_S$. In a general sequence of unitaries with memory, however, the initial and the final dimension do not necessarily agree, since the dimensions of the agents' ancillary systems might change after they apply their unitary actions. As we model the agents' actions as unitaries, a change in the ancillary dimension of the agents is equivalent to a change of the dimension of the comb system that the agents interact with. We will also refer to sequences of unitary channels with memory that have such changes in dimension as *combs*. One way to avoid this problem might be to impose well-motivated physical restrictions. For example, one might demand that the physical information carrier that the comb provides to the agents is not replaced by the agents, only its state. And that the agents give that physical information carrier with the changed state back to the comb. Or that the agents can use some reversible operations that do not require access to an ancillary system. While such assumptions may be appealing operationally, mathematically they are restrictions. Can we weaken these restrictions?

In the following, we will consider two different strategies to weaken the restrictions. The first strategy requires extending our formalism to the case that the dimensions of subsystems can change if the agents apply their unitaries. Such an extension is important in its own right. However, our considerations here will be rather speculative. They suggest that it must be operationally possible to make sense of superpositions involving different physical systems in each branch of the superposition. Mathematically, they guide us to an enlarged Hilbert space that combines the situations involving different physical systems in a direct sum.

The second strategy is mathematically more rigorous and operationally less exotic, but it requires the introduction of extra environment systems that run parallel to the combs. These changes fix the dimension mismatch at the cost of letting the agents act on these new environment systems too, i.e. giving the agents more power.

2. Strategy 1 to deal with changing dimensions: Extending our formalism

To treat causal structures in which some of the dimensions can change in between, one needs a generalization of the formalization of the scenario considered in Section IV. In this generalization, one allows the dimensions of the main system and the agents' ancillary systems to change.

The generalization affects the input spaces and output spaces of the normalization operators and the agents' perspectival time-evolution operators, as well as the history state. The Hilbert spaces of the inputs and outputs of the operators and the history state depend on whether some agents' operations have already been applied or not, and in an indefinite causal structure several of these options might be combined coherently. The fact that we consider a purified description of quantum theory with unitary evolution implies that the overall dimension does not change. Therefore, the global Hilbert spaces will remain isomorphic. However, their factorization into constituent parts depends on whether the agents' operations have been applied or not. Therefore, one needs a way to deal with the situation that several factorizations of the global Hilbert space might matter simultaneously.

To provide a specific example, let us say that at time $|t\rangle_c$ agent A has not applied their operation yet. At this time, the input from the causal structure lives on a space $\mathcal{H}_A$ and has state $|\omega\rangle_A$, while the agent's ancillary system is $\mathcal{H}_{A'}$ and has state $|0\rangle_{A'}$. At the next time, i.e. $|t+1\rangle_c$, the agent has applied their unitary operation, and the output to the causal structure lives on Hilbert space $\mathcal{H}_{\widetilde{A}}$ and has state $|\widetilde{\omega}\rangle_{\widetilde{A}}$, while the ancillary system is now $\mathcal{H}_{\widetilde{A}'}$ and has state $|1\rangle_{\widetilde{A}'}$. Furthermore, we assume that the indefinite causal structure allows both times to appear simultaneously. This means we would like to make sense of a state of the form

$$|t\rangle_c \otimes |\omega\rangle_A \otimes |0\rangle_{A'} + |t+1\rangle_c \otimes |\widetilde{\omega}\rangle_{\widetilde{A}} \otimes |1\rangle_{\widetilde{A}'} \tag{C3}$$

One approach to this issue would be to introduce a canonical way to identify the different factorizations with each other, i.e. to use an isomorphism. Another approach might be to consider a direct sum of the different Hilbert space factorizations, i.e. $(\mathcal{H}_A \otimes \mathcal{H}_{A'}) \oplus (\mathcal{H}_{\widetilde{A}} \otimes \mathcal{H}_{\widetilde{A}'})$. The latter allows us to directly make mathematical sense of Eq. (C3), but it comes at the cost of making the Hilbert space larger.

However, this direct sum does have an operational meaning as the Hilbert space describing situations in which the systems before the agents' operation and the systems after the agents' operations might appear in superposition. Therefore, after repeating the same direct sum procedure for the other agents, one might postulate that this is the right global Hilbert space to consider. To extend the normalization operators to this enlarged Hilbert space, one might demand in addition that the normalization operators are block-diagonal, i.e. that they cannot mix up the systems before the agent's operation and the systems after. Postulates **S.1** and **S.2** only consider the beginning and the end at which only one factorization matters (i.e. none of the agents has used their operation yet, or all of them have, respectively). Therefore these postulates can be directly adapted without changing their meaning. Furthermore, one can directly carry over Postulates **N.1**, **N.2** and **N.3** by demanding them for each block separately, without changing their motivation. Similarly, Postulates **U.2** and **U.3** can be directly carried over: **U.2** just expresses an essential property of time evolution operators, while **U.3** just says that at the time of action, the agent uses their operation and ancilla, while the rest evolves in an uncorrelated way, not caring what that entails. The only real question seems to be whether time evolution should still be unitary on the enlarged global Hilbert space. Since our starting configuration only lives on one block (the one in which none of the agents have used their operation), the effective Hilbert space probed by time evolution might be much smaller than the enlarged global Hilbert space, at least for fixed choices of agent unitaries. Therefore, for fixed agent unitaries, one might restrict the input space and output space of the time evolution operators to just the effective subspace, and in particular demand unitarity just on this subspace.

However, the above considerations are speculative and should only be understood as hints or guidelines for a real formalization. We will leave the precise formalization and detailed investigation of these ideas for future work. We just keep the basic lessons from the previous paragraphs: Superpositions of different physical systems can be operationally meaningful (for example, they appear after scattering processes in relativistic quantum field theory, where it is unclear which particles will be detected before actually using the detector) and they can be mathematically described using a direct sum.

Furthermore, as we will see, for a fixed value of the control system the different factorizations will not appear simultaneously. The normalization operators will only be constant positive multiples of the identity depending only on the control value, while the perspectival time-evolutions can still be interpreted as controlled operations that act unitarily on/between the relevant factorizations, and respect the time of action requirements. Therefore, while we lack a rigorous formalization of our postulates for the case that the dimensions can change, our protocol will still

satisfy the purpose/intention of the postulates.

3. Strategy 2 to deal with changing dimensions: Add environment systems

As the above considerations are rather speculative, both mathematically and considering their operational meaning, we provide another approach to solve the issue of the mismatching dimensions. We indicated in the beginning that there is some ambiguity about the operational meaning of putting different chains of unitaries with memory in superposition: How are the memories and ancillary systems implemented physically? Can the physical systems implementing the memories be shared between the combs, or do they have to be separate, only existing for the right value of the control system and even then only at the right time? Such ambiguities about the operational meaning of superpositions of chains of unitaries with inconsistent dimensions give us some freedom, and we can use this freedom to choose the operational scenario that is the most natural.

Specifically, we are interested in extensions of the dimensions of the physical systems, with the requirement that these extensions are operationally rather trivial. Some issues related to mismatching dimensions might be fixed this way. For example, one can enlarge the Hilbert space on which a unitary acts by adding extra basis states, and letting the unitary act trivially on these extra basis states, while not changing its action on the other basis states. Alternatively, one can introduce additional environment systems. One can use this strategy to modify the combs and the agents to enforce that the dimensions match.

For example, let us say that the open end of $V_1^{(k)}$ towards agent $A_{\pi_k(2)}$ has a smaller dimension than the open end of $V_2^{(k)}$ towards the same agent. We can fix this mismatch by making the dimension of $V_1^{(k)}$ larger, letting $V_1^{(k)}$ act trivially on the extra dimensions as described above. However, we change the power of agent $A_{\pi_k(2)}$ by allowing this agent to act on these extra dimensions.

Alternatively one can add extra environment systems that run parallel to the combs, but are considered a part of an enlarged comb. Arguably, this changes the comb itself only in a trivial way: After all, we just add parallel running environment systems, only performing identity transformations. However, if we choose the dimensions right ("smallest common multiple"), and allow the agents to act on these environment systems too, then we can achieve that the open ends within each slot of the enhanced comb have the same dimension. One can use similar tricks to achieve that there is a unique memory dimension that is shared by all combs at all times.

Since the enlargement of the combs by adding environment systems that implement identity channels has a clear operational meaning, this is the variant we commit to from now on.

So far, we only focused on pure combs in the form of unitary sequences with memory. However, the ideas of this section also allow us to discuss mixed combs given as sequences of channels with memory. Here, the operational ambiguities are even larger: Our goal is to describe a coherent process involving superpositions of combs that are not fully coherent by themselves. First of all, we need to purify the combs by purifying the channels, compare Figure 9. This procedure is not fully unique. Our choice of purification procedures will be part of our definition for what it means to put these combs into a coherently controlled superposition.

We can use our trick to enlarge the combs by parallel environment systems to enforce that the dilation systems used to purify the combs have the same dimension for all combs. Therefore, in Figure 9, we can assume that the discarding operation is the same for all combs. This discarding operation will not be considered part of the protocol, but will be applied after the protocol. Furthermore, we can assume that the Hilbert space of the dilation state ($|\nu\rangle$ in Figure 9) is the same for all combs. Then, by modifying $V_0^{(k)}$ by adding an extra unitary on the wire of the dilation input, we can assume that the dilation input is the same for all combs. In Figure 9, this dilation input is exactly the state called $|\nu\rangle$, which is independent of the comb index k. This state will not be considered part of the protocol itself. Instead, we consider it to be a part of the input state to the causal structure (i.e. a part of $|\psi(0,\dots,0)\rangle$ in the notation of Section IV A).

Our trick of enlarging the combs with non-evolving parallel environments allowed us to enforce that the relevant dimensions agree. We will now provide a mathematical formalization of this situation.

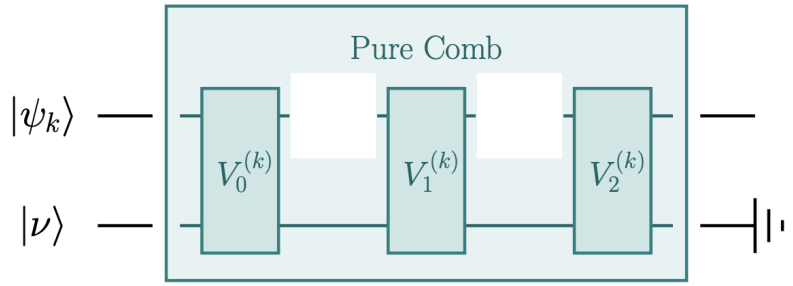


FIG. 9: The relation between the pure and mixed combs considered here: The dilation environment input $|\nu\rangle$ is treated as part of an extended main system that is the input to the causal structure. The partial trace over the environment output is only applied after the main protocol has finished. The unitaries with memory are a pure comb that we handle just as in the previous sections. One can assume that the dilation inputs for all combs are all the same, or that each comb has its own one and that the dilation inputs of the other combs get discarded.

The comb memory wire parallel to the action of agent A_j is called E_j. As argued before, it is independent of the the comb index k. Furthermore, we also argued that the ancillary system dimension of the agents do not change. We will call the ancillary systems $\mathcal{H}_{A'_j}$ and their collection $\mathcal{H}_{S'} := \mathcal{H}_{A'_1} \otimes \cdots \otimes \mathcal{H}_{A'_N}$. Hence, the combs can be written as

$$\widetilde{\mathcal{G}}_k(U_1, U_2, \dots U_N) = \tag{C4}$$
$$(V_N^{(k)} \otimes \mathbb{1}_{S'})(U_{\pi_k(N)} \otimes \mathbb{1}_{E_{\pi_k(N)}} \otimes \mathbb{1}_{S'\setminus\pi_k(N)})(V_{N-1}^{(k)} \otimes \mathbb{1}_{S'}) \dots (V_1^{(k)} \otimes \mathbb{1}_{S'})(U_{\pi_k(1)} \otimes \mathbb{1}_{E_{\pi_k(1)}} \otimes \mathbb{1}_{S'\setminus\pi_k(1)})(V_0^{(k)} \otimes \mathbb{1}_{S'}),$$

compare to [56–58]. Here, $\mathbb{1}_{S'\setminus\pi_k(j)}$ is the identity on all ancillary systems except on the ancillary system of agent $\pi_k(j)$. Leaving identity operations on the ancillas implicit for notational convenience we thus arrive at Equation (38) from the main text.

In what follows we describe a general procedure for writing down a history state complying with our axioms from Section IV A. While there are potentially many ways to write down such history states, our goal was to pick one with with a notation that is as simple as possible for an arbitrary number of agents. This means the procedure will not be as efficient or short as possible, but will use indices and notation that make it easier to discuss the local perspectives later on.

As explained in the main text, the history state decomposes into three parts as

$$|\Psi\rangle\rangle = |\Psi_{\text{desync}}\rangle\rangle + |\Psi_{\text{combs}}\rangle\rangle + |\Psi_{\text{resync}}\rangle\rangle, \tag{C5}$$

and we will now consider each part separately.

4. Desynchronizing the clocks

In the first step of the protocol we use the control degree of freedom to desynchronize the clocks such that the agents are put into the right order. It will be helpful to manipulate the clocks such that consecutive agents are *two* ticks apart because between the actions of two consecutive agents there is a unitary $V_j^{(k)}$ of the comb. To desynchronize the clocks, we will partially freeze them in time. More specifically, we start from $|0, 0, \dots, 0\rangle \otimes |\psi\rangle_S$. At first all the clocks make two synchronized step to $|2, 2, \dots, 2\rangle \otimes |\psi\rangle_S$. For the desynchronization procedure we consider a history state of the following form:

$$|\Psi_{\text{desync}}\rangle\rangle = |0, 0, \dots, 0\rangle_c \otimes |\psi\rangle_S + |1, 1, \dots, 1\rangle_c \otimes |\psi\rangle_S + \sum_{k=1}^{M} \sum_{j=2}^{T_0} |t(j)_1^{(k)}, t(j)_2^{(k)}, \dots, t(j)_N^{(k)}\rangle_c \otimes (|k\rangle\langle k| \otimes \mathbb{1})|\psi\rangle_S \tag{C6}$$

There are many desynchronization procedures one can choose from. Our goal is to pick one with a notation that is as simple as possible for an arbitrary amount of agents. This means the procedure will not be as efficient or short as possible, but will use indices and notation that make it easier to discuss the local perspectives later on. One such

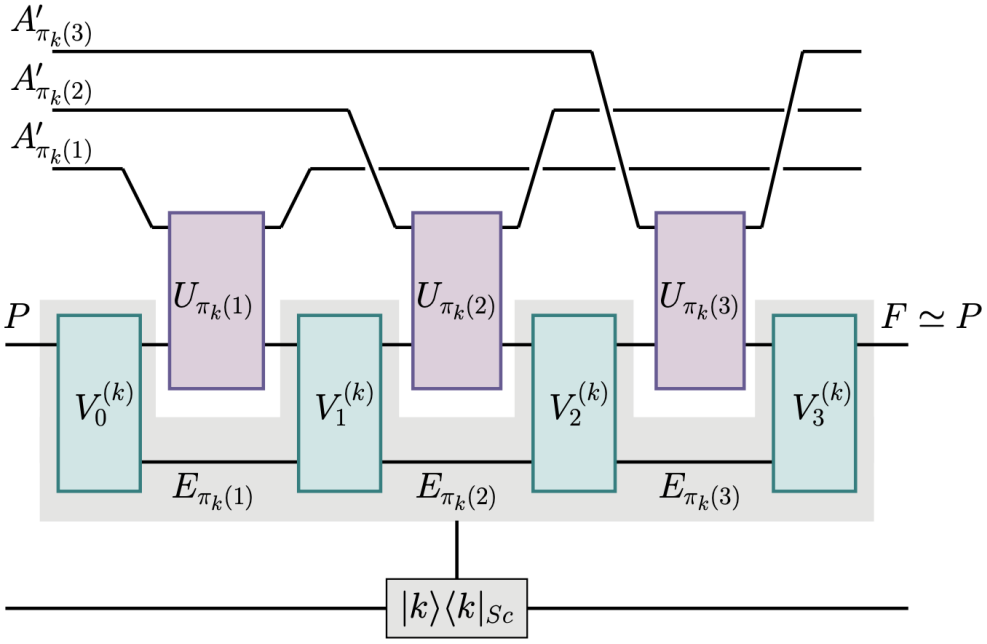


FIG. 10: This figure shows the general scenario for coherently controlled causal order in the tripartite case (see also [56–58]). The process $\mathcal{G}$ consists of a control degree of freedom, whose value k controls which comb $\widetilde{\mathcal{G}}_k$ is implemented. The order of the agents A_j in comb $\widetilde{\mathcal{G}}_k$ is described by a permutation π_k, i.e. the m-th agent in comb $\widetilde{\mathcal{G}}_k$ is agent $A_{\pi_k(m)}$. The combs $\widetilde{\mathcal{G}}_k$ are assumed to be pure, implemented via unitaries $V_j^{(k)}$ with memories and have have been extended (e.g. via ancillas) such that all the relevant dimensions are independent of k. This means the global past P and the global future F are independent of k. Furthermore, the dimension of the comb memory $E_{\pi_k(m)}$ running parallel to agent $A_{\pi_k(m)}$ is independent of k. The agents' operations U_j are shown in purple. They can act on the ancilla A'_j of the respective agent, but not on the ancillas of the other agents or the comb memory.

procedure works as follows:

The clock of the fastest agent, i.e. $\pi_k(1)$, continues to tick at the same rate as before. This we describe via

$$t(j)^{(k)}_{\pi_k(1)} = j \tag{C7}$$

For notational simplicity, we will desynchronize the clocks one after the other. Consider integers $2 \leq m \leq N$. We use the time range described by $2(m-2)\cdot N + 2 \leq j \leq 2(m-1)\cdot N + 1$ to slow down the clock of agent $\pi_k(m)$. More specifically, at times $2(m-2)\cdot N + 2 \leq j \leq 2(m-2)\cdot N + 2(m-1) + 2$, the clock of agent $\pi_k(m)$ completely freezes, while the clocks of the other agents march on. Except for that freezing period, the clock ticks at a normal rate. Overall, this can be described as follows ($m \geq 2$):

$$t(j)^{(k)}_{\pi_k(m)} = \begin{cases} j & \text{for } j \leq 2(m-2)N + 2 \\ 2(m-2)\cdot N + 2 & \text{for } 2(m-2)N + 2 \leq j \leq 2(m-2)N + 2(m-1) + 2 \\ j - 2(m-1) & \text{for } j \geq 2(m-2)N + 2(m-1) + 3 \end{cases} \tag{C8}$$

We choose the largest j to be

$$T_0 := 2(N-2)N + 2(N-1) + 4 + 4(N+1) = 2N^2 + 2(N+1) + 4, \tag{C9}$$

which includes $4(N+1)$ more well-synchronized ticks to make sure that for all k the clocks freezes are far way from the application of the combs. The desynchronization procedure is shown for $N = 4$ in Figure 11.

First, we note that two consecutive agents $\pi_k(m)$ and $\pi_k(m+1)$ are indeed two time steps apart at the end:

$$[j - 2([m+1]-1)] - [j - 2(m-1)] = -2$$

$\pi_k(1)$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
$\pi_k(2)$	0	1	2	2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
$\pi_k(3)$	0	1	2	3	4	5	6	7	8	9	10	10	10	10	10	11	12	13	14	15	16	17	18	19	20	21	22
$\pi_k(4)$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	18	18	18	18	18	18	19	20

FIG. 11: This figure shows the clock times during the desynchronization procedure for the special case $N = 4$. Time passes from left to right. Time freezing is marked in color. After the shown times only well-synchronized ticks happen.

Moreover, we made sure to construct the history state such that only one clock freezes simultaneously, i.e. time freezes of different clocks are well-separated from each other, and that no times are skipped.

Next, let us consider the local perspectives, i.e. ${}_{c_a}\langle t|\Psi_{\text{desync}}\rangle\rangle$. Let us expand $|\psi\rangle_S = \sum_k |k\rangle_{S_C}|\psi_k\rangle_{S_P}$. Then we have

$$|\Psi_{\text{desync}}\rangle\rangle = |0,0,\ldots,0\rangle_c \otimes |\psi\rangle_S + |1,1,\ldots,1\rangle_c \otimes |\psi\rangle_S + \sum_{k=1}^{M}\sum_{j=2}^{T_0} |t(j)_1^{(k)}, t(j)_2^{(k)}, \ldots, t(j)_N^{(k)}\rangle_c \otimes |k\rangle_{S_C}|\psi_k\rangle_{S_P} \tag{C10}$$

We now need to define the normalisation operators $N_t^{(A_j)}$ and the unitaries $\mathcal{U}_{A_j}(t,t')$ relating the perspectival states at different times. Without loss of generality, we will show how to construct those for the point of view of A_1. Define

$$\alpha_k(t) = \|\langle t|_{c_1} \sum_{j=2}^{T_0} |t(j)_1^{(k)}, t(j)_2^{(k)}, \ldots, t(j)_N^{(k)}\rangle_c\|. \tag{C11}$$

For the local perspectives, let us consider the non-trivial case given by $t \geq 2$. Furthermore, let us neglect for now the overlap with the comb-phase, as we will treat this later. Then we have that $\alpha_k(t) \neq 0$ because no time is skipped during desynchronization. We can therefore define

$$N_t^{(A_1)} = \sum_k \frac{1}{\alpha_k(t)} |k\rangle\langle k|_{S_C}, \tag{C12}$$

as the normalization operator, which then gives the perspectival state

$$|\psi^{A_1}(t)\rangle = \sum_k |\xi_k(t)\rangle_{c_{\setminus 1}} |k\rangle_{S_C}|\psi_k\rangle_{S_P}, \tag{C13}$$

where $|\xi_k(t)\rangle_{c_{\setminus 1}}$ is a normalized state proportional to $\langle t|_{c_1} \sum_{j=2}^{T_0} |t(j)_1^{(k)}, t(j)_2^{(k)}, \ldots, t(j)_N^{(k)}\rangle_c$. There exists a unitary relating $|\psi^{A_1}(t)\rangle$ with $|\psi^{A_1}(t+1)\rangle$. This unitary can be chosen to have the form $\mathcal{U}_{A_1}(t+1,t) = \sum_k u_{c,k}^{A_1} \otimes |k\rangle\langle k|_{S_c} \otimes \mathbb{1}_{S_P}$ because our protocol only changes the clock times conditioned on comb index k, but nothing else. Indeed, we can choose $u_{c,k}^{A_1}$ to be any unitary mapping $|\xi_k(t)\rangle_c \mapsto |\xi_k(t+1)\rangle$, and acting arbitrarily on other states. Furthermore, given comb index k, the protocol is independent of the initial state of the system, therefore the unitaries $\mathcal{U}_{A_1}(t+1,t)$ are independent of the initial system state since they are defined as controlled unitaries with the comb index k as control. At last, the actions of the agents have not been used yet.

The generality of our last arguments shows that there are many possible desynchronization protocols. We have merely chosen one that is easier to write down for an arbitrary number of agents. The price is that our protocol requires a lot of time steps. It is very likely that there exist more efficient protocols, e.g. protocols that freeze several clocks at once or insert less steps that only have synchronized clock ticks.

5. Application of the combs

Now we consider the application of the combs. The starting point is

$$\sum_{k=1}^{M} (|k\rangle\langle k| \otimes \mathbb{1})|\psi\rangle_S \otimes |T_0, T_0 - 2, \ldots, T_0 - 2(N-1)\rangle_{c_{\pi_k(1)},\ldots,c_{\pi_k(N)}}, \tag{C14}$$

with

$$|t_1, t_2, \ldots, t_N\rangle_{c_{\pi_k(1)},\ldots,c_{\pi_k(N)}} := \widetilde{U}_{\pi_k}|t_1, t_2, \ldots, t_N\rangle_c, \tag{C15}$$

where $\widetilde{U}_{\pi_k}$ is the unitary implementing the permutation on the Hilbert spaces of the local clocks.

For this part of the protocol, the clocks will always tick in synchronization. Then all agents see the following sequence of time evolutions:

$$V_0^{(k)} \otimes T^{\otimes(N-1)},\ U_{\pi_k(1)} \otimes T^{\otimes(N-1)},\ V_1^{(k)} \otimes T^{\otimes(N-1)},\ U_{\pi_k(2)} \otimes T^{\otimes(N-1)},\ \ldots,\ U_{\pi_k(N)} \otimes T^{\otimes(N-1)},\ V_N^{(k)} \otimes T^{\otimes(N-1)} \tag{C16}$$

So the time of action for each agent is $t^* = T_0 + 2$. For completeness, let us describe the time-evolutions the agents see in more detail. For that purpose, we start at $\tau := T_0 - 2(N-1)$. This is the time that the slowest clock shows right before entering $|\Psi_{\text{combs}}\rangle\rangle$. Then the unitary time evolution that agent j sees is given by (p a non-negative integer)

$$\mathcal{U}_{A_j}(\tau + p + 1, \tau + p) = \sum_{k=1}^{M} |k\rangle\langle k|_{Sc} \otimes T^{\otimes(N-1)} \otimes W_{A_j}^{(k)}(p+1, p) \tag{C17}$$

Here, the $W_{A_j}^{(k)}(p+1,p)$ are either $V_x^{(k)}$ for some x, or $U_{\pi_k(y)}$ for some y, or $\mathbb{1}$. The case $\mathbb{1}$ appears if for this specific time and comb index the agent is still in the desynchronization phase or already in the resynchronization phase, while for other comb indices the agent is the $|\psi_{\text{combs}}\rangle\rangle$ phase. This potential overlap is the reason why we pad $|\psi_{\text{desync}}\rangle\rangle$ and $|\psi_{\text{resync}}\rangle\rangle$ with many steps at which nothing happens, except for well-synchronized clock ticks. This overlap can happen because for different comb indices k, the agent may enter $|\psi_{\text{combs}}\rangle\rangle$ at very different times.

Let us analyze p in more detail. First, for a given comb index k, we determine the time of agent A_j, right before agent A_j enters $|\psi_{\text{combs}}\rangle\rangle$. For that purpose, let us mention the place of agent A_j in the causal order of comb k. This is $\pi_k^{-1}(j)$. Since τ is the time of the slowest clock right before entering the comb phase, and consecutive agents are two clock ticks apart from each other, the clock of agent A_j shows $\tau + 2\cdot(N - \pi_k^{-1}(j))$ right before entering the comb phase. In particular, $m_j^{(k)} := (N - \pi_k^{-1}(j))$ is the number of places in comb k that agent A_j is ahead of the last agent who acts. This can be rewritten as $j = \pi_k(N - m_j^{(k)})$.

Now we can determine the unitaries $W_{A_j}^{(k)}(p+1,p)$ in detail. Since we already pad $|\psi_{\text{desync}}\rangle\rangle$ and $|\psi_{\text{resync}}\rangle\rangle$ with many steps that have just synchronized clock ticks, we can directly begin with applying the unitaries of the combs and the agents. Since p is the value that the agent is ahead of the smallest comb entering time τ, p is of the form:

$$p = 2m_j^{(k)} + \text{“ number of the comb/agent unitary that gets applied ”}$$

Here, $2m_j^{(k)}$ is again the offset to the comb entering time provided by having a faster clock. Then the unitary $W_{A_j}^{(k)}(p+1,p)$ is given by ($N \geq x \geq 0$ a non-negative integer, $N \geq y \geq 1$ a positive integer)

$$\begin{aligned}
&W_{A_j}^{(k)}(2m_j^{(k)} + 2x + 1, 2m_j^{(k)} + 2x) = V_x^{(k)},\\
&W_{A_j}^{(k)}(2m_j^{(k)} + 2y, 2m_j^{(k)} + 2y - 1) = U_{\pi_k(y)},\\
&W_{A_j}^{(k)}(p+1, p) = \mathbb{1} \text{ for other values of } p.
\end{aligned} \tag{C18}$$

The corresponding part of the history state looks as follows

$$\begin{aligned}
|\Psi_{\text{combs}}\rangle\rangle = &\sum_{k=1}^{M} |T_0 + 1, T_0 - 1, \ldots, T_0 - 2(N-1) + 1\rangle_{c_{\pi_k(1)},\ldots c_{\pi_k(N)}} \otimes \Big[|k\rangle\langle k| \otimes V_0^{(k)}\Big]|\psi\rangle_S +\\
&+ \sum_{k=1}^{M}\sum_{y=1}^{N} |T_0 + 2y, T_0 - 2 + 2y, \ldots, T_0 - 2(N-1) + 2y\rangle_{c_{\pi_k(1)},\ldots c_{\pi_k(N)}}\\
&\qquad \otimes \Big[|k\rangle\langle k| \otimes \Big(U_{\pi_k(y)} V_{y-1}^{(k)} \ldots U_{\pi_k(1)} V_0^{(k)}\Big)\Big]|\psi\rangle_S +
\end{aligned}$$

$$+\sum_{k=1}^{M}\sum_{x=1}^{N}|T_0+1+2x,T_0-1+2x,\ldots,T_0-2(N-1)+1+2x\rangle_{c_{\pi_k(1)},\ldots c_{\pi_k(N)}}$$
$$\otimes\left[|k\rangle\langle k|\otimes\left(V_x^{(k)}U_{\pi_k(x)}\ldots U_{\pi_k(1)}V_0^{(k)}\right)\right]|\psi\rangle_S \tag{C19}$$

Hence, all the combs get applied, see Equation (C19), each agent has a well defined time of action and the other agent's unitaries appear at most linearly in each parties perspective, see Equation (C16).

6. Resynchronization

Now we consider the final part of the protocol, the resynchronization step. We define

$$T_1 := T_0 + 2N + 1 \tag{C20}$$

such that the starting point is given by

$$\sum_{k=1}^{M}|T_1,T_1-2,\ldots,T_1-2(N-1)\rangle_{c_{\pi_k(1)},\ldots c_{\pi_k(N)}}\otimes\left[|k\rangle\langle k|\otimes\left(V_N^{(k)}U_{\pi_k(N)}\ldots U_{\pi_k(1)}V_0^{(k)}\right)\right]|\psi\rangle_S$$

To make sure that for all k the clock freezes are far apart from the application of the combs, we first insert $4(N+1)$ well-synchronized ticks. Afterwards, we choose the resynchronization to proceed exactly as the desynchronization, but with the order of agents reversed. By using the function $t(j)^{(k)}_{\pi_k(m)}$ from Equation (C8), this can be described by the history state

$|\Psi_{\text{resync}}\rangle\rangle =$

$$\sum_{k=1}^{M}\sum_{j=0}^{4(N+1)}|T_1+1+j,T_1-1+j,\ \ldots,\ T_1+1-2(N-1)+j\rangle_{c_{\pi_k(1)},\ \ldots c_{\pi_k(N)}}$$
$$\otimes\left[|k\rangle\langle k|\otimes\left(V_N^{(k)}U_{\pi_k(N)}\ldots U_{\pi_k(1)}V_0^{(k)}\right)\right]|\psi\rangle_S+ \tag{C21}$$
$$+\sum_{k=1}^{M}\sum_{j=0}^{T_0}|T_1+4N+6+t(j)^{(k)}_{\pi_k(N)},\ T_1+4N+4+t(j)^{(k)}_{\pi_k(N-1)},\ \ldots,\ T_1+2N+8+t(j)^{(k)}_{\pi_k(1)}\rangle_{c_{\pi_k(1)},\ \ldots c_{\pi_k(N)}}$$
$$\otimes\left[|k\rangle\langle k|\otimes\left(V_N^{(k)}U_{\pi_k(N)}\ldots U_{\pi_k(1)}V_0^{(k)}\right)\right]|\psi\rangle_S$$

Just as during the desynchronization process, nothing happens on the system and we can write the perspectival states and unitaries as $|\psi^{A_j}(t)\rangle = \sum_k |\xi_k(t)\rangle_c|k\rangle_{S_C}\widetilde{\mathcal{G}}_k(U_1,U_2,\ldots U_N)|\psi_k\rangle_{S_P}$ and $\mathcal{U}_{A_j}(t+1,t) = \sum_k \widetilde{V}_c^k\otimes|k\rangle\langle k|_{S_C}\otimes\mathbb{1}_{S_P}$. At the end, all clocks show the same time $T_1+T_0+2N+8.$, i.e. we end with the state

$$\sum_{k=1}^{M}|T_1+T_0+2N+8,\ldots T_1+T_0+2N+8\rangle\otimes\left[|k\rangle\langle k|\otimes\left(V_N^{(k)}U_{\pi_k(N)}\ldots U_{\pi_k(1)}V_0^{(k)}\right)\right]|\psi\rangle_S \tag{C22}$$

Since directly before the end we have several time steps at which nothing but clock ticks happen, we see that we obtain the coherent control of causal order we wanted in a way that is compatible with our postulates.

With this generic protocol we can implement any process describing coherent control of causal order within our Page-Wootters framework.

Chapter 3

Composition rules for quantum processes: a no-go theorem

Authors: Philippe Allard Guérin, Marius Krumm, Costantino Budroni, Časlav Brukner
Published in: New J. Phys. **21**, 012001 (2019)
DOI: https://doi.org/10.1088/1367-2630/aafef7
Preprint available as: arXiv:1806.10374
https://arxiv.org/abs/1806.10374

Contribution: I made essential contributions to all parts of the research. I provided the leading input for extending convex bilinearity to full bilinearity, and to extending the proof to all dimensions.

Abstract: A quantum process encodes the causal structure that relates quantum operations performed in local laboratories. The process matrix formalism includes as special cases quantum mechanics on a fixed background space-time, but also allows for more general causal structures. Motivated by the interpretation of processes as a resource for quantum information processing shared by two (or more) parties, with advantages recently demonstrated both for computation and communication tasks, we investigate the notion of composition of processes. We show that under very basic assumptions such a composition rule does not exist. While the availability of multiple independent copies of a resource, e.g. quantum states or channels, is the starting point for defining information-theoretic notions such as entropy (both in classical and quantum Shannon theory), our no-go result means that a Shannon theory of general quantum processes will not possess a natural rule for the composition of resources.

I. INTRODUCTION

Experimental tests with elementary quantum systems, most notably Bell tests, radically challenge the very notions of physical reality and cause-effect relations [1, 2]. Notwithstanding such fundamental novel effects, quantum mechanics still assumes a definite causal order of events. Namely, given two events, i.e. two operations performed locally in two quantum laboratories, say A and B, we always assume that they are either time-like separated, hence, A cannot signal to B or vice versa, or they are space-like separated, hence, they cannot signal in either direction.

Motivated by the problem of quantum gravity, operational formalisms have been proposed for computing the joint probabilities for the outcome of local experiments, without the assumption of a fixed space-time background [3–8]. *Process matrices* [6] are introduced as the most general class of multilinear mappings of local quantum operations into probability distributions. The process matrix formalism provides a unified description of causally ordered quantum mechanics (quantum states and quantum channels), but also includes experimentally relevant non-causal processes such as the quantum switch [7, 9–14]. Furthermore, the formalism predicts novel and potentially observable phenomena, such as the violation of so-called *causal inequalities* [6, 14–17].

Moreover, it has been proven that such processes are able to provide advantages for quantum information processing tasks, both for computation and communication [7, 18–24]. One would, then, expect that a theory of information can be developed also for processes. Such a theory would deal with, e.g., rates of information compression and communication, i.e., a process-analog of the classical and quantum Shannon theory. A fundamental assumption in classical and quantum Shannon theory [25, 26] is the availability of multiple independent copies of a resource (for example a classical source of random variables, a quantum state, or a channel), which is at the basis of the definition of information-theoretic entropy, i.e., Shannon or von Neumann entropy. To be more concrete, in the example of Schumacher's compression [25, 27], the optimal data compression of n samples of an independent and identically distributed quantum source ρ into $nS(\rho) + \delta$ qubits (with $\delta \to 0$ for $n \to \infty$), and the subsequent transmission, can be achieved only if the sender can act globally on multiple copies of the quantum state in which the information is encoded.

A natural question then arises, namely, whether a process matrix can be understood as a resource available in multiple (possibly identical) copies to experimenters, similarly to the example of Schumacher's compression above. Answering this question will provide us with deeper insight into the nature of process matrices. For instance, if we consider an experimental realization of a process, e.g., consisting of a sequence of optical elements as in photonic experiments [10, 11], one can easily imagine that it is possible to create two identical copies of the setup, and share them among the two parties. Alternatively, if one imagines that a process matrix does not only represent an experimental setup, but also the space-time structure [28–30], then it is harder to imagine how two "copies of spacetime" may be shared between the two parties. More generally, such a composition rule should not be only about identical copies, but it should also allows us to combine different processes.

It is important, at this point, to distinguish two different scenarios and their corresponding composition rules. On the one hand, one may simply ask what is the rule for composing different processes *independently*, with the requirement that experimenters act locally on each copy of the process; this rule is given by the tensor product. On the other hand, going back to the example of Schumacher's compression protocol, one may require that a single experimenter (or many experimenters for multipartite systems) has access to multiple copies of a process, in order to perform a protocol that involves global operations. We will see that the latter notion is incompatible with the definition of a process.

For quantum states, quantum channels, or for any collection of processes with the same definite causal order [31, 32], the parallel composition can be described by the tensor product. However, it is known that a parallel composition of process matrices via the tensor product can fail [33], as the resulting process matrix contains causal "double-loops" [6], which give rise to the "grandfather paradox", or equivalently, to unormalised probabilites.

In this work, we show that under weak assumptions (bilinearity, every output is a valid process matrix, reduction to the usual tensor product for definite causal structure) there exists no composition that allows the experimenters to have access to multiple shared processes. This result means that many information theoretic protocols relying on many copies of a resource have no straightforward generalization to process matrices.

II. PRELIMINARY NOTIONS

The most general operation that can be performed on a quantum system is represented by a quantum instrument, namely, a collection $\{\mathcal{M}_a\}_a$ of completely positive trace-nonincreasing maps that sum up to a trace-preserving

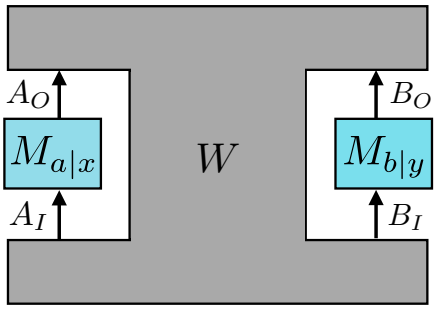


FIG. 1. Graphical representation of the probability rule Eq. (2).

map $\mathcal{M} = \sum_a \mathcal{M}_a$. An operation represented by the instrument $\{\mathcal{M}_a\}_a$ will give an output a with probability $P(a) = \mathrm{tr}[\mathcal{M}_a(\rho)]$ and transformation of the state $\rho \mapsto \mathcal{M}_a(\rho)/P(a)$. We admit the possibility of an input x, and label the corresponding operations as $\{\mathcal{M}_{a|x}\}_{a,x}$. Such maps can be represented as matrices via the Choi-Jamiołkowski isomorphism [34, 35]

$$\mathcal{M}_{a|x} \mapsto M_{a|x} = \sum_{ij} |i\rangle\langle j|^{A_I} \otimes \mathcal{M}_{a|x}(|i\rangle\langle j|)^{A_O}. \tag{1}$$

We will call $M_{a|x}$ the Choi matrix of $\mathcal{M}_{a|x}$ [36]. Consider a set of local operations, i.e., Choi matrices, $\{M^{\mathrm{A}}_{a|x}\}_{a,x}$ and $\{M^{\mathrm{B}}_{b|y}\}_{b,y}$, associated with Alice's and Bob's laboratories, where A denotes Alice's input-output space $\mathcal{H}_{\mathrm{A_I}} \otimes \mathcal{H}_{\mathrm{A_O}}$, and similarly for B. A process W is understood as the most general linear mapping of such operations into probabilities, which can be represented using the trace inner product as

$$p(ab|xy) = \mathrm{tr}\left[\left(M^{\mathrm{A}}_{a|x} \otimes M^{\mathrm{B}}_{b|y}\right) W^T\right], \tag{2}$$

where T denotes the transposition in the computational basis. A visual representation of this probability rule is given in Figure 1. In order to obtain valid probabilities, i.e., non-negative numbers summing up to one, for arbitrary operations $\{M^{\mathrm{A}}_{a|x}\}_{a,x}$, $\{M^{\mathrm{B}}_{b|y}\}_{b,y}$ (including operations that involve shared entangled ancillary systems), it can be proven [9] that the following constraints must be satisfied

$$W \geq 0\,, \tag{3}$$
$$\mathrm{tr}\, W = d_O = d_{\mathrm{A_O}} d_{\mathrm{B_O}}, \tag{4}$$
$${}_{\mathrm{B_I B_O}}W = {}_{\mathrm{A_O B_I B_O}}W, \tag{5}$$
$${}_{\mathrm{A_I A_O}}W = {}_{\mathrm{A_I A_O B_O}}W, \tag{6}$$
$$W = {}_{\mathrm{B_O}}W + {}_{\mathrm{A_O}}W - {}_{\mathrm{A_O B_O}}W, \tag{7}$$

where ${}_X W := \frac{\mathbb{1}^X}{d_X} \otimes \mathrm{tr}_X W$. The linear constraints in Eqs. (4)-(7) can be written in a more compact form as

$$L_V(W) = W, \tag{8}$$

where L_V is the projector onto the subspace of operators in $\mathcal{L}(\mathcal{H}_{\mathrm{AB}})$ that satisfy Eqs. (5)-(7). We will denote such a linear subspace as $L_V(\mathcal{L}(\mathcal{H}_{\mathrm{AB}}))$. This projector enforces the normalisation of probabilites, and can be interpreted as preventing the paradoxes that would occur in processes with "causal loops" [6]. It is also convenient to define $\mathcal{W} \subset \mathcal{L}(\mathcal{H}_{\mathrm{AB}})$ as the set of matrices that satisfy conditions in Eqs. (3)-(7), and similarly $\mathcal{W}'$ for the spaces $\mathrm{A'B'} := \mathcal{H}_{\mathrm{A_I'}} \otimes \mathcal{H}_{\mathrm{A_O'}} \otimes \mathcal{H}_{\mathrm{B_I'}} \otimes \mathcal{H}_{\mathrm{B_O'}}$. If ${}_{\mathrm{B_O}}W = W$, one can show that Bob cannot signal to Alice, i.e., $p(a|x,y) = p(a|x,y')$ for all a, x, y, y', we denote it as $\mathrm{A} \preceq \mathrm{B}$ and we say that the process is causally ordered [9]. Similarly, the case ${}_{\mathrm{A_O}}W = W$ correspond to the opposite causal order and it is denoted as $\mathrm{B} \preceq \mathrm{A}$. If ${}_{\mathrm{A_O B_O}}W = W$ we have at the same time $\mathrm{A} \preceq \mathrm{B}$ and $\mathrm{B} \preceq \mathrm{A}$, then W represents a bipartite quantum state and we have no-signaling in both directions.

Similarly, in the case of N parties $\mathrm{A}^1, \ldots, \mathrm{A}^N$, linear constraints can be written in the compact form [9]

$$L_{V_N}(W) := {}_{\left[1-\prod_{i=1}^N (1-\mathrm{A_O}^i + \mathrm{A_I}^i \mathrm{A_O}^i) + \prod_{i=1}^N \mathrm{A_I}^i \mathrm{A_O}^i\right]} W = W, \tag{9}$$

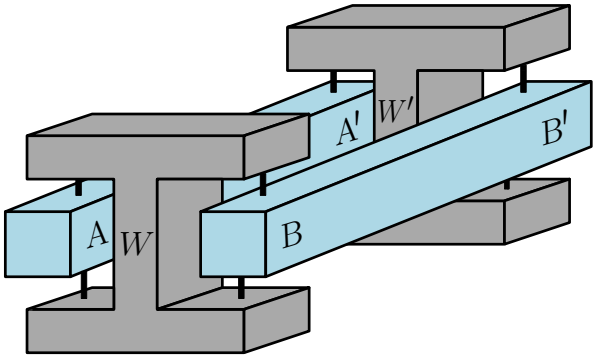


FIG. 2. The tensor product compostion rule $\mu(W, W') = W \otimes W'$. Here AA' is a composite party that can perform general quantum operations $\mathcal{L}(\mathcal{H}_{A_I A_I'}) \to \mathcal{L}(\mathcal{H}_{A_O A_O'})$, and similarly for BB'; the corresponding probabilities are given by Eq. (2). We shall show that this composition rule does not satisfy all requirements that we demand on such a rule.

where the index i runs through the different parties. Notice that if $W = W^1 \otimes W^2$, then the set $\{1, \ldots, N\}$ can be split as $\chi_1 \cup \chi_2$, with $\chi_1 \cap \chi_2 = \emptyset$, where χ_k indexes the parties appearing in W^k. Then

$$\begin{aligned} L_{V_N}(W_1 \otimes W_2) = & W_1 \otimes W_2 \Leftrightarrow \\ \left[1-\prod_{i\in\chi_1}(1-\mathrm{A_O}^i+\mathrm{A_I}^i\mathrm{A_O}^i)+\prod_{i\in\chi_1}\mathrm{A_I}^i\mathrm{A_O}^i\right] W_1 = W_1 & \text{ and } \left[1-\prod_{i\in\chi_2}(1-\mathrm{A_O}^i+\mathrm{A_I}^i\mathrm{A_O}^i)+\prod_{i\in\chi_2}\mathrm{A_I}^i\mathrm{A_O}^i\right] W_2 = W_2. \end{aligned} \tag{10}$$

A. Examples

The process matrix formalism allows one to treat quantum states, quantum channels, and even situations where the causal order is indefinite, in a unified way. For example, the process matrix associated to a quantum state ρ can be described as a single party process matrix, as $W = \rho^{A_I} \otimes \mathbb{1}^{A_O}$. The process matrix associated to N spatially separated copies of the state is a N-partite process $W = \prod_{i=1}^N \rho^{A_I^i} \otimes \mathbb{1}^{A_O^i}$, where each of the A_I^i and A_O^i are isomorphic. However, one could also consider the same W as a global single party process, with input Hilbert space $A_I = \prod_i A_I^i$, and output Hilbert space $A_O = \prod_i A_O^i$.

A quantum channel $\mathcal{C} : \mathcal{L}(\mathcal{H}_{A_O}) \to \mathcal{L}(\mathcal{H}_{B_I})$, connecting the output Hilbert space of Alice to Bob's input Hilbert space, can be described in process matrix language as $W = C^{A_O B_I}$, where C is the Choi matrix of the channel $\mathcal{C}$, as defined by Eq. (1). The process matrix describing N parallel uses of the channel $\mathcal{C}$ is simply $W = \prod_{i=1}^N C^{A_O^i B_I^i}$. Again, this process can be considered as a $2N$-partite process, or as a bipartite process with $A_O := \prod_i A_O^i$ and $B_I = \prod_i B_I^i$.

III. COMPOSITION RULES

From the above considerations, it seems that one could simply take the tensor product as a composition rule to obtain multipartite processes representing multiple independent copies of a resource. In fact, Eq. (10) implies that whenever the linear constraints are satisfied for both W_1 and W_2, then the corresponding multipartite constraints will be satisfied for $W_1 \otimes W_2$.

The situation is different, however, if we require W_1 and W_2 to be shared by *the same* parties. To keep the discussion simple, consider only two parties, Alice and Bob, who share two possible processes, $W_1 \in \mathcal{W}$ and $W_2 \in \mathcal{W}'$. We want now to create the composite process $\mu(W^1, W^2)$ such that it is still a bipartite one, i.e., Alice can access both the systems $\mathrm{A_I A_O}$ and $\mathrm{A_I}'\mathrm{A_O}'$, and Bob both $\mathrm{B_I B_O}$ and $\mathrm{B_I}'\mathrm{B_O}'$. If both processes have the same definite order, i.e., ${}_{\mathrm{A_O}}W_1 = W_1$ and ${}_{\mathrm{A_O}'}W_2 = W_2$, or the analogous condition with $\mathrm{B_O}, \mathrm{B_O}'$, then, we know from standard quantum theory that the right operation for composing such processes is $W_1 \otimes W_2$. This composition rule is represented in Fig. 2. One can easily prove that whenever the two processes do not have the same definite causal order, then $L_V(W^1 \otimes W^2) \neq W^1 \otimes W^2$, where L_V is taken with respect to the bipartition $(\mathrm{AA}', \mathrm{BB}')$ [33]. For

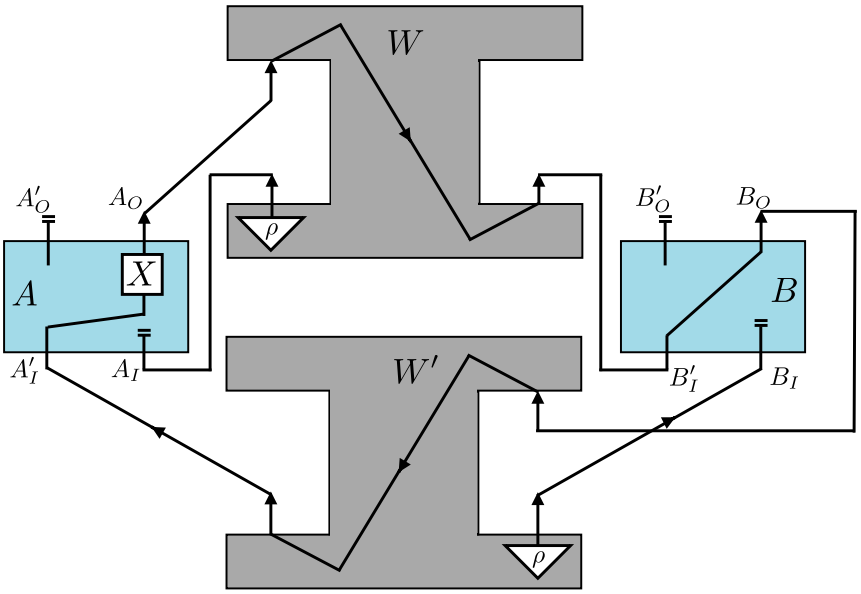


FIG. 3. The tensor product compostion rule $\mu(W, W') = W \otimes W'$ does not produce valid processes for all choices of W and W'. Here the process W corresponds to Alice recieving a state ρ, with an identity channel connecting her output system to Bob's input; W' is the same thing with the order of the parties reversed. The specific choice of local maps (X being the Pauli-X unitary gate) have zero probability under the "generalised Born rule" Eq. (2), instead of one, as it should be for deterministic operations.

instance, consider the process

$$W \otimes W, \text{ with } W = \frac{1}{2}(W^{A \preceq B} + W^{B \preceq A}) \tag{11}$$

then, it is sufficient to check directly the violation of Eq. (7) with respect to the bipartition (AA', BB'), namely, $W \otimes W \neq {}_{B_O B_O'}(W \otimes W) + {}_{A_O A_O'}(W \otimes W) - {}_{A_O A_O' B_O B_O'}(W \otimes W)$. This problem is illustrated in Figure 3, where two processes W, W' corresponding to channels in different directions can be seen to lead to a "loop", and to unnormalised probabilites. It is then natural to ask whether the tensor product can be replaced with another composition rule.

One may, however, argue that it is in principle possible to define more general composition rules that take this problem into account. For instance, one could take the tensor product and then "project" back the corresponding operator onto the space of valid process, or one could first decompose the process into a linear combination of processes in a definite order, then take the tensor product of each term and then recombine them. There are infinitely many possible recipes to define a composition rule; an abstract prescription for general composition rules is provided in Refs. [37, 38]. In the following, we will ask three reasonable and physically motivated requirements and show that there is no way of satisfying all three.

To define our composition rule μ, we may ask the following minimal requirements:

R.1 $\mu(W_1, W_2)$ is a valid process w.r.t. the bipartition (AA', BB'), for $W_1 \in \mathcal{W}, W_2 \in \mathcal{W}'$ (Validity).

R.2 $\mu(W_1, W_2) = W_1 \otimes W_2$ if $W_1 \in \mathcal{W}, W_2 \in \mathcal{W}'$, and W_1, W_2 are in the same order, i.e., $({}_{A_O} W_1 = W_1$ and ${}_{A_O'} W_2 = W_2)$, or $({}_{B_O} W_1 = W_1$ and ${}_{B_O'} W_2 = W_2)$ (Consistency).

R.3 $\mu(W_1, W_2)$ is convex linear in both arguments (Convex Linearity);

Requirement R.1 is needed for the composition of two processes to still belong to a bipartite scenario, i.e., where Alice has access to both systems AA', and Bob to BB'. R.2 is a consistency condition, i.e., the case of definite order should coincide with standard quantum theory. R.3 can be derived by requiring that our composition is well-behaved with respect to statistical mixtures, i.e., classical randomness, as explained in Appendix A.

It will be interesting to first consider a weaker assumption than R.1, because it will help us to single out the usual mathematical tensor product as a composition rule:

R'.1 $\mu(W_1, W_2) \geq 0$ for $W_1 \in \mathcal{W}, W_2 \in \mathcal{W}'$ (Positivity);

Assume that μ is a composition rule satisfying R'.1 (or R.1), R.2, R.3. Then there is a *unique* real-linear extension μ^L that satisfies $\mu^L(W_1, W_2) = \mu(W_1, W_2)$, for all $W_1 \in \mathcal{W}, W_2 \in \mathcal{W}'$. By construction this extension satisfies:

R'.3 $\mu(W_1, W_2)$ is real linear in both arguments (Linearity);

For the linear extension, we only demand R.1' (or R.1) for process matrices as inputs, so it will trivially continue to be satisfied. As R.2 itself is a (bi)linear condition, the linear extension will satisfy it even when it is extended to the linear span of process matrices:

R'.2 $\mu(W_1, W_2) = W_1 \otimes W_2$ if $W_1 \in L_V(\mathcal{L}(\mathcal{H}_{AB})), W_2 \in L_V(\mathcal{L}(\mathcal{H}_{A'B'}))$, and W_1, W_2 satisfy (${}_{A_O}W_1 = W_1$ and ${}_{A_O'}W_2 = W_2$), or (${}_{B_O}W_1 = W_1$ and ${}_{B_O'}W_2 = W_2$) (Consistency)

Details can be found in Appendix A.

With our axioms, we will be able to prove

Theorem 1. *The only function satisfying R'.1, R'.2, R'.3 is* $\mu(W_1, W_2) := W_1 \otimes W_2$.

Theorem 1 can be applied to the linear extension μ^L, implying that $\mu(W_1, W_2) = W_1 \otimes W_2$, and from that it will follow

Theorem 2. *There exists no function satisfying R.1-R.3.*

In particular, Th. 1 will imply that for the multipartite case the choice of the composition rule is unique. We will prove Th. 1 for the simple case of local systems consisting of n-qubits, i.e., with local dimension 2^n for each one of $A_I, A_I', \ldots, B_O, B_O'$, the general proof can be found in Appendix B. Given Th. 1, for the proof of Th. 2 it is sufficient to use the result of Ref. [33], or the example in Eq. (11).

First, we need the following

Lemma 1. *Given* A_1, A_2 *Hermitian operators such that* $A_1 \in L_V(\mathcal{L}(\mathcal{H}_{AB}))$ *and* $A_2 \in L_V(\mathcal{L}(\mathcal{H}_{A'B'}))$, *and let* μ *be a composition rule satisfying R'.1-3. Then* $\mu(A_1, A_2) = \mu(A_1, A_2)^\dagger$ *and* $\|\mu(A_1, A_2)\| \leq \|A_1 \otimes A_2\|$.

Proof.–For A Hermitian, its norm can be written as: $\|A\| = \min\{\lambda \mid -\lambda \mathbb{1} \leq A \leq \lambda \mathbb{1}\}$. Consider $A_1 \in L_V(\mathcal{L}(\mathcal{H}_{AB}))$ and $A_2 \in L_V(\mathcal{L}(\mathcal{H}_{A'B'}))$ Hermitian and with $\lambda_i = \|A_i\|$ for $i = 1, 2$. We define

$$W_1^{\pm} = \lambda_1 \mathbb{1} \pm A_1, \quad W_2^{\pm} = \lambda_2 \mathbb{1} \pm A_2, \tag{12}$$

which are valid processes, up to a normalization factor, on the spaces AB and A'B'. We then have,

$$\begin{aligned} 0 \leq \frac{\mu(W_1^+, W_2^+) + \mu(W_1^-, W_2^-)}{2} &= \lambda_1 \lambda_2 \mathbb{1} + \mu(A_1, A_2), \\ 0 \leq \frac{\mu(W_1^+, W_2^-) + \mu(W_1^-, W_2^+)}{2} &= \lambda_1 \lambda_2 \mathbb{1} - \mu(A_1, A_2), \end{aligned} \tag{13}$$

which implies $\mu(A_1, A_2) = \mu(A_1, A_2)^\dagger$ and $\|\mu(A_1, A_2)\| \leq \lambda_1 \lambda_2$. In the above, we used R'.1 for positivity, then R'.3 to split the different terms, and finally, R.2' to take the identity out of μ. □

For the following, we need to specialize the form of the operator A_1 and A_2. We define the set of tensor products of either traceless operators or the identity as

$$\mathrm{PTI}_{AB} := \{M = X^1_{A_I} \otimes X^2_{A_O} \otimes X^3_{B_I} \otimes X^4_{B_O} \mid M \in L_V(\mathcal{L}(\mathcal{H}_{AB})),\ X^i \text{ identity or traceless } \}, \tag{14}$$

and analogously for A'B'. For $M \in \mathrm{PTI}_{AB}$, an operator of the form $\mathbb{1} + M$ is, up to normalization, a causally ordered process. With the above definition, we prove the following

Lemma 2. *Let* μ *be a composition rule satisfying R'.1-3, and let* $M \in \mathrm{PTI}_{AB}$ *and* $N \in \mathrm{PTI}_{A'B'}$ *be Hermitian operators with eigenvalues in the interval* $[-1, 1]$. *Given an eigenvector* $|k\rangle$ *of* M *with eigenvalue* $(-1)^k$ *and an eigenvector* $|j\rangle$ *of* N *with eigenvalue* $(-1)^j$, *we have*

$$\mu(M, N)|k, j\rangle = (-1)^{k+j}|k, j\rangle \tag{15}$$

Proof.—To prove the lemma, it is sufficient to consider the (unnormalized) processes $W_1^k := \mathbb{1} + (-1)^{k+1}M$ and $W_2^j := \mathbb{1} + (-1)^{j+1}N$. By R'.2, $\mu(\mathbb{1},\mathbb{1}) = \mathbb{1}\otimes\mathbb{1}$ and $\mu(M,\mathbb{1}) = M\otimes\mathbb{1}$, since for $M \in \mathrm{PTI}_{\mathrm{AB}}$, either ${}_{\mathrm{A_O}}M = M$ or ${}_{\mathrm{B_O}}M = M$. Then,

$$\mu(W_1^k, W_2^j) = \mathbb{1} + (-1)^{k+1}M\otimes\mathbb{1} + (-1)^{j+1}\mathbb{1}\otimes N + (-1)^{k+j}\mu(M,N). \tag{16}$$

by R'.2 and R.3', and finally, by R'.1,

$$0 \leq \langle k,j|\mu(W_1^k, W_2^j)|k,j\rangle = 1 - 1 - 1 + (-1)^{j+k}\langle k,j|\mu(M,N)|k,j\rangle, \tag{17}$$

which implies $\mu(M,N)|k,j\rangle = (-1)^{j+k}|k,j\rangle$, since $\|\mu(M,N)\| \leq 1$, by Lemma 1. □

A straightforward corollary of Lemma 2 is that $\mu(M,N) = M\otimes N$ whenever M, N have eigenvalues only in $\{-1,1\}$. By linearity, this is enough to prove Th. 1 for all processes defined on n-qubit systems (i.e., local dimension 2^n) since we have a basis of operators, given by tensor products of Pauli matrices and the identity, that satisfy the assumptions. The same reasoning can be extended to arbitrary dimensions, see the details in Appendix B.

IV. DISCUSSION AND CONCLUSIONS

In this letter, we considered the parallel composition of process matrices. As the tensor product is known to lead to invalid process matrices, we investigated whether there is another map that can describe this parallel composition. We only asked for three weak desiderata: First of all, in contrast to the usual tensor product, it should always result in a valid process matrix. Furthermore, it should reduce to the familiar tensor product in the case of definite causal order. At last, we demanded bilinearity for compatibility with the interpretation of convex mixtures as statistical mixtures. However we have seen that even those reasonable desiderata are incompatible with each other.

Our results imply that an information theory of general quantum processes cannot rely on the assumption that multiple independent processes can be shared between two (or more) parties. In information theory, it is typical to assume that many independent samples of a random source, many independent uses of a channel, etc. are available, and that agents can perform global operations on many independent copies of the resource; this will not be possible in an information theory of general quantum processes. Rather, these results suggest that the proper setting for defining information-theoretic quantities such as entropies, capacities, etc., for process matrices is *single-shot* information theory [39–41].

One can infer from the main proof that even the case of two channels with opposing signalling direction will lead to a contradiction, which is perhaps unsurprising in the usual case of quantum mechanics on a fixed background spacetime. Indeed, suppose that an event A is in the causal past of an event B, and that A' is in the causal future of B'. Our desiderata that A and A' correspond to the same party can be interpreted as requiring that the events A, A' occur at the same space-time point p. This could be the case, but then B must be in the future light-cone of p, while B' must be in it's past light-cone. It is thus impossible to satisfy the requirement that B and B' also occur at the same spacetime point.

Therefore any composition rule for process matrices must take care of removing the two-way signalling terms, whose impossibility has a clear interpretation as discussed above. We have shown that there is no linear way of doing so, if we ask for that our composition rule reduces to the usual tensor product in the case of two processes with the same definite causal order.

However, there might exist reasonable non-linear composition rules, in the cases where processes have a concrete physical interpretation. A meaningful way to define an event for the composite party AA' is by the "simultaneous" entering of both systems $\mathcal{H}_A$ and $\mathcal{H}_{A'}$ in a localised laboratory, and similarly for BB'. There can be a probability that the systems do not enter the laboratories simultaneously, in which case it is necessary to post-select on the runs of the experiment where this was indeed the case. Since the post-selection probability depends on the two processes that we wish to compose, the map will be non-linear. An important issue with such a post-selected composition map for information-theoretic applications is that the parallel composition of resources is usually a "free operation", while in the post-selected case it would have a probability of failure.

V. ACKNOWLEDGEMENTS

We thank Ilya Kull, Ding Jia, Fabio Costa, Paolo Perinotti, and Lucien Hardy for useful discussions and comments. We acknowledge the support of the Austrian Science Fund (FWF) through the Doctoral Programme CoQuS, the projects I-2562-N27 and I-2906, the project M 2107 (Meitner-Programm) and the research platform TURIS, as well as support from the European Commission via Testing the Large-Scale Limit of Quantum Mechanics (TEQ) (No. 766900) project. P.A.G. acknowledges support from the Fonds de Recherche du Québec – Nature et Technologies (FRQNT). This publication was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

Appendix A: Linearity and convex linearity

In this appendix, we discuss convex-linearity and the linear extension of convex maps. First, let us argue why convex-linearity is a reasonable physical assumption. In operational approaches to physical theories [42, 43], one studies the probabilities that can be obtained from an abstract set of preparations and measurements. Given two preparations α, and β, there exists an another preparation γ that consists of preparing α with classical probability p, and preparing β with probability $(1-p)$. The probability for any measurement on γ is the weighted sum of the probabilities associated with preparations α and β. If we associate "states" with preparations, this means that the state space is convex linear. For example, the density matrix formalism can be seen to arise by adding classical uncertainty to the pure state formalism (i.e. kets in a Hilbert space). If one knows that with probability p_j, one prepares $|j\rangle$, then the density matrix is given by $\rho = \sum_j p_j |j\rangle\langle j|$. Another motivation for allowing arbitrary probabilistic mixtures appears in Refs. [44, 45], where it is shown that it implies that optimal compression is equivalent to linear compression.

The same interpretation can be used for process matrices: if the process matrices W_j are prepared with probabilities p_j, then all expectations values (and by that all statistics) can be calculated with the effective process matrix $W = \sum_j p_j W_j$. This can be seen by noting that $p(a,b) = \mathrm{Tr}[W M_a^{(A)} \otimes M_b^{(B)}]$ is a linear function in W and applying the law of total probability.

Consistency demands that the composition rule μ remains compatible with this interpretation of convex mixtures: If the first process is W_j with probability p_j and the second process is W'_k with probability p'_k, then the effective process matrices determining the statistics are $W = \sum_j p_j W_j$ and $W' = \sum_k p'_k W'_k$. The resulting combined process would be $\mu\Big(\sum_j p_j W_j, \sum_k p'_k W'_k\Big)$. However, a different point of view would be to say: With probabilities p_j and p'_k we combined the processes W_j and W'_k to $\mu(W_j, W'_k)$. So we prepared $\mu(W_j, W'_k)$ with probability $p_j p'_k$. Now, the effective process matrix is described by $\sum_{jk} p_j p'_k \mu(W_j, W'_k)$. As both points of view describe the same operational scenario, they have to be consistent:

$$\mu\Big(\sum_j p_j W_j, \sum_k p'_k W'_k\Big) = \sum_{jk} p_j p'_k \mu(W_j, W'_k). \tag{A1}$$

Next, we explain in further detail how to extend a function satisfying R.1 (or R′.1), R.2 and R.3 to a function satisfying R.1 (or R′.1), R′.2 and R′.3 on the linear span of all the process matrices.

Constructing the (bi)linear extension itself is a standard procedure in quantum information theory and is explained e.g. in Refs. [42, 43] for general abstract state spaces. Let S_1, S_2 be two convex sets, and let $f : S_1 \to S_2$ be a convex linear map. Let V_1, V_2 be the real vector spaces obtained respectively from S_1, S_2 by taking their linear span. Then f can be extended in the obvious way to a linear function $f^L : V_1 \to V_2$, defined by $f^L(\lambda a + b) = \lambda f(a) + f(b)$, for all $a, b \in S_1$, $\lambda \in \mathbb{R}$.

However, we still need to check that the bilinear extension still satisfies our postulates: We do not change R.1 (or R′.1), i.e. we only demand the output to be a process matrix (or positive) if the inputs are process matrices. Therefore R.1 (or R′.1) trivially continues to hold as the extension does not change the function on inputs that are process matrices.

Less trivial is how to generalize R.2. We will explicitly show that it still holds for the cases we need. Let us assume we have operators $M_1 \in L_V(\mathcal{L}(\mathcal{H}_{AB}))$ and $M_2 \in L_V(\mathcal{L}(\mathcal{H}_{A'B'}))$ with ${}_{A_O}M_1 = M_1$ and ${}_{A'_O}M_2 = M_2$ (or

alternatively ${}_{B_O}M_1 = M_1$ and ${}_{B'_O}M_2 = M_2$). We now show that

$$\mu^L(M_1, M_2) = M_1 \otimes M_2. \tag{A2}$$

By definition, M_1 and M_2 are allowed terms satisfying the projective condition (8). Therefore there exist λ_1, λ_2 such that $\frac{\mathbb{1}}{d_I} + \lambda_1 M_1$ and $\frac{\mathbb{1}}{d_{I'}} + \lambda_2 M_2$ are valid process matrices. Similarly $\frac{\mathbb{1}}{d_I d_{I'}}$ itself is a valid process matrix, with no signaling at all. Using R.2 for the original μ on valid process matrices, we find for the linear extension:

$$\begin{aligned}\mu^L\left(\frac{\mathbb{1}}{d_I} + \lambda_1 M_1, \frac{\mathbb{1}}{d_{I'}}\right) =& \mu\left(\frac{\mathbb{1}}{d_I} + \lambda_1 M_1 \frac{\mathbb{1}}{d_{I'}}\right) = \left(\frac{\mathbb{1}}{d_I} + \lambda_1 M_1\right) \otimes \frac{\mathbb{1}}{d_{I'}} = \frac{1}{d_I d_{I'}} \mathbb{1} \otimes \mathbb{1} + \frac{\lambda_1}{d_{I'}} M_1 \otimes \mathbb{1} \\ =& \mu\left(\frac{\mathbb{1}}{d_I}, \frac{\mathbb{1}}{d_{I'}}\right) + \frac{\lambda_1}{d_{I'}} M_1 \otimes \mathbb{1} = \mu^L\left(\frac{\mathbb{1}}{d_I}, \frac{\mathbb{1}}{d_{I'}}\right) + \frac{\lambda_1}{d_{I'}} M_1 \otimes \mathbb{1}.\end{aligned}$$

Therefore by bilinearity we find $\mu^L(M_1, \mathbb{1}) = M_1 \otimes \mathbb{1}$ and similarly $\mu^L(\mathbb{1}, M_2) = \mathbb{1} \otimes M_2$. Similarly, applying R.2 to the process matrices $\frac{\mathbb{1}}{d_I} + \lambda_1 M_1$ and $\frac{\mathbb{1}}{d_{I'}} + \lambda_2 M_2$, which have the same signaling direction, we obtain

$$\mu^L\left(\frac{\mathbb{1}}{d_I} + \lambda_1 M_1, \frac{\mathbb{1}}{d_{I'}} + \lambda_2 M_2\right) = \mu\left(\frac{\mathbb{1}}{d_I} + \lambda_1 M_1, \frac{\mathbb{1}}{d_{I'}} + \lambda_2 M_2\right) = \left(\frac{\mathbb{1}}{d_I} + \lambda_1 M_1\right) \otimes \left(\frac{\mathbb{1}}{d_{I'}} + \lambda_2 M_2\right). \tag{A3}$$

Collecting our results and using bilinearity on the left hand side of Eq. (A3) above, we finally see that R'.2 is satisfied:

$$\mu^L(M_1, M_2) = M_1 \otimes M_2. \tag{A4}$$

Appendix B: Proof of Th.1 in arbitrary dimension

In this appendix, we will extend the proof of Th. 1 to the case of arbitrary dimension. We start with the following **Lemma 3.** *Let $M \in \mathrm{PTI}_{\mathrm{AB}}$ and $N \in \mathrm{PTI}_{\mathrm{A'B'}}$ be Hermitian operators such that $|k\rangle = |k_1\rangle_{\mathrm{A_I}} \otimes |k_2\rangle_{\mathrm{A_O}} \otimes |k_3\rangle_{\mathrm{B_I}} \otimes |k_4\rangle_{\mathrm{B_O}}$ is an eigenvector for M, with eigenvalues given, according to the above factorization, by the products $\lambda_k = \lambda_k^{(1)}\lambda_k^{(2)}\lambda_k^{(3)}\lambda_k^{(4)}$, with $\lambda_k^{(i)} \in \{-1, 0, 1\}$ $i = 1, 2, 3, 4$, and, similarly, $|j\rangle = |j_1\rangle_{\mathrm{A_I'}} \otimes |j_2\rangle_{\mathrm{A_O'}} \otimes |j_3\rangle_{\mathrm{B_I'}} \otimes |j_4\rangle_{\mathrm{B_O'}}$ is an eigenvector of N, with eigenvalue $\eta_j = \eta_j^{(1)}\eta_j^{(2)}\eta_j^{(3)}\eta_j^{(4)}$, with $\eta_j^{(i)} \in \{-1, 0, 1\}$ $i = 1, 2, 3, 4$. We then have*

$$\mu(M, N)|k, j\rangle = \lambda_k \eta_j |k, j\rangle. \tag{B1}$$

Proof.— The case $\lambda_k, \eta_j = \pm 1$ are included in Lemma 2. Let us consider the case $M|k\rangle = 0$ and $N|j\rangle \neq 0$, the case $M|k\rangle = N|j\rangle = 0$ can be obtained in a similar way, by applying the same argument first to M, then to N. Since M is in $\mathrm{PTI}_{\mathrm{AB}}$, we can write it as $M = X^1_{\mathrm{A_I}} \otimes X^2_{\mathrm{A_O}} \otimes X^3_{\mathrm{B_I}} \otimes X^4_{\mathrm{B_O}}$. Let us now further assume $X^1|k_1\rangle_{\mathrm{A_I}} = 0$, and $X^i|k_i\rangle_{\mathrm{Y}} \neq 0$ for $i = 2, 3, 4$, $\mathrm{Y} = \mathrm{A_O}, \mathrm{B_I}, \mathrm{B_O}$, in particular, this implies that $|k_i\rangle$ are eigenvectors for eigenvalues ± 1 for $i = 2, 3, 4$. We can then write:

$$X^1 = \left(X^1 + |k_1\rangle\langle k_1| - |(k+1)_1\rangle\langle(k+1)_1|\right) + \left(|(k+1)_1\rangle\langle(k+1)_1| - |k_1\rangle\langle k_1|\right) =: X'^1 + X''^1, \tag{B2}$$

where $|(k+1)_1\rangle$ is a vector orthogonal to $|k_1\rangle$. Then X'^1, X''^1 are both traceless and $X'^1|k_1\rangle = |k_1\rangle$, $X''^1|k_1\rangle = -|k_1\rangle$. We then have that $M' := X'^1_{\mathrm{A_I}} \otimes X^2_{\mathrm{A_O}} \otimes X^3_{\mathrm{B_I}} \otimes X^4_{\mathrm{B_O}}$ and $M'' := X''^1_{\mathrm{A_I}} \otimes X^2_{\mathrm{A_O}} \otimes X^3_{\mathrm{B_I}} \otimes X^4_{\mathrm{B_O}}$ are again in $\mathrm{PTI}_{\mathrm{AB}}$. Thus, by Lemma 2,

$$\mu(M, N)|k, j\rangle = \mu(M' + M'', N)|k, j\rangle = \mu(M', N)|k, j\rangle + \mu(M'', N)|k, j\rangle = M' \otimes N|k, j\rangle + M'' \otimes N|k, j\rangle = 0. \tag{B3}$$

If another operator, say X^2, is zero on the corresponding eigenvector, say $|k_2\rangle_{\mathrm{A_O}}$, we can again repeat the construction in Eq. (B2) to construct X'^2, X''^2 with $+1, -1$ eigenvalues and use again linearity and Lemma 2. Similarly, the same argument can be extended to all X^i and to N. □

To conclude the proof of Th. 1, it is sufficient to construct a basis of operators containing the identity and where each elements, except the identity, is traceless and with eigenvalues in $\{-1, 0, 1\}$. Let $\mathcal{H}$ be a Hilbert space with

dimension d, and let $\{|k\rangle\}_{k=1}^{d}$ be a basis for $\mathcal{H}$. The space of Hermitian operators on $\mathcal{H}$ is a real vector space of dimension d^2. We define the following operators

$$Z_i = |i\rangle\langle i| - |i+1\rangle\langle i+1|, \quad 1 \leq i \leq d-1 \tag{B4}$$
$$X_{jk} = |j\rangle\langle k| + |k\rangle\langle j|, \quad 1 \leq j < k \leq d \tag{B5}$$
$$Y_{jk} = i(|j\rangle\langle k| - |k\rangle\langle j|), \quad 1 \leq j < k \leq d, \tag{B6}$$

which are traceless, hermitian and with eigenvalues in $\{-1, 0, 1\}$. The X_{jk} and Y_{jk} are also known as part of an operator basis called *Generalized Gell-Mann matrices* [46]. For completeness we now show that, together with $\mathbb{1}$, the above set of matrices form a basis for the space of Hermitian operators on $\mathcal{H}$. It is clear that the $\{X_{jk}\}$ and $\{Y_{jk}\}$ span the space of Hermitian operators whose diagonal is zero in the $|k\rangle$ basis. All that remains to be shown is that $\{\mathbb{1}, Z_i\}$ forms a basis for the space of diagonal real matrices, which we prove by expressing the basis $\{|k\rangle\langle k|\}$ in terms of the new basis $\{\mathbb{1}, Z_i\}$.
Notice that for $1 \leq i \leq d-1$,

$$\sum_{j=i}^{j=d-1} Z_j = |i\rangle\langle i| - |d\rangle\langle d|, \tag{B7}$$

and also that

$$\sum_{j=1}^{d-1} jZ_j = \sum_{j=1}^{d-1} |j\rangle\langle j| - (d-1)|d\rangle\langle d| = \mathbb{1} - d|d\rangle\langle d|. \tag{B8}$$

Combining the above two expressions gives

$$|d\rangle\langle d| = \frac{1}{d}\mathbb{1} - \frac{1}{d}\sum_{j=1}^{d-1} jZ_j, \tag{B9}$$

$$|i\rangle\langle i| = \frac{1}{d}\mathbb{1} + \sum_{j=i}^{d-1} Z_j - \frac{1}{d}\sum_{j=1}^{d-1} jZ_j, \quad 1 \leq i < d, \tag{B10}$$

which concludes that $\{\mathbb{1}, Z_i, X_{jk}, Y_{jk}\}$ is a basis for the space of Hermitian operators of $\mathcal{H}$.
We can use the above construction to build a basis for $\mathcal{L}(\mathcal{H}_{A_I} \otimes \mathcal{H}_{A_O} \otimes \mathcal{H}_{B_I} \otimes \mathcal{H}_{B_O})$ consisting of tensor products of local Hermitian operators whose eigenvalues are in $\{-1, 0, 1\}$. We then remove from this basis all the terms that do not satisfy the linear constraints L_V. This gives us a basis for the linear space of valid Ws, which is contained in $\mathrm{PTI}_{\mathrm{AB}}$. We will call this basis simply $\{M_i\}_{i\in I}$, and by Lemma 3, we have

$$\mu(M_i, M_j) = M_i \otimes M_j. \tag{B11}$$

We can then decompose any pair W, W' as

$$W = \mathbb{1} + \sum_i c_i M_i, \quad W' = \mathbb{1} + \sum_i d_i M_i, \tag{B12}$$

and apply μ, namely

$$\begin{aligned} \mu(W, W') &= \mathbb{1} + \sum_i c_i M_i \otimes \mathbb{1} + \mathbb{1} \otimes \sum_i d_i M_i + \sum_{ij} c_i d_j \mu(M_i, M_j) \\ &= \mathbb{1} + \sum_i c_i M_i \otimes \mathbb{1} + \mathbb{1} \otimes \sum_i d_i M_i + \sum_{ij} c_i d_j M_i \otimes M_j = W \otimes W', \end{aligned} \tag{B13}$$

which concludes the proof of Th. 1.

[1] J. S. Bell, "On the Einstein–Podolsky–Rosen Paradox," *Physics* **1**, 195 (1964).

[2] C. J. Wood and R. W. Spekkens, "The lesson of causal discovery algorithms for quantum correlations: causal explanations of Bell-inequality violations require fine-tuning," *New J. Phys.* **17**, 033002 (2015), arXiv:1208.4119 [quant-ph].
[3] L. Hardy, "Probability Theories with Dynamic Causal Structure: A New Framework for Quantum Gravity," arXiv:gr-qc/0509120 [gr-qc].
[4] L. Hardy, "Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure," *J. Phys. A: Math. Theor.* **40**, 3081 (2007), arXiv:gr-qc/0608043 [gr-qc].
[5] L. Hardy, "Quantum Gravity Computers: On the Theory of Computation with Indefinite Causal Structure," in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony*, pp. 379–401. Springer Netherlands, Dordrecht, 2009. arXiv:quant-ph/0701019 [quant-ph].
[6] O. Oreshkov, F. Costa, and Č. Brukner, "Quantum correlations with no causal order," *Nat. Commun.* **3**, 1092 (2012), arXiv:1105.4464 [quant-ph].
[7] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, "Quantum computations without definite causal structure," *Phys. Rev. A* **88**, 022318 (2013), arXiv:0912.0195 [quant-ph].
[8] R. Oeckl, "A local and operational framework for the foundations of physics," arXiv:1610.09052 [quant-ph].
[9] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, "Witnessing causal nonseparability," *New J. Phys.* **17**, 102001 (2015), arXiv:1506.03776 [quant-ph].
[10] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, "Experimental superposition of orders of quantum gates," *Nat. Commun.* **6**, 7913 (2015), arXiv:1412.4006 [quant-ph].
[11] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, "Experimental verification of an indefinite causal order," *Sci. Adv.* **3**, (2017), arXiv:1608.01683 [quant-ph].
[12] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, "Indefinite Causal Order in a Quantum Switch," arXiv:1803.04302 [quant-ph].
[13] O. Oreshkov, "On the whereabouts of the local operations in physical realizations of quantum processes with indefinite causal order," arXiv:1801.07594.
[14] O. Oreshkov and C. Giarmatzi, "Causal and causally separable processes," *New Journal of Physics* **18**, 093020 (2016), arXiv:1506.05449 [quant-ph].
[15] C. Branciard, M. Araújo, A. Feix, F. Costa, and Č. Brukner, "The simplest causal inequalities and their violation," *New J. Phys.* **18**, 013008 (2016), arXiv:1508.01704 [quant-ph].
[16] A. A. Abbott, C. Giarmatzi, F. Costa, and C. Branciard, "Multipartite causal correlations: Polytopes and inequalities," *Phys. Rev. A* **94**, 032131 (2016), arXiv:1608.01528 [quant-ph].
[17] N. Miklin, A. A. Abbott, C. Branciard, R. Chaves, and C. Budroni, "The entropic approach to causal correlations," *New J. Phys.* **19**, 113041 (2017), arXiv:1608.01528 [quant-ph].
[18] M. Araújo, F. Costa, and Č. Brukner, "Computational Advantage from Quantum-Controlled Ordering of Gates," *Phys. Rev. Lett.* **113**, 250402 (2014).
[19] A. Feix, M. Araújo, and Č. Brukner, "Quantum superposition of the order of parties as a communication resource," *Phys. Rev. A* **92**, 052326 (2015), arXiv:1508.07840.
[20] P. A. Guérin, A. Feix, M. Araújo, and Č. Brukner, "Exponential Communication Complexity Advantage from Quantum Superposition of the Direction of Communication," *Phys. Rev. Lett.* **117**, 100502 (2016), arxiv:1605.07372 [quant-ph].
[21] Ä. Baumeler and S. Wolf, "Non-Causal Computation," *Entropy* **19**, (2017), arXiv:1601.06522.
[22] M. Araújo, P. A. Guérin, and A. Baumeler, "Quantum computation with indefinite causal structures," *Phys. Rev. A* **96**, 052315 (2017), arXiv:1706.09854.
[23] Ä. Baumeler and S. Wolf, "Computational tameness of classical non-causal models," *Proc. R. Soc. A* **474**, (2018), arXiv:1611.05641.
[24] D. Ebler, S. Salek, and G. Chiribella, "Enhanced Communication with the Assistance of Indefinite Causal Order," *Phys. Rev. Lett.* **120**, 120502 (2018), arxiv:1711.10165 [quant-ph].
[25] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, New York, NY, USA, 10th ed., 2011.
[26] M. M. Wilde, *Quantum Information Theory*. Cambridge University Press, 2 ed., 2017.
[27] B. Schumacher, "Quantum coding," *Phys. Rev. A* **51**, 2738–2747 (1995).
[28] M. Zych, *Quantum Systems under Gravitational Time Dilation*. Springer Theses. Springer International Publishing, 2017.
[29] A. Feix and Č. Brukner, "Quantum superpositions of common-cause and direct-cause causal structures," *New J. Phys.* **19**, 123028 (2017), arXiv:1606.09241 [quant-ph].
[30] M. Zych, F. Costa, I. Pikovski, and C. Brukner, "Bell's Theorem for Temporal Order," arXiv:1708.00248 [quant-ph].
[31] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Theoretical framework for quantum networks," *Phys. Rev. A* **80**, 022339 (2009), arXiv:0904.4483 [quant-ph].
[32] F. Costa and S. Shrapnel, "Quantum causal modelling," *New J. Phys.* **18**, 063032 (2016), arXiv:1512.07106 [quant-ph].
[33] D. Jia and N. Sakharwade, "Tensor products of process matrices with indefinite causal structure," *Phys. Rev. A* **97**, 032110 (2018).
[34] A. Jamiołkowski, "Linear transformations which preserve trace and positive semidefiniteness of operators," *Rep. Math. Phys.* **3**, 275 – 278 (1972).

[35] M.-D. Choi, "Completely positive linear maps on complex matrices," *Linear Algebra Its Appl.* **10**, 285 – 290 (1975).
[36] Alternatively, one could define $\mathcal{M}_{a|x}$ with a global transposition ^{t}, taken w.r.t. the $\{|ij\rangle\}_{ij}$ basis, as in Ref. [9]. This allow one to write the process matrix associated to a quantum state ρ shared between the parties simply as $\rho_I \otimes \mathbb{1}_O$, instead of $\rho_I^t \otimes \mathbb{1}_O$. We will not use this convention here.
[37] P. Perinotti, "Causal structures and the classification of higher order quantum computations," in *Time in Physics (Tutorials, Schools, and Workshops in the Mathematical Sciences)*, R. Renner and S. Stupar, eds. Birkhäuser Basel, 1st ed., 2017. arXiv:1612.05099 [quant-ph].
[38] A. Bisio and P. Perinotti, "Axiomatic theory of Higher-Order Quantum Computation," arXiv:1806.09554 [quant-ph].
[39] R. Renner, S. Wolf, and J. Wullschleger, "The Single-Serving Channel Capacity," in *2006 IEEE International Symposium on Information Theory*. IEEE, 2006. arXiv:cs/0608018 [cs.IT].
[40] R. Konig, R. Renner, and C. Schaffner, "The Operational Meaning of Min- and Max-Entropy," *IEEE Transactions on Information Theory* **55**, 4337–4347 (2009), arXiv:0807.1338 [quant-ph].
[41] G. Chiribella and D. Ebler, "Optimal quantum networks and one-shot entropies," *New Journal of Physics* **18**, 093053 (2016), arXiv:1606.02394 [quant-ph].
[42] L. Hardy, "Quantum theory from five reasonable axioms," arxiv:quant-ph/0101012.
[43] J. Barrett, "Information processing in generalized probabilistic theories," *Phys. Rev. A* **75**, 032304 (2007), arxiv:quant-ph/0508211.
[44] L. Hardy, "Foliable Operational Structures for General Probabilistic Theories," arXiv:0912.4740 [quant-ph].
[45] L. Hardy, "Reformulating and Reconstructing Quantum Theory," arXiv:1104.2066 [quant-ph].
[46] R. A. Bertlmann and P. Krammer, "Bloch vectors for qudits," *J. Phys. A: Math. Theor.* **41**, 235303 (2008), arXiv:0806.1174 [quant-ph].

Chapter 4

Quantum reference frame transformations as symmetries and the paradox of the third particle

Authors: Marius Krumm, Philipp A. Höhn, Markus P. Müller
Submitted to *Quantum* for review (4th of Feb., 2021).
Preprint available as: arXiv:2011.01951
https://arxiv.org/abs/2011.01951

Contribution: I made contributions to all aspects of the research. Markus P. Müller initiated the project, I joined in the beginning, and Philipp Höhn joined after a few months. The paper developed as a continuous, dynamical interaction between Philipp Höhn, Markus P. Müller and me, providing conceptual considerations, mathematical lemmas, proofs, proof-reading and corrections for each other.

Abstract: In a quantum world, reference frames are ultimately quantum systems too — but what does it mean to "jump into the perspective of a quantum particle"? In this work, we show that quantum reference frame (QRF) transformations appear naturally as symmetries of simple physical systems. This allows us to rederive and generalize known QRF transformations within an alternative, operationally transparent framework, and to shed new light on their structure and interpretation. We give an explicit description of the observables that are measurable by agents constrained by such quantum symmetries, and apply our results to a puzzle known as the 'paradox of the third particle'. We argue that it can be reduced to the question of how to relationally embed fewer into more particles, and give a thorough physical and algebraic analysis of this question. This leads us to a generalization of the partial trace ('relational trace') which arguably resolves the paradox, and it uncovers important structures of constraint quantization within a simple quantum information setting, such as relational observables which are key in this resolution. While we restrict our attention to finite Abelian groups for transparency and mathematical rigor, the intuitive physical appeal of our results makes us expect that they remain valid in more general situations.

I. INTRODUCTION

All physical quantities are described relative to some frame of reference. But since all physical systems are fundamentally quantum, reference frames must ultimately be quantum systems, too. This simple insight is of fundamental importance in a variety of physical fields, including quantum information theory [1–8], quantum thermodynamics [9–17], quantum gravity [18–26], and in the foundations of quantum physics [27–37] more generally.

Recently, there has been a wave of interest in a specific approach to quantum reference frames (QRFs) that we can broadly classify as *structural* in nature, including e.g. Refs. [38–48]. This approach extends the usual concept of reference frames by associating them with quantum systems, and by describing the physical situation of interest from the "internal perspective" of that quantum system. For example, if an interferometer has a particle travelling in a superposition of paths, how "does the particle see the interferometer" [49]?

A central topic in this approach is the QRF dependence of observable properties like superposition, entanglement [38–40], classicality [39, 71, 72], or of quantum resources [73]. The corresponding QRF transformations admit an unambiguous definition of spin in relativistic settings by transforming to a particle's rest frame [46, 47], they describe the comparison of quantum clock readings [42, 45], and they yield an alternative approach to indefinite causal structure [48, 76]. Among other conceived applications, they are furthermore conjectured to play a crucial rule in the implementation of a "quantum equivalence principle" [75] as well as in spacetime singularity resolution [57] and the description of early universe power spectra [58, 59] in quantum gravity and cosmology.

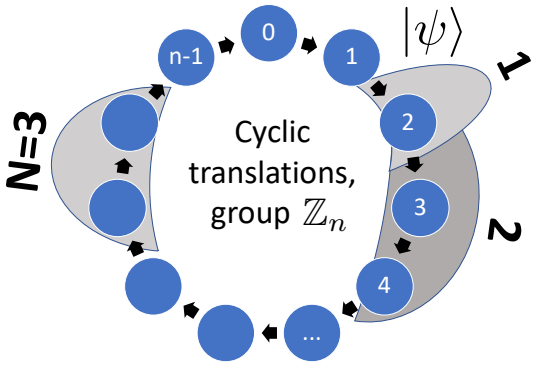


FIG. 1. The simplest example of this article's setup: a discretization of wave functions in one spatial dimension under translation symmetry. The configuration space is the cyclic group $\mathbb{Z}_n$, and the one-particle Hilbert space is $\mathcal{H} = \ell^2(\mathbb{Z}_n) \simeq \mathbb{C}^n$. We have N distinguishable particles in a joint quantum state $|\psi\rangle \in \mathcal{H}^{\otimes N}$, and we study QRF transformations that switch between the "perspectives of the particles".

Despite the broad appeal, several fundamental and conceptual questions remain open. For example, how should we make concrete sense of the idea of "jumping into the reference frame of a particle"? How are QRF transformations different from any other unitary change of basis in Hilbert space? What kind of physical symmetry claim is associated with the intuition that QRF changes "leave the physics invariant"? Furthermore, there are reported difficulties to extend basic quantum information concepts into this context. For example, Ref. [49] describes a 'paradox of the third particle', an apparent inconsistency arising from determining reduced quantum states in different QRFs.

In this article, we shed considerable light on all of these questions. We introduce a class of physical systems subject to simple principles, and derive the QRF transformations as the physical symmetries of these systems. On the one hand, this gives us a transparent operational framework for QRFs that makes sense of the 'jumping' metaphor. On the other hand, it allows us to identify QRF transformations as elements of a natural symmetry group, and to describe the structure of the observables that are invariant under such transformations. This algebraic structure turns out to be key to elucidate the paradox of the third particle, which we do by introducing a relational notion of the partial trace.

To keep the mathematical structures as transparent and accessible as possible, we restrict our attention in this article to finite Abelian groups. But this already includes interesting physical settings like the discretization of translation-invariant quantum particles on the real line (see Figure 1), admitting the formulation of intriguing thought experiments. Within this familiar quantum information regime of finite-dimensional Hilbert spaces, we uncover a variety of structures that not only shed light on the questions raised above, but that also reflect important aspects of constraint quantization [60, 61], which for example underlies canonical approaches to quantum gravity and cosmology. This includes the notions of a "physical Hilbert space" encoding the relational states of the theory [19, 66, 67], of relational and Dirac observables [18–26, 42–45], and a simple demonstration of how constraints can in general arise from symmetries. In particular, these notions will assume key roles in our proposed resolution of the paradox of the third particle.

Overview and summary of results. Our article is organized as follows. In Section II, we begin with a thorough operational comparison of this structural approach to QRFs with the more common quantum information approach. This sets the stage by embedding the notion of QRF transformations into a broader conceptual framework.

In Section III, we specialize to a concrete class of physical systems ("$\mathcal{G}$-systems") which hold a finite Abelian group as their classical configuration space. We prove the existence and elucidate the group structure of QRF

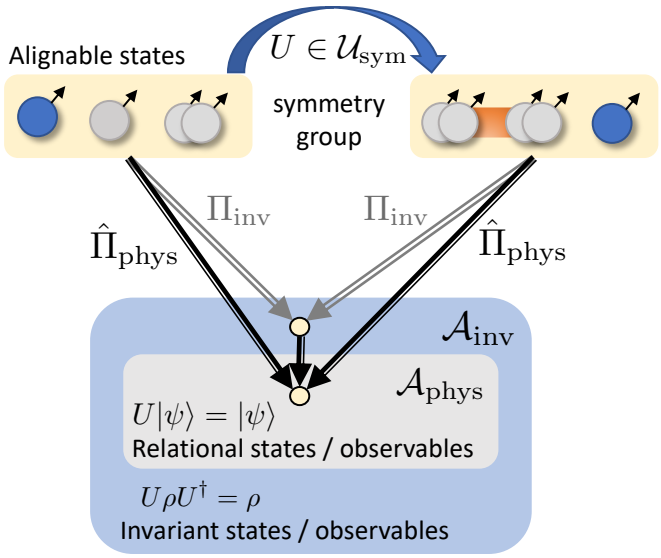


FIG. 2. Some of the structures we uncover in Section III. We axiomatically derive and analyze the quantum symmetry group $\mathcal{U}_{\rm sym}$, and characterize a class of "alignable states" that can be transformed into a form that is "relative to one of the particles". As described in Refs. [38–40], "jumping from the first to the third particle", for example (sketched on top), can transform separable into entangled states, owing to the fact that, as we will show, $\mathcal{U}_{\rm sym}$ is larger than the classical group of translations. We identify two subalgebras of operators that are invariant under all quantum symmetries, $\mathcal{A}_{\rm phys} \subset \mathcal{A}_{\rm inv}$, and corresponding projections that extract the "invariant part" of a state.

transformations for such systems, and introduce a notion of "alignable states" which are those that can be described "relative to one of the particles". We determine the invariant observables measurable by observers constrained by such symmetries. As sketched in Figure 2, we find that there are two important, but distinct notions of invariant observables, depending on whether symmetry transformations may induce superselection sector dependent phases or not. While the role of invariant observables in the structural approach has been stressed before [39, 41–45], attention was thus far restricted to their description on the space of invariant pure states ("physical Hilbert space"). Furthermore, we uncover important aspects of constraint quantization, and obtain representation-theoretic notions of physical concepts like the "total momentum" and its role as a constraint.

In Section IV, we apply our insights to the paradox of the third particle. We argue that the problem reduces to the physical question of when two groups of particles hold "the same relation" to each other within two distinct configurations, such that the corresponding branches should interfere (see Figure 6 on page 19). Mathematically, this corresponds to the question of how to embed the algebra of invariant N-particle observables into that of $N + M$ particles. We show that no unique answer to this question exists for the full set of invariant observables in $\mathcal{A}_{\rm inv}$: the answer always depends on the physical choice of how to determine the particle group interrelations operationally.

However, we show that a unique and natural embedding *does* exist for the subset of relational observables in $\mathcal{A}_{\rm phys}$. The trick is to use a *coherent superposition* of all operationally conceivable particle group relations, and it turns out this construction preserves the algebraic structure of the N-particle observables. We use this to define a relational notion of the partial trace which arguably resolves the paradox, and we compare this resolution to the one proposed by Angelo et al. [49] before concluding in Section V.

II. QUANTUM INFORMATION VS. STRUCTURAL APPROACH TO REFERENCE FRAMES

Let us begin with the main element that both the quantum information as well as the structural approach to QRFs have arguably in common: a physical system with a symmetry such that all observable quantities are invariant, or even fully relational. This is also the starting point of Refs. [2, 8, 33–36].

A. Describing physics with or without external relatum

Consider a physical system S of interest. We assume that there is a set $\mathcal{S}$ of *states* in which the system S can be prepared. Furthermore, there is a group of *symmetry transformations* $\mathcal{G}_{\rm sym}$ that acts on $\mathcal{S}$. Specifying $\mathcal{S}$ and $\mathcal{G}_{\rm sym}$ amounts to making a specific physical claim:

Assumption 1. If the system S is considered *in isolation*, then it is impossible to distinguish (even probabilistically) whether it has been prepared in some state ρ or in another state $G\rho$. This is true for all states $\rho \in \mathcal{S}$ and all symmetry transformations $G \in \mathcal{G}_{\rm sym}$.

'In isolation' here means that any other physical structure to which S could be related is disregarded, either because it does not exist in the first place, one does not have access to it, or it is deliberately ignored. This setting is schematically depicted in Figure 3. Examples include:

(i) Minkowski spacetime of special relativity, with $\mathcal{S}$ the set of all possible states of matter (say, of classical point particles), and the Poincaré group $\mathcal{G}_{\rm sym}$ as the group of symmetry transformations.

(ii) Electromagnetism in some bounded region of spacetime. This is a gauge theory with $\mathcal{G}_{\rm sym}$ the group of local $\mathrm{U}(1)$-transformations as its symmetry group.

(iii) A spin in quantum mechanics with Hilbert space $\mathcal{H}$ and projective representation $g \mapsto U_g$ of the rotation group $\mathcal{G} = \mathrm{SO}(3)$. Here, $\mathcal{G}_{\mathrm{sym}}$ consists of all maps of the form $\rho \mapsto U_g \rho U_g^\dagger$.

These three examples illustrate an important subtlety: to claim that ρ and $G\rho$ are physically indistinguishable, one needs to speak about ρ and $G\rho$ as different objects in the first place. In other words, one has to *somehow* define ρ and $G\rho$ as distinct states. But in order to do so, one would need something *external* to the system S to refer to.

FIG. 3. What both approaches have in common: a system S with a symmetry group $\mathcal{G}_{\mathrm{sym}}$ acting on its states $\rho \in \mathcal{S}$. States are implicitly defined via some (physical or fictional) external relatum, but *internally* (that is, for observers without access to the relatum) ρ and $G\rho$ are indistinguishable, for all $G \in \mathcal{G}_{\mathrm{sym}}$.

In example (i), there simply is no material external relatum, while in example (ii), it is given by electromagnetism *outside* of the bounded region. As emphasized in Ref. [50], while gauge symmetries do not change the physics of a given system, they alter the way that the system interacts with other systems. This observation is at the heart of the recent pivot to edge modes in gauge theory and gravity [51–56] and our resolution of the paradox of the third particle in Section IV can also be viewed in this light. In case (iii), the external relatum would be best described as an external classical reference frame, for example the laboratory of an agent experimenting with S. This illustrates that to consider a system "in isolation" in the sense of Assumption 1 does *not* imply that the system S is literally a physically isolated system. It simply means that we have chosen to describe the system without the external relatum relative to which the action of the symmetry group is defined. Moreover, the setting does *not* imply that the agent who treats ρ and $G\rho$ as indistinguishable is herself part of the system S, but only that the agent considers S without the external relatum.

Here we argue that the essential difference between the two approaches to quantum references frames can succinctly be stated as follows:

The **quantum information (QI) approach** as in e.g. Refs. [1–6] emphasizes the fact that quantum states are often only defined relative to an external relatum (as in Figure 3), and that this relatum may ultimately be a quantum system, too. This leads to questions like: how can quantum information-theoretic protocols be performed in the absence of a shared reference frame [1]? How well can quantum states stand in as resources of asymmetry if there is no shared frame [3, 4]? What are fundamental quantum limits for communicating or aligning reference frames [1]? Addressing questions as these often involves encoding information in quantum states in an external relatum independent manner and, as such, requires external relatum independent descriptions of states.

The **structural approach** as in e.g. Refs. [38–40] is not primarily concerned with operational protocols. While it shares the aim of external relatum independent descriptions of states with the QI approach, it goes further: it disregards the external relatum altogether, and instead asks whether and how *physical subsystems* of S can be promoted to an *internal* reference frame. This emphasizes the fact that the distinction between quantum systems and their reference frames is not fundamental, but merely conventional. It leads to questions like: what is the description of the quantum state relative to one of its particles? Can we find a Hilbert space basis in which the description of the physics is simplified, e.g., in which superpositions of subsystems of interest may be removed? More generally, what are the "QRF transformations" that relate the descriptions relative to different choices of internal reference frame?

In the QI approach, it is usually not necessary to take the extra step to internal frame choices and to ask how a system is described relative to one of its subsystems, as we will explain shortly. It suffices to focus on invariant properties of S which have a meaning relative to an arbitrary choice of external frame in order to successfully carry out communication tasks in the absence of a shared frame. It is also worth emphasizing that there does not exist a sharp distinction between the two approaches in the body of literature on QRFs. Since the structural approach shares external relatum independent state descriptions with the QI approach, there exist "hybrid" works which arguably incorporate elements from both. For example, Refs. [2, 7, 8, 33–36, 49] use standard quantum information techniques to define external relatum independent states, but also use the latter to explore to some degree the question of how a quantum state is described relative to a subsystem. However, these works do not study the relations between the different such descriptions and thus, in particular, do not study the QRF transformations.

The structural approach to QRFs is sometimes illustrated in ways that seem at first sight to be in conflict to the characterization above. For example, Figure 1 in Ref. [38] suggests to think of QRFs as physically attached to an observer and its laboratory (defined by its own quantum state), similarly as reference frames in Special Relativity are often thought of as being at-

tached to an observer (defined by its state of motion). QRF transformations would then relate the descriptions of "quantum" observers who are relative to each other in superposition in a Wigner's-friend-type fashion.

However, we will show below that the structural framework of QRFs can be derived and analyzed exactly under an alternative and simpler interpretation. As we will elaborate and generalize below, choosing a QRF amounts to *aligning one's description of the physics with respect to some choice of internal quantum subsystem*, such as the position of one of the particles. Two different observers can choose two different subsystems (say, particles) that are relative to each other in superposition, even if the observers themselves are fully classical. Their descriptions will then be related by QRF transformations. The observer who assigns the quantum state may thus retain the status of a *classical* entity external to the quantum system (at least in laboratory situations), as illustrated in Figure 3. While more conservative, this new interpretation is operationally more immediate, and it is sufficient to reconstruct and extend the full machinery of QRF transformations, as we will see.

The characterization above is also in line with another version of the structural approach: the so-called perspective-neutral approach [39, 41–45] which, motivated by quantum gravity, is formulated in the language of constrained Hamiltonian systems [60, 61]. The starting point of this approach is a deep physical and operational motivation: *take the idea seriously that there are no reference frames, such as rods or clocks, that are external to the universe*. To implement this idea, one starts with a "kinematical Hilbert space" that defines all the involved quantum degrees of freedom and some gauge symmetry, but is interpreted as purely auxiliary. The absence of external references is then implemented by restricting to the gauge-invariant subset of states where the description becomes purely relational.

The actual mathematical machinery applied in this approach still fits the description above: the kinematical Hilbert space can be viewed as being described relative to a *fictional* external relatum. The insight that there is nothing external to the universe motivates to ask — purely formally at first — whether some of the *internal* degrees of freedom of the theory can be promoted to a frame of reference, such as a rod or clock. One may finally ask whether observers who are part of the theory may in fact have good operational access to that chosen frame of reference, but this is an *additional* (though important) question that we here regard as secondary.

B. Communication scenarios illustrating the two approaches

Before we turn to the structural approach in detail, and relate the verbal description above to the mathematical formalization, let us elaborate on the distinction by means of two communication scenarios. To do so, let us informally introduce some piece of notation that we will later on define more formally. If $\rho \in \mathcal{S}$ is some state, denote by $[\rho]$ the set of all states that are symmetrically equivalent to ρ, i.e. $[\rho] := \{G\rho \mid G \in \mathcal{G}_{\text{sym}}\}$. The $[\rho]$ can be viewed as equivalence classes of states, or as orbits of the symmetry group.

Adapting the quantum information terminology from Ref. [1], we refer to physical properties of S that only depend on the equivalence class $[\rho]$ as *speakable information*. Being invariant under the action of $\mathcal{G}_{\text{sym}}$ and thus not requiring an external relatum in order to be defined, two agents can agree on the description of these properties by classical communication even in the absence of a shared frame. By contrast, we refer to physical properties of S that depend on the concrete representative ρ from an equivalence class $[\rho]$ of states as *unspeakable information*. These properties thus require the external relatum to be meaningful and cannot be communicated purely classically between two agents who do not share a frame.

1. *The quantum information approach: communicating quantum systems*

Consider the scenario in Figure 4. Alice holds a quantum system S that she has prepared in some state $\rho \in \mathcal{S}(\mathcal{H})$, and $\mathcal{S}(\mathcal{H})$ denotes the density matrices on the corresponding Hilbert space $\mathcal{H}$. We assume that there is a (for now, for simplicity) compact group $\mathcal{G}$ of symmetries and a projective representation $\mathcal{G} \ni g \mapsto U_g$ such that $\mathcal{G}$ acts on $\mathcal{S}$ via $\mathcal{U}_g(\rho) = U_g \rho U_g^\dagger$. In this case, the symmetry group is $\mathcal{G}_{\text{sym}} = \{\mathcal{U}_g \mid g \in \mathcal{G}\}$. If we assume that Alice's quantum system S has the properties of Assumption 1, then the very definition of ρ is relative to her local frame of reference.

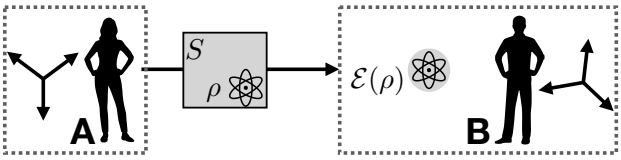


FIG. 4. A communication scenario within the **quantum information approach** as in Ref. [1]. The focus is on sending and recovering actual physical (quantum) states that are defined (as in Assumption 1) with respect to some external relatum, i.e. that may contain unspeakable information. This task becomes interesting if Alice's and Bob's reference frames are initially unaligned.

Suppose that Alice sends the quantum system physically to Bob. Since Bob's reference frame is not aligned with Alice's, he will describe the situation as receiving a randomly sampled representative of the equivalence class $[\rho]$. Thus, he will assign the state $\mathcal{E}(\rho) := \int_{\mathcal{G}} U_g \rho U_g^\dagger \, dg$ to the incoming quantum system.

The QI approach is concerned with the possibility to devise protocols that can be performed even in the absence of a shared reference frame. For example, the task to send quantum information from Alice to Bob can be accomplished by encoding it into a *decoherence-free subspace*, i.e. a subsystem within the set of $\rho \in \mathcal{S}(\mathcal{H})$ for which $\mathcal{E}(\rho) = \rho$ (see e.g. Ref. [1, Sec. A.2] for a concrete example). Another possibility to do so is by sending several quantum systems (e.g. spin-coherent states) that break the symmetry, and that allow Bob to partially correlate his reference frame with Alice's via suitable measurements on those states. The key to carrying out communication protocols without a shared frame is thus to focus on invariant physical properties that are meaningful in any external laboratory frame. This does not require describing S relative to one of its subsystems.

Nevertheless, in the QI approach, the quantum nature of reference frames is sometimes taken into account, for example, by "quantizing" them to overcome superselection rules that arise in the absence of a shared classical frame [1]. This "quantization" of a frame means *adding* a reference quantum system R to the system of interest S in order to define relative quantities between R and S, such as relative phases [1] or relative distances [2, 8], that are invariant under $\mathcal{G}_{\text{sym}}$ and thereby meaningful relative to *any* external laboratory frame. In a communication scenario between two parties Alice and Bob who do not share a classical frame, the reference system R will typically be communicated together with S. While this also constitutes an internalization of a frame in the sense thats the reference system R is now a quantum system too, it is still external to S. Furthermore, since the relative quantities between R and S are meaningful relative to any external laboratory frame with respect to which a measurement will be carried out, it is not necessary to take an extra step and ask how S is described "from the perspective" of R in order for Alice and Bob to succeed in their communication task.

In summary, in the QI approach, the quantum system S of interest (say, a set of spins) is treated as a distinct entity from the reference frame (say, a gyroscope). Thus, "QRF transformations" relating descriptions relative to different subsystems (which may be in relative superposition) are typically not studied in this approach.[1] The focus is on *correlating* (aligning) Alice's and Bob's frames, and it is the absence of alignment that is modelled by the $\mathcal{G}$-twirl, $\rho \mapsto \mathcal{E}(\rho)$. The external relatum independent (or relational) state descriptions of the QI approach are thus the *incoherently* group-averaged states.

[1] This includes Ref. [7], where transformations between different "quantized" reference systems R_1 and R_2 are studied. However, in the spirit of the QI approach, the derived transformations proceed between different invariant states (i.e. essentially $\mathcal{G}$-twirls of $\rho_S \otimes \rho_{R_i}$, $i = 1, 2$) and are thus *not* transformations between descriptions of the quantum state of S relative to different choices of subsystem, as we will see them later. In particular, the descriptions of the quantum state of S relative to different subsystems will be different descriptions of one and the same invariant state.

2. *The structural approach: agreeing on a redundancy-free internal description of quantum states*

The structural approach does not stop at an external relatum independent state description. It also asks for a description of a quantum state relative to an internal frame that is part of the system of interest.

A transparent way to understand the structural approach operationally is as follows. Alice and Bob in their respective labs would like to agree on a concrete description of the quantum state of a system *without* external relatum, i.e. in particular without shared reference frame. They have the option of describing S in terms of the equivalence classes $[\rho]$ of quantum states. However, there is an evident *redundancy* in the description of each equivalence class in terms of concrete quantum states: each member of the equivalence class is a legitimate (and non-unique) description of it. In order to break this redundancy and succeed in this task, they can take advantage of the fact that any equivalence class $[\rho]$ of states admits certain "canonical choices" for its description which are associated with different internal reference frame choices. The transformations relating these different canonical choices amount to "QRF transformations" and they will be elements of the symmetry group $\mathcal{G}_{\text{sym}}$ defining the equivalence classes.

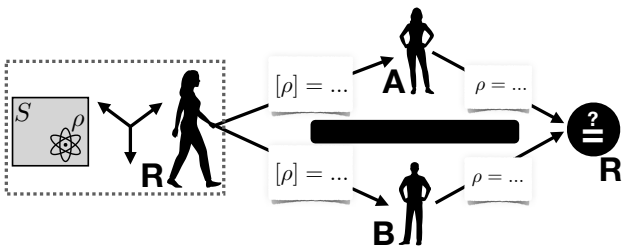


FIG. 5. A simple communication scenario which we choose for illustrating the operational essence of the **structural approach** as in Refs. [38–40]. The focus is on agents agreeing on a (redundancy-free) *description* of quantum states in the absence of an external relatum.

For example, one could imagine the following communication scenario depicted in Figure 5 to illustrate the role of "canonical choices":

- Referee Refaella informs Alice and Bob in their separate laboratories that she will prepare quantum states of a particular system S (subject to Assumption 1) relative to her (freely aligned) frame, but that she will only communicate the *description* of the respective equivalence classes $[\rho]$ to Alice and Bob separately.
- Alice's and Bob's task is to separately return a concrete (redundancy-free) description of each quan-

tum state to Refaella and they will win the game provided their descriptions always agree (either for all states of S, or for a particular class $\mathcal{C}$ of states).

- Alice and Bob are only permitted to communicate prior to the beginning of the game to agree on a strategy.

Let us consider two examples for how this can be accomplished. These examples illustrate that there will generally exist multiple "canonical choices" for describing $[\rho]$ in terms of concrete quantum states, however, that Alice and Bob can always agree in their communication beforehand which such choice to pick. This will also give a hint on the relation to quantum reference frames as described in Refs. [38–40], and we will elaborate on this further in the following sections.

Example 1. *Consider a single quantum spin-*$1/2$ *particle, with state space* $\mathcal{S}(\mathbb{C}^2)$*. Let us assume that the symmetry group is the full projective unitary group, i.e.* $\mathcal{G}_{\text{sym}} = \{\rho \mapsto U\rho U^\dagger \mid U^\dagger U = \mathbf{1}\}$*, which is isomorphic to the rotation group* $\mathrm{SO}(3)$*.*

Let ρ *be an arbitrary state that Refaella is for some reason interested in preparing. The equivalence class* $[\rho]$ *consists of all states with the same eigenvalues* λ_1, λ_2 *as* ρ*. Describing* $[\rho]$ *is equivalent to listing the eigenvalues* λ_1, λ_2 *and this is what Refaella may communicate to Alice and Bob. Clearly, there are many ways to represent this information in terms of a concrete quantum state* ρ*.*

The strategy that Alice and Bob can agree on in order to win the game, but prior to it starting, is trivial: they can agree to always choose a basis (i.e. a specific reference frame alignment) such that ρ *described relative to it is a diagonal matrix. This leaves two "canonical choices" of representation: ordering the eigenvalues such that* $\lambda_1 \geq \lambda_2$ *they could decide to always return either* $\rho = \mathrm{diag}(\lambda_1, \lambda_2)$ *or* $\rho = \mathrm{diag}(\lambda_2, \lambda_1)$ *to Refaella. The transformation relating the two descriptions is the unitary "QRF transformation"* $U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$*.*

This trivial example relies on the simple fact that every quantum state has a *canonical description*: the matrix representation in its own eigenbasis (up to a choice of order of eigenvalues). In some sense, every quantum state defines a finite set of natural representations of itself. It is in this sense that the structural approach interprets quantum systems as *quantum reference frames*: the system's state breaks the fundamental symmetry, and admits, at least on the level of classical descriptions, a canonical choice of representation.

Example 1 illustrates a general consequence of the symmetry structure: for any particular choice of QRF, the *set* of state descriptions relative to that QRF corresponds in general only to a *subset* or *subspace* of states. In this example, any such choice only allows to describe a *subset of states that corresponds to a classical bit*: namely, the convex hull of the density matrices $\mathrm{diag}(1, 0)$ and $\mathrm{diag}(1/2, 1/2)$. The following example demonstrates how a full subspace of states can be encoded.

Example 2. *Consider* two *spin -*$1/2$ *particles with rotational symmetry. That is, the symmetry group is* $\mathcal{G}_{\text{sym}} = \{\rho \mapsto U \otimes U\rho U^\dagger \otimes U^\dagger \mid U \in \mathrm{SU}(2)\}$*, acting on states in* $\mathcal{S}(\mathbb{C}^2 \otimes \mathbb{C}^2)$*. Let us make a somewhat arbitrary, but nonetheless illustrative choice of a class* $\mathcal{C}$ *of states for which the above communication game can be played. These will be the pure states*

$$\mathcal{C} = \left\{ \cos\frac{\theta}{2}|\phi_-\rangle + e^{i\varphi}\sin\frac{\theta}{2}|\phi\rangle \otimes |\phi\rangle \right\},$$

where $0 \leq \theta \leq \pi$*,* $0 \leq \varphi < 2\pi$*,* $|\phi_-\rangle$ *is the singlet state, and* $|\phi\rangle \in \mathbb{C}^2$ *an arbitrary normalized state. The set of states* $\mathcal{C}$ *is the disjoint union of the sets* $\mathcal{C}_{\theta,\varphi}$ *for which the two angles are fixed and* $|\phi\rangle$ *is still an arbitrary qubit state. Since* $U \otimes U|\phi_-\rangle = |\phi_-\rangle$*, the* $\mathcal{C}_{\theta,\varphi}$ *are orbits of the symmetry group, i.e. equivalence classes of states.*

If Refaella gives Alice and Bob a description of such an equivalence class $[|\psi\rangle] = \mathcal{C}_{\theta,\varphi}$*, they can agree on returning the standard description* $|\psi'\rangle = \cos\frac{\theta}{2}|\phi_-\rangle + e^{i\varphi}\sin\frac{\theta}{2}|0\rangle \otimes |0\rangle$*, for example. This prescription has the added benefit of* preserving superposition across different equivalence classes*. Namely, if for* $i = 1, 2$*, we have* $|\psi_i\rangle = \alpha_i|\phi_-\rangle + \beta_i|\phi\rangle \otimes |\phi\rangle$ *such that* $|\alpha_1| \neq |\alpha_2|$*, then* $|\psi_1\rangle$ *and* $|\psi_2\rangle$ *are in different equivalence classes, and so are (in general) their superpositions. But the states that Alice and Bob return respect superpositions: if* $|\psi\rangle = \kappa|\psi_1\rangle + \lambda|\psi_2\rangle$*, then the returned states satisfy* $|\psi'\rangle = \kappa|\psi'_1\rangle + \lambda|\psi'_2\rangle$*. That is, this choice of QRF admits the description of a subspace, a* qubit*, inside the joint state space. Other choices of QRF do so as well. These would correspond to canonical descriptions where* $|0\rangle \otimes |0\rangle$ *is replaced by some arbitrary* $|\phi_0\rangle \otimes |\phi_0\rangle$*, and they are related by "QRF transformations"* $U \otimes U$*.*

There are also seemingly natural choices of QRF that, however, are deficient in that the set of admissible descriptions relative to them cannot encompass a state space, as the following example illustrates.

Example 3. *Consider again two spin-*$1/2$ *particles, but now under slightly different circumstances. There is a canonical choice of factorization of the Hilbert space: by looking at the system in isolation, observers can determine the decomposition into two distinguishable particles. If we assume that this is the only structure that can be determined by such observers, then we have the symmetry group* $\mathcal{G}_{\text{sym}} = \{\rho \mapsto U \otimes V\rho U^\dagger \otimes V^\dagger \mid U^\dagger U = V^\dagger V = \mathbf{1}\}$*. Under these circumstances, a canonical choice of frame is such that any pure state* $|\psi\rangle$ *becomes identical to its own Schmidt representation,* $|\psi\rangle = \sum_{i=0}^{1} \sqrt{\alpha_i}|ii\rangle$*, where* $\alpha_0 \geq \alpha_1$*.*

While Alice and Bob could easily agree on such a convention, the ensuing canonical description would not preserve complex superpositions and, in particular, not lead to a subspace of states as its image, owing to the real nature and ordering of the Schmidt-coefficients.

A priori, a choice of QRF in the structural approach can therefore be quite arbitrary. However, as the examples above motivate, a "good" choice of QRF will correspond to one that admits the description of a set of states relative to it which carries sufficient convex or linear structure to encode classical or quantum information. Preferably, that set of states should correspond to a subspace of maximal size within $\mathcal{C}$.

In the remainder of this article, we will focus on a more interesting realization of such a scenario which reproduces the notion of QRFs in the structural picture. We will define particular systems S that we call "$\mathcal{G}$-systems", and we will see that these carry an interesting group of symmetries $\mathcal{G}_{\text{sym}}$. If we ask what kind of canonical choices of (redundancy-free) description $\mathcal{G}$-systems admit, such that Alice and Bob can succeed in the communication scenario of Figure 5, we will find that these correspond to choosing one of the subsystems of S as a reference system and to describing the remaining degrees of freedom relative to it. In this manner, we will recover and generalize the "quantum states relative to a particle" of Refs. [38–40]. In particular, the transformations among the canonical choices of description of S are elements of the symmetry group $\mathcal{G}_{\text{sym}}$ and exactly the QRF transformations of Ref. [40], which are also equivalent to the ones in [38, 39] (restricted to a discrete setting). In Ref. [62] we will further explicitly demonstrate the equivalence with the perspective-neutral approach to QRFs [39] and elucidate that any equivalence class $[\rho]$ of quantum states above corresponds precisely to a perspective-neutral quantum state. As we will see, this means that the relational states of the structural approach are *coherently* (not incoherently as in the QI approach) group-averaged states.

III. FROM SYMMETRIES TO QRF TRANSFORMATIONS AND INVARIANT OBSERVABLES

Quantum reference frames as described in Refs. [38–40] have first been considered for the case of wave functions on the real line. We have a Hilbert space of square-integrable functions, $\mathcal{H} = L^2(\mathbb{R})$, and a physical claim that there is no absolute notion of origin. In other words, the "physics" does not change under translations (we will soon formulate what this means in detail). If we have N particles on the real line, the total Hilbert space is $L^2(\mathbb{R})^{\otimes N}$.

As noted in Ref. [40], the real numbers $\mathbb{R}$ play a double role in this case: on the one hand, they label the configuration space on which the wave functions are supported; on the other hand, they also label the possible translations, i.e. the fundamental symmetry group $(\mathbb{R}, +)$.

In this section, we will analyze this particular situation in a simplified setting: one in which the group is finite and Abelian. In the simplest case, we discretize the real line and make it periodic, as in Figure 1. Formally, for some $n \in \mathbb{N}$, we consider the *cyclic group*

$$\mathbb{Z}_n := \{0, 1, 2, \ldots, n-1\}$$

with addition modulo n as its group operation. To this, we associate a single-particle Hilbert space

$$\mathcal{H} = \ell^2(\mathbb{Z}_n) = \text{span}\{|0\rangle, |1\rangle, \ldots, |n-1\rangle\}$$

and a total Hilbert space $\mathcal{H}^{\otimes N}$ for N distinguishable particles. We will denote the particles with labels $A, B, C, \ldots$, and later in this paper with integers $1, 2, 3, \ldots$. Within this formalism, we can realize the main ideas of quantum references frames as in Refs. [38–40]. For the case $N = 2$, consider the quantum state

$$|\psi\rangle_{AB} = |0\rangle_A \otimes \frac{1}{\sqrt{2}} (|1\rangle + |2\rangle)_B .$$

We are interested in a situation where "only the relation between the particles" matters, but not their total position. That is, in some sense, "applying elements of $\mathbb{Z}_n$ to a quantum state doesn't change the physics". Intuitively, this means, for example, that the quantum state

$$|\psi'\rangle_{AB} = |1\rangle_A \otimes \frac{1}{\sqrt{2}} (|2\rangle + |3\rangle)_B$$

should be an equivalent description of the system's properties, since it is related to $|\psi\rangle$ by a translation. Motivated by Ref. [38], we can do something more interesting. First, in the terminology of Refs. [38–40], the form of $|\psi\rangle$ can be interpreted as saying that "particle B, as seen by A, is in the state $\frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)$". Second, we can then use the prescription of Refs. [38–40] to "jump into B's reference frame", and consider the state

$$|\psi''\rangle = \frac{1}{\sqrt{2}}(|n-2\rangle + |n-1\rangle)_A \otimes |0\rangle_B$$

and conclude that "particle A, as seen by B, is in the state $\frac{1}{\sqrt{2}}(|n-1\rangle + |n-2\rangle)$". After all, this still expresses the fact that with amplitudes $\frac{1}{\sqrt{2}}$, B is either one or two positions to the right of A.

We will now show that we can understand these transformations as natural symmetry transformations in a simple class of physical systems which we call "$\mathcal{G}$-systems". Choosing one of the particles as a reference frame (as sketched above) will correspond to a choice of canonical representation of a state as in the structural approach outlined above. This will give the idea of "jumping into a particle's perspective" a thorough operational interpretation.

A. $\mathcal{G}$-systems and their symmetries

We begin by considering a specific physical system which is motivated by translation-invariant quantum

physics on the real line with Hilbert space $L^2(\mathbb{R})$. Here we consider a finite, discrete group-theoretic analogue, again using the group as both the configuration space and set of transformations. Some aspects of QRF transformations in this case were also considered in Ref. [40]. In contrast to Ref. [40], we restrict our attention to finite Abelian groups $\mathcal{G}$ for simplicity. Due to the structure theorem [63], every such $\mathcal{G}$ can be interpreted as the group of translations of a discrete torus of some dimension. In the simplest case where $\mathcal{G} = \mathbb{Z}_n$, this torus is the circle[2], and we are in the setting of Figure 1.

Definition 4 ($\mathcal{G}$-system)**.** *Fix some finite Abelian group $\mathcal{G}$, interpreted as a classical configuration space. That is, we regard the $g \in \mathcal{G}$ as perfectly distinguishable orthonormal basis vectors $|g\rangle$, spanning a Hilbert space $\mathcal{H}$. Formally, this Hilbert space is $\mathcal{H} = \ell^2(\mathcal{G})$, and it carries a distinguished basis $\{|g\rangle\}_{g\in\mathcal{G}}$, similarly as quantum mechanics on the real line carries a distinguished position basis.*

Consider N distinguishable particles on such a classical configuration space, where $N \in \mathbb{N}$. That is, the total Hilbert space is $\mathcal{H}^{\otimes N}$, and it carries a natural orthonormal basis

$$\mathcal{H}^{\otimes N} = \mathrm{span}\{|g_1, \ldots, g_N\rangle \mid g_i \in \mathcal{G}\}.$$

The physical system S described by this Hilbert space will carry a group of symmetries $\mathcal{G}_{\mathrm{sym}}$ as introduced in Assumption 1 and Figure 3. Clearly, the basic Hilbert space structure of S, i.e. the notion of linearity and the inner product, must not depend on the orientation of the external reference frame. Hence, the symmetry group will be of the form

$$\mathcal{G}_{\mathrm{sym}} = \{U \bullet U^\dagger \mid U \in \mathcal{U}_{\mathrm{sym}}\},$$

for $\mathcal{U}_{\mathrm{sym}}$ some group of unitaries. Furthermore, we assume that the classical configuration space, i.e. the set *of basis vectors, $\{|g_1, \ldots, g_N\rangle \mid g_i \in \mathcal{G}\}$, is an internal structure of S that is defined without the external reference frame. We now postulate that the classical configurations carry $\mathcal{G}$-symmetry. In particular, any given configuration*

$$|\mathbf{g}\rangle := |g_1, g_2, \ldots, g_N\rangle$$

and its "translated" version

$$U_g^{\otimes N}|\mathbf{g}\rangle = |g\mathbf{g}\rangle := |gg_1, gg_2, \ldots, gg_n\rangle$$

are internally indistinguishable. On the other hand, we postulate that the relation between the particles **is** accessible to observers without the external frame. *To formalize this, consider some tuple $\mathbf{h} \in \mathcal{G}^{N-1}$ of group elements, i.e. $\mathbf{h} = (h_1, \ldots, h_{N-1})$. Any state of the form*

$$|g, h_1 g, h_2 g, \ldots, h_{N-1} g\rangle =: |g, \mathbf{h}g\rangle \tag{1}$$

has the same pairwise relations between its particles, no matter what the state $|g\rangle$ of the first particle is. We now define $\mathcal{G}_{\mathrm{sym}}$ as the largest possible symmetry group that is compatible with these postulates. To this end, $\mathcal{U}_{\mathrm{sym}}$ must be the group of unitary transformations with the following properties:

1. *U maps classical configurations to classical configurations, i.e. $U|g_1, ..., g_n\rangle = |g'_1, ..., g'_n\rangle$.*
2. *On classical configurations, U preserves relative positions, i.e. $U|g, \mathbf{h}g\rangle = |g', \mathbf{h}g'\rangle$.*
3. *If two classical configurations are g-translations of each other, then U preserves this fact[3], i.e.*

$$|\mathbf{g}\rangle = U_g^{\otimes N}|\mathbf{j}\rangle \Rightarrow U|\mathbf{g}\rangle = U_g^{\otimes N}\left(U|\mathbf{j}\rangle\right).$$

A few words of justification are in place. While two choices of external reference frame may yield a different description of any configuration, they must agree on the *set* of all possible configurations that S can be in, for otherwise their descriptions of S cannot be placed in full relation with one another.[4] The *set* of basis vectors $\{|\mathbf{g}\rangle\}_{\mathbf{g}\in\mathcal{G}^N}$ must thus be independent of the external frame and hence should remain invariant under $\mathcal{G}_{\mathrm{sym}}$. It is also clear that the symmetry group must preserve the linear and probabilistic structure of quantum theory and thereby leave the inner product on $\mathcal{H}^{\otimes N}$ invariant.[5] After all, by Assumption 1, symmetry related quantum states should be indistinguishable even probabilistically. Furthermore, the $\mathbf{h}$ label the 'relative positions' among the N particles. These are internal properties of S and so independent of any external relatum. Finally, configurations that are g-translations of each other are by assumption internally indistinguishable. The symmetry group must preserve this indistinguishability.

The symmetry group of a $\mathcal{G}$-system can now easily be written down. To this end, define the subspaces

$$\mathcal{H}_{\mathbf{h}} := \mathrm{span}\{|g, \mathbf{h}g\rangle \mid g \in \mathcal{G}\}$$

and the corresponding orthogonal projectors $\Pi_{\mathbf{h}}$. Note that $\mathcal{H}^{\otimes N} = \bigoplus_{\mathbf{h}\in\mathcal{G}^{N-1}} \mathcal{H}_{\mathbf{h}}$, and the $\{\Pi_{\mathbf{h}}\}_{\mathbf{h}\in\mathcal{G}^{N-1}}$ define a projective measurement.

[2] This representation is not unique. For example, we can interpret $\mathbb{Z}_6$ as the translation group of six points on a circle, but the structure theorem tells us that $\mathbb{Z}_6 \simeq \mathbb{Z}_2 \times \mathbb{Z}_3$. Thus, we can also interpret this group as the translations of a two-dimensional (2×3)-torus.

[3] It is possible to drop this condition, and to assume only 1. and 2. In this case, one obtains similar results to those presented here, but with modified structures: the algebra of invariant operators then becomes what we call $\mathcal{A}_{\mathrm{alg}}$ in Lemma 26, and the symmetry group becomes the group of conditional *permutations* (not only conditional translations). Physically, this does not seem particularly well-motivated, and it leads to the loss of certain uniqueness results, including the uniqueness of $U \in \mathcal{U}_{\mathrm{sym}}$ in Theorem 18.

[4] This assumes that the external frame choices in the ambient laboratory that an agent may have access to are complete in the sense that all quantum properties of S can be described relative to them.

[5] In constraint quantization, $\mathcal{H}^{\otimes N}$ corresponds to the kinematical Hilbert space and so the preservation refers here to the kinematical inner product. While one is usually only interested in the physical inner product (i.e. the inner product on the space of solutions to the constraints), it nevertheless holds that also the kinematical inner product is left invariant by the group generated by the (self-adjoint) constraints.

Lemma 5. *The symmetry group of a $\mathcal{G}$-system is*

$$\mathcal{U}_{\mathrm{sym}} = \left\{ U = \bigoplus_{\mathbf{h}\in\mathcal{G}^{N-1}} U_{g_{\mathbf{h}}}^{\otimes N} \;\middle|\; g_{\mathbf{h}} \in \mathcal{G} \right\}, \tag{2}$$

where $U_{g_{\mathbf{h}}}^{\otimes N}$ denotes the global translation by $g_{\mathbf{h}}$, but restricted to the subspace $\mathcal{H}_{\mathbf{h}}$.

That is, the symmetries in $\mathcal{U}_{\mathrm{sym}}$ act as *relation-conditional global translations*: every classical configuration is globally translated via some $U_{g_{\mathbf{h}}}^{\otimes N}$, but the amount of translation $g_{\mathbf{h}}$ may depend on the relation $\mathbf{h}$ between the particles. We will soon identify the QRF transformations of Refs. [38–40] with elements of this group. Thus, the above highlights that these transformations make sense in a purely classical context; indeed, the corresponding classical frame transformations were also studied in [39, 40] and shown to be conditional on the interparticle relations.[6] For example, they can also be applied if one deals with statistical mixtures of particle positions instead of superpositions. Nonetheless, their unitary extension to all of $\mathcal{H}^{\otimes N}$ leads to interesting quantum effects like the frame-dependence of entanglement [38–40]. This is similar to the behavior of the CNOT gate in quantum information theory, which is defined by its classical action on two bits, but nonetheless can create entanglement.

Proof. Due to conditions 1. and 2. of Definition 4, the $U \in \mathcal{U}_{\mathrm{sym}}$ leave every $\mathcal{H}_{\mathbf{h}}$ invariant. Thus, U decomposes into a direct sum $U = \bigoplus_{\mathbf{h}\in\mathcal{G}^{N-1}} U_{\mathbf{h}}$. Fix some $\mathbf{h} \in \mathcal{G}^{N-1}$. Since $\mathcal{H}_{\mathbf{h}}$ is invariant, there exists some $g_{\mathbf{h}} \in \mathcal{G}$ such that $U|e,\mathbf{h}\rangle = |g_{\mathbf{h}}, \mathbf{h}g_{\mathbf{h}}\rangle$. Now, for every $g \in \mathcal{G}$, we have $|g,\mathbf{h}g\rangle = U_g^{\otimes N}|e,\mathbf{h}\rangle$. Thus, condition 3. of Definition 4 implies that

$$\begin{aligned} U|g,\mathbf{h}g\rangle &= U_g^{\otimes N}\left(U|e,\mathbf{h}\rangle\right) = U_g^{\otimes N}|g_{\mathbf{h}},\mathbf{h}g_{\mathbf{h}}\rangle \\ &= U_{gg_{\mathbf{h}}}^{\otimes N}|e,\mathbf{h}\rangle = U_{g_{\mathbf{h}}}^{\otimes N}|g,\mathbf{h}g\rangle. \end{aligned}$$

This shows that $U_{\mathbf{h}}$ acts like $U_{g_{\mathbf{h}}}^{\otimes N}$ on $\mathcal{H}_{\mathbf{h}}$. □

When working with pure state vectors, we sometimes want to allow global phases. Thus, we use the notation

$$\mathcal{U}_{\mathrm{sym}}^{*} := \mathcal{U}_{\mathrm{sym}} \times \mathrm{U}(1) = \{e^{i\theta}U \mid U \in \mathcal{U}_{\mathrm{sym}}, \theta \in \mathbb{R}\}.$$

Above, we have decided to denote the state of the particles *relative to the first particle*, but this also defines the relations between all other pairs of particles: the equation $|\mathbf{g}\rangle = |g,\mathbf{h}g\rangle \in \mathcal{H}_{\mathbf{h}}$ means that $g_i = h_{i-1}g_1$ for $i \geq 2$, but this implies that $g_i = (h_{i-1}h_{j-1}^{-1})g_j$ for all i,j if we set $h_0 := e$, the unit element of the group. Thus, the $\mathcal{H}_{\mathbf{h}}$ decompose the global Hilbert space into sectors of equal pairwise relations.

It is clear that global $\mathcal{G}$-translations are elements of the symmetry group, but they do not exhaust it:

Example 6. *Given any $\mathcal{G}$-system, the global translations $U_g^{\otimes N}$ are symmetry transformations. Since they represent the global action of $\mathcal{G}$ on the N-particle Hilbert space, this can be written as*

$$\mathcal{G} \subset \mathcal{G}_{\mathrm{sym}}.$$

However, there are other symmetries that are not global translations. For example, for $N = 2$ particles, the unitary U which acts on all basis vectors $|g_1,g_2\rangle$ as

$$U|g_1,g_2\rangle := |g_2, g_1^{-1}g_2^2\rangle$$

is a symmetry transformation, i.e. $U \in \mathcal{U}_{\mathrm{sym}}$. Namely, $|g_1,g_2\rangle \in \mathcal{H}_h$ for $h = g_1^{-1}g_2$, and U implements the global translation $U_{g(h)}^{\otimes 2}$ on $\mathcal{H}_h$, where $g(h) = h$.

On the other hand, the transformation

$$V|g_1,g_2\rangle := |g_2^{-1}, g_1^{-1}\rangle$$

is not *a symmetry transformation: it satisfies conditions 1. and 2. of Definition 4, but violates condition 3.*

We will later see that QRF transformations correspond to elements in $\mathcal{G}_{\mathrm{sym}} \setminus \mathcal{G}$.

B. Invariant observables and Hilbert space decomposition

Which observables can we internally measure in a $\mathcal{G}$-system, i.e. without access to the external relatum that was used to define the state space and the symmetry group? These must be the observables that are invariant under all symmetry transformations and which thus correspond to speakable information:

Definition 7 (Invariant observable)**.** *We define the* invariant subalgebra $\mathcal{A}_{\mathrm{inv}}$ *as*

$$\mathcal{A}_{\mathrm{inv}} = \{A \in \mathcal{L}(\mathcal{H}^{\otimes N}) \mid [U,A] = 0 \text{ for all } U \in \mathcal{U}_{\mathrm{sym}}\},$$

where $\mathcal{L}(\mathcal{H})$ denotes the set of linear operators on Hilbert space $\mathcal{H}$. These are the operators A that are invariant under all symmetry transformations $A \mapsto UAU^{\dagger}$. A self-adjoint element $A = A^{\dagger} \in \mathcal{A}_{\mathrm{inv}}$ is called an invariant observable.

Since all observable properties of our system are assumed to be invariant under $\mathcal{G}_{\mathrm{sym}}$, it follows that the observables in Definition 7 comprise the set of *all* observables that can be physically measured by an observer who does not have access to the external reference frame.

Clearly, all the $\Pi_{\mathbf{h}}$ are invariant observables, i.e. $\Pi_{\mathbf{h}} \in \mathcal{A}_{\mathrm{inv}}$. However, due to the fact that we have declared a classical basis to be a distinguished structure of the $\mathcal{G}$-system, there are many more invariant observables. To

[6] More precisely, in the perspective-neutral approach these classical reference frame transformations correspond to *conditional* gauge transformations, i.e. the gauge flow distance depends on the subsystem relations, see Appendix B of Ref. [39] and also Refs. [41–43].

determine the algebra $\mathcal{A}_{\text{inv}}$, recall the decomposition of $U \in \mathcal{U}_{\text{sym}}$ from Lemma 5. We can regard $\mathcal{U}_{\text{sym}}$ as a representation of several copies of the group $\mathcal{G}$, and thus further refine this decomposition via basic representation theory of finite Abelian groups [63].

A major role is played by the *characters* of $\mathcal{G}$. These are the homomorphisms $\chi : \mathcal{G} \to S^1$, i.e. the maps from $\mathcal{G}$ to the complex unit vectors $S^1 := \{z \in \mathbb{C}, \mid |z| = 1\}$ with $\chi(gh) = \chi(g)\chi(h)$. In other words, the characters are the one-dimensional irreducible representations (irreps) of $\mathcal{G}$, and these turn out to exhaust all irreps. The set of all characters of $\mathcal{G}$ is denoted $\widehat{\mathcal{G}}$.

Denote the order of the group by $n := |\mathcal{G}|$, then $g^n = e$ for all $g \in \mathcal{G}$ [63]. Thus, every $\chi(g)$ must be among the n-th roots of unity: $\chi(g)^n = 1$. Moreover, there are exactly n characters, i.e. $|\widehat{\mathcal{G}}| = n$.

Furthermore, note that $\dim \mathcal{H}_{\mathbf{h}} = n$. We claim that these subspaces can be decomposed as follows:

$$\mathcal{H}_{\mathbf{h}} = \bigoplus_{\chi \in \widehat{\mathcal{G}}} \mathcal{H}_{\mathbf{h};\chi}$$

with $\mathcal{H}_{\mathbf{h};\chi}$ the one-dimensional subspace spanned by the vector

$$|\mathbf{h};\chi\rangle := \frac{1}{\sqrt{|\mathcal{G}|}} \sum_{g \in \mathcal{G}} \chi(g^{-1})|g, \mathbf{h}g\rangle. \tag{3}$$

Indeed, due to Ref. [63, Proof of Corollary III.2.3], we have the well-known orthogonality relations $\sum_{g\in\mathcal{G}} \overline{\chi(g)}\chi'(g) = n\delta_{\chi,\chi'}$. Using this, direct calculation shows that the $|\mathbf{h};\chi\rangle$ are orthonormalized states, and

$$U_g^{\otimes N}|\mathbf{h};\chi\rangle = \chi(g)|\mathbf{h};\chi\rangle \quad \text{for all } g \in \mathcal{G}. \tag{4}$$

Example 8. *As a simple example, consider the cyclic group $\mathcal{G} = \mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ with addition modulo n, see Figure 4. This group can be interpreted as a finite analogue of a part of the real line with periodic boundary conditions, by distributing finitely many possible positions along a ring. Its irreducible representations and the respective characters are given by $\chi_k(g) := e^{i\frac{2\pi}{n}kg}$ with $k \in \{0, 1, \ldots, n-1\}$ [70]. Indeed, one can directly verify that the χ_k form one-dimensional representations of $\mathbb{Z}_n$, and they are inequivalent. We explicitly obtain*

$$|\mathbf{h};\chi_k\rangle = \frac{1}{\sqrt{n}} \sum_{g=0}^{n-1} e^{-i\frac{2\pi}{n}kg} |g, g+\mathbf{h}\rangle ,$$

where $g + \mathbf{h}$ means that g is added to each component of $\mathbf{h}$, modulo n. Similarly, one can directly verify that

$$U_g^{\otimes N} |\mathbf{h};\chi_k\rangle = e^{i\frac{2\pi}{n}kg} |\mathbf{h};\chi_k\rangle .$$

One can see that the $|\mathbf{h};\chi_k\rangle$ are obtained via a kind of discrete Fourier transform [69] from the classical configurations, and therefore they are reminiscent of momentum eigenstates.

Since elements of $\mathcal{U}_{\text{sym}}$ translate all particles by the same amount, and momentum is the generator of translations, one may identify $|\mathbf{h};\chi_k\rangle$ with the eigenstates of total momentum. However, since we are not explicitly interested in dynamics in this paper, we will postpone any elaboration on this analogy to our upcoming work, Ref. [62].

From Eqs. (2)–(4) it is clear that the subspace spanned by the eigenstates with trivial character $\chi = \mathbf{1}$ is the subspace of $\mathcal{U}_{\text{sym}}$-invariant states, $|\psi\rangle = U|\psi\rangle$ for all $U \in \mathcal{U}_{\text{sym}}$. We denote it by

$$\mathcal{H}_{\text{phys}} := \bigoplus_{\mathbf{h}\in\mathcal{G}^{N-1}} \mathcal{H}_{\mathbf{h};\mathbf{1}} = \text{span}\left\{|\mathbf{h};\mathbf{1}\rangle \mid \mathbf{h} \in \mathcal{G}^{N-1}\right\}. \tag{5}$$

We have equipped the total invariant subspace with the label "phys" because it is the finite group version of the so-called physical Hilbert space of constraint quantization. When the symmetry group is generated by (self-adjoint) constraints, the physical Hilbert space corresponds to the set of solutions to the quantum constraints and is thereby precisely the Hilbert space on which the group acts trivially. It is usually called 'physical' because quantum states of a gauge system are required to satisfy the constraints imposed by gauge symmetry. Nevertheless, we will see that we can give quantum states that are not invariant under the symmetry group a useful physical interpretation, and we will clarify their relation with the 'physical' states in $\mathcal{H}_{\text{phys}}$ in Ref. [62]. Being spanned by the states $|\mathbf{h};\mathbf{1}\rangle$ which encode the particle relations in an invariant manner, we shall henceforth also refer to $\mathcal{H}_{\text{phys}}$ as the subspace of *relational states*. Its dimension is $|\mathcal{G}|^{N-1}$, and thus:

Lemma 9. *The subspace of relational states $\mathcal{H}_{\text{phys}}$ is isomorphic to $\mathcal{H}^{\otimes(N-1)}$.*

To determine the invariant subalgebra, consider any $A \in \mathcal{L}(\mathcal{H}^{\otimes N})$ and develop it into the $|\mathbf{h};\chi\rangle$-eigenbasis: $A = \sum_{\mathbf{h},\mathbf{h}',\chi,\chi'} a_{\mathbf{h},\mathbf{h}',\chi,\chi'}|\mathbf{h};\chi\rangle\langle\mathbf{h}';\chi'|$. Using Eqs. (2) and (4), conjugation with some $U \in \mathcal{U}_{\text{sym}}$ yields

$$UAU^\dagger = \sum_{\mathbf{h},\mathbf{h}',\chi,\chi'} \chi(g_{\mathbf{h}})\chi'(g_{\mathbf{h}'})^{-1} a_{\mathbf{h},\mathbf{h}',\chi,\chi'}|\mathbf{h};\chi\rangle\langle\mathbf{h}';\chi'|.$$

This is equal to A for all U if and only if for all $\mathbf{h}, \mathbf{h}', \chi, \chi'$, one of the following is true: either $a_{\mathbf{h},\mathbf{h}',\chi,\chi'} = 0$ or $\chi(g_{\mathbf{h}}) = \chi'(g_{\mathbf{h}'})$ for all possible choices of $g_{\mathbf{h}}$, $g_{\mathbf{h}'}$. The latter condition is automatically satisfied if $\chi = \chi' = \mathbf{1}$. Thus, all operators A that are fully supported on the relational subspace $\mathcal{H}_{\text{phys}}$ will be elements of $\mathcal{A}_{\text{inv}}$. Let us denote the set of such operators by $\mathcal{A}_{\text{phys}}$, then we have just shown that $\mathcal{A}_{\text{phys}} \subset \mathcal{A}_{\text{inv}}$. For reasons that will become clear later, we will call the observables in $\mathcal{A}_{\text{phys}}$ *relational observables*.

Now consider the other cases in which at least one of χ or χ' differs from $\mathbf{1}$. Clearly, if $\mathbf{h} = \mathbf{h}'$ and $\chi = \chi'$ then the character condition $\chi(g_{\mathbf{h}}) = \chi'(g_{\mathbf{h}'})$ is trivially satisfied, and $a_{\mathbf{h},\mathbf{h},\chi,\chi}$ does not need to be zero. Consider the case $\mathbf{h} = \mathbf{h}'$ and $\chi \neq \chi'$. Choosing any $g_{\mathbf{h}}$ with $\chi(g_{\mathbf{h}}) \neq$

$\chi'(g_{\mathbf{h}})$ shows that we must have $a_{\mathbf{h},\mathbf{h},\chi,\chi'} = 0$. Finally, if $\mathbf{h} \neq \mathbf{h}'$ and at least one of χ or χ' (say, χ) differs from $\mathbf{1}$, choose $g_{\mathbf{h}'} = e$ and $g_{\mathbf{h}}$ such that $\chi(g_{\mathbf{h}}) \neq 1$. This violates the character condition and implies $a_{\mathbf{h},\mathbf{h}',\chi,\chi'} = 0$. In summary, we have proven the following:

Lemma 10. *The invariant algebra consists exactly of the block matrices of the form*

$$\mathcal{A}_{\text{inv}} = \left\{ A_{\text{phys}} \oplus \bigoplus_{\mathbf{h}\in\mathcal{G}^{N-1}} \bigoplus_{\chi\neq\mathbf{1}} a_{\mathbf{h};\chi} |\mathbf{h};\chi\rangle\langle\mathbf{h};\chi| \right\},$$

where $A_{\text{phys}} \in \mathcal{A}_{\text{phys}}$ is supported on the relational subspace $\mathcal{H}_{\text{phys}}$ defined in Eq. (5), and the $a_{\mathbf{h};\chi}$ are complex numbers.

A few words are in place regarding the physical interpretation of these observables. Due to Eq. (3), χ labels the irreps of the global translations on state space. As already mentioned in Example 8, they can thus be interpreted as a discrete analog of (an exponentiated version of) *total momentum*. We can hence interpret the operator $|\mathbf{h};\chi\rangle\langle\mathbf{h};\chi|$ as describing a projective measurement that asks *whether the relation between the particles is* $\mathbf{h}$, *and whether the total momentum corresponds to* χ. Since this operator is contained in $\mathcal{A}_{\text{inv}}$, this measurement can be performed by an observer without access to the external reference frame. In the special case if $\chi = \mathbf{1}$, i.e. on the relational subspace $\mathcal{H}_{\text{phys}}$ which corresponds to "total momentum zero", such an observer can also perform measurements that correspond to superpositions of different particle relations $\mathbf{h}$. However, for "non-zero total momentum" ($\chi \neq \mathbf{1}$), we obtain an emergent superselection rule that forbids such superpositions and the corresponding measurements.

The reader familiar with constraint quantization will notice that the invariant observables A_{phys} on the subspace $\mathcal{H}_{\text{phys}}$ are the finite group analog of so-called Dirac observables [18, 19, 60]. Given some continuous group that is generated by an algebra of constraints, Dirac observables are operators that commute with the constraint operators (up to terms proportional to the constraints themselves). As such, they are invariant under the group generated by the constraints and observables on solutions to the constraints, i.e. on the so-called physical Hilbert space.

There is, however, a subtlety in this analogy: usually the (continuous) group generated by the constraints would be the analog of the 'classical' group $\mathcal{G}$ given here which is a strict subgroup of $\mathcal{G}_{\text{sym}}$. Thus, it is natural to ask whether the $\mathcal{G}_{\text{sym}}$-invariant subspace $\mathcal{H}_{\text{phys}}$ is a strict subset of the subspace of $\mathcal{G}$-invariant states. Accordingly, one may wonder whether the entire invariant algebra $\mathcal{A}_{\text{inv}}$ defined in terms of invariance under the larger group $\mathcal{G}_{\text{sym}}$ in Definition 7 is a strict subset of the algebra that is invariant under the smaller group $\mathcal{G}$. It is this latter algebra which thus gives rise to the actual analog of Dirac observables for the finite groups considered here. We will address these questions in the next subsection.

C. Group averaging states

Although we work with a representation of the larger group $\mathcal{G}_{\text{sym}}$, Eqs. (2)–(5) indicate that the total Hilbert space decomposes naturally in terms of the representation of the smaller group $\mathcal{G}$; e.g., the physical Hilbert space is also precisely the subspace invariant under $\mathcal{G}$. We will now clarify this observation by considering the corresponding (coherent) group averaging operations,

$$\Pi_{\text{phys}} := \frac{1}{|\mathcal{U}_{\text{sym}}|} \sum_{U\in\mathcal{U}_{\text{sym}}} U, \qquad \Pi'_{\text{phys}} := \frac{1}{|\mathcal{G}|} \sum_{g\in\mathcal{G}} U_g^{\otimes N}, \quad (6)$$

which are standard in constraint quantization [19, 66, 67], and for which the following holds.

Lemma 11. *The two coherent group averaging operations coincide, $\Pi_{\text{phys}} = \Pi'_{\text{phys}}$, and Π_{phys} is the orthogonal projector onto the relational subspace $\mathcal{H}_{\text{phys}}$.*

Proof. Direct calculation shows that $\Pi_{\text{phys}}^\dagger = \Pi_{\text{phys}}$ and $\Pi_{\text{phys}}'^\dagger = \Pi'_{\text{phys}}$, and that $\Pi'_{\text{phys}} = \Pi'^2_{\text{phys}}$ and $\Pi_{\text{phys}} = \Pi^2_{\text{phys}}$. Thus, Π_{phys} and Π'_{phys} are orthogonal projectors. Since $\mathcal{H}^{\otimes N}$ is spanned by the $|g,\mathbf{h}g\rangle$ for $g \in \mathcal{G}$ and $\mathbf{h} \in \mathcal{G}^{N-1}$, the image $\text{im}(\Pi'_{\text{phys}})$ of Π'_{phys} is spanned by

$$\Pi'_{\text{phys}} |g,\mathbf{h}g\rangle = \frac{1}{|\mathcal{G}|} \sum_{g'\in\mathcal{G}} U_{g'}^{\otimes N} |g,\mathbf{h}g\rangle = \frac{1}{\sqrt{|\mathcal{G}|}} |\mathbf{h};\mathbf{1}\rangle. \quad (7)$$

Since these states span $\mathcal{H}_{\text{phys}}$, this proves that Π'_{phys} is the orthogonal projector onto the physical subspace. By construction, every $|\psi\rangle \in \text{im}(\Pi_{\text{phys}})$ is invariant under every $U \in \mathcal{U}_{\text{sym}}$, and thus in particular under every $U_g^{\otimes N} \in \mathcal{U}_{\text{sym}}$. Thus, $\text{im}(\Pi_{\text{phys}}) \subseteq \mathcal{H}_{\text{phys}}$. On the other hand, decomposing $U \in \mathcal{U}_{\text{sym}}$ as in (2), we get

$$\Pi_{\text{phys}}|\mathbf{h};\mathbf{1}\rangle = \frac{1}{|\mathcal{U}_{\text{sym}}|} \sum_{U\in\mathcal{U}_{\text{sym}}} U_{g_{\mathbf{h}}}^{\otimes N} |\mathbf{h};\mathbf{1}\rangle = |\mathbf{h};\mathbf{1}\rangle$$

since $|\mathbf{h};\mathbf{1}\rangle$ is invariant under global translations. Thus, $\text{im}(\Pi_{\text{phys}}) \supseteq \mathcal{H}_{\text{phys}}$, and so $\Pi_{\text{phys}} = \Pi'_{\text{phys}}$. □

In conclusion, any basis state in $\mathcal{H}_{\mathbf{h}}$ projects to the same invariant subnormalized state $\Pi_{\text{phys}} |g,\mathbf{h}g\rangle = \Pi_{\text{phys}} |g',\mathbf{h}g'\rangle = \frac{1}{\sqrt{|\mathcal{G}|}} |\mathbf{h};\mathbf{1}\rangle$ under coherent group averaging, and it does not matter whether one averages with respect to the larger group $\mathcal{G}_{\text{sym}}$ or the smaller $\mathcal{G}$.

However, we will now see that the set of invariant observables, i.e. the observables resulting from *incoherent* group averaging ($\mathcal{G}$-twirling), differs for the two choices, but only outside of the relational subspace $\mathcal{H}_{\text{phys}}$. These operations are defined by

$$\Pi_{\text{inv}}(\rho) := \frac{1}{|\mathcal{U}_{\text{sym}}|} \sum_{U\in\mathcal{U}_{\text{sym}}} U\rho U^\dagger,$$

$$\Pi'_{\text{inv}}(\rho) := \frac{1}{|\mathcal{G}|} \sum_{g\in\mathcal{G}} U_g^{\otimes N} \rho (U_g^{\otimes N})^\dagger.$$

It is well-known, and easy to check by direct calculation, that these maps are projectors, i.e. $\Pi_{\text{inv}}^2 = \Pi_{\text{inv}}$ and $\Pi_{\text{inv}}'^2 = \Pi_{\text{inv}}'$, and that they are orthogonal with respect to the Hilbert-Schmidt inner product, i.e.

$$\text{tr}\left(A^\dagger \Pi_{\text{inv}}(B)\right) = \text{tr}\left(\Pi_{\text{inv}}(A)^\dagger B\right) \text{ for all } A, B \in \mathcal{L}(\mathcal{H}^{\otimes N}).$$

If $A \in \mathcal{L}(\mathcal{H}^{\otimes N})$ satisfies $[U, A] = 0$ for all $U \in \mathcal{U}_{\text{sym}}$, then $\Pi_{\text{inv}}(A) = A$. Conversely, if $B \in \text{im}(\Pi_{\text{inv}})$, then $UBU^\dagger = B$, i.e. $[U, B] = 0$, for all $U \in \mathcal{U}_{\text{sym}}$. Thus, Π_{inv} projects into the invariant algebra $\mathcal{A}_{\text{inv}}$. Similarly, Π_{inv}' projects into

$$\mathcal{A}_{\text{inv}}' = \{A \in \mathcal{L}(\mathcal{H}^{\otimes N}) \mid [U_g^{\otimes N}, A] = 0 \text{ for all } g \in \mathcal{G}\}.$$

Clearly, $\mathcal{A}_{\text{inv}} \subseteq \mathcal{A}_{\text{inv}}'$, but are these algebras equal? The following lemma collects the above insights, and answers this question in the negative.

Theorem 12. *Π_{inv} is the orthogonal projector onto the invariant subalgebra $\mathcal{A}_{\text{inv}}$. It can also be written in the form*

$$\Pi_{\text{inv}}(\rho) = \Pi_{\text{phys}} \rho \Pi_{\text{phys}} + \sum_{\mathbf{h}, \chi \neq \mathbf{1}} \langle \mathbf{h}; \chi | \rho | \mathbf{h}; \chi \rangle |\mathbf{h}; \chi\rangle\langle \mathbf{h}; \chi|.$$

Similarly, Π_{inv}' is the orthogonal projector onto the strictly larger subalgebra

$$\mathcal{A}_{\text{inv}}' = \left\{ \bigoplus_{\chi \in \widehat{\mathcal{G}}} A_\chi \right\} = \left\{ A_{\text{phys}} \oplus \bigoplus_{\chi \neq \mathbf{1}} A_\chi \right\},$$

where every A_χ is an arbitrary operator supported on the subspace $\mathcal{H}_\chi := \text{span}\{|\mathbf{h}; \chi\rangle \mid \mathbf{h} \in \mathcal{G}^{N-1}\}$ (note that $\mathcal{H}_{\mathbf{1}} = \mathcal{H}_{\text{phys}}$, so $A_{\text{phys}} = A_{\mathbf{1}}$).

Proof. To see the claimed form of Π_{inv}, note that the combination of projections is a Hilbert-Schmidt-orthogonal projection with image $\mathcal{A}_{\text{inv}}$. It remains to be shown that $\mathcal{A}_{\text{inv}}'$ has the claimed form. Note that $g \mapsto U_g^{\otimes N}$ is a representation of the finite Abelian group $\mathcal{G}$. It thus decomposes into one-dimensional irreps, and the equivalence classes of irreps are labelled by the characters χ. Thus, the form of $\mathcal{A}_{\text{inv}}'$ follows again from Schur's lemma. □

This theorem has interesting implications for the physical properties of $\mathcal{G}$-systems S. Recall our initial scenario as depicted in Figure 3. Suppose that we *only* demand symmetry of S with respect to ordinary, *unconditional* global translations $U_g^{\otimes N}$, and ask which observables can be measured by an observer without access to the external relatum. The answer is: all observables in $\mathcal{A}_{\text{inv}}'$. On the other hand, if we demand symmetry with respect to all *conditional* global translations in $\mathcal{U}_{\text{sym}}$ — and we will soon see that the QRF transformations of Refs. [38–40] are among those — then this turns out to be a more stringent requirement. In this case, fewer observables are measurable, namely only those in $\mathcal{A}_{\text{inv}}$.

In this sense, QRF transformations have fewer frame-independent observables than classical transformations: if all QRF transformations are symmetries, then superpositions of different particle relations $\mathbf{h}$ are forbidden by an emergent superselection rule whenever the "total momentum is non-zero", i.e. $\chi \neq \mathbf{1}$. On the other hand, if we only demand that global classical translations are symmetries, then these superpositions remain allowed.

Let us return to the analogy with contraint quantization discussed in the previous subsection. Lemma 11 and Theorem 12 show that an observable A_{phys}, i.e. the analog of a Dirac observable in our context, does *not* depend on whether it is constructed relative to $\mathcal{G}_{\text{sym}}$ or its subgroup $\mathcal{G}$. Later, we will also see that $\mathcal{A}_{\text{phys}}$ is the finite group analog of the algebra generated by so-called *relational* Dirac observables. These are invariant observables that encode relations between the subsystems, and they are common use in canonical quantum gravity [18–26, 42–45]. This explains why we have called the observables in $\mathcal{A}_{\text{phys}}$ "relational observables". They will become crucial in the resolution of the paradox of the third particle in Section IV, and they will turn out to be tomographically complete for the QRF states which we introduce in the next subsection.

Recall the notion of equivalence classes $[\rho]$ from Section II. We are now ready to introduce this notion formally for $\mathcal{G}$-systems:

Definition 13. *We call two quantum states $\rho, \sigma \in \mathcal{S}(\mathcal{H}^{\otimes N})$* symmetry-equivalent, *and write $\rho \simeq \sigma$, if there exists some symmetry $U \in \mathcal{U}_{\text{sym}}$ such that $\sigma = U\rho U^\dagger$. We call them* observationally equivalent, *and write $\rho \sim \sigma$, if $\text{tr}(A\rho) = \text{tr}(A\sigma)$ for all invariant observables (and thus all operators) $A \in \mathcal{A}_{\text{inv}}$.*

Clearly, if $\rho \simeq \sigma$ then $\rho \sim \sigma$, but the converse is not in general true. The equivalence class $[\rho]$ from Section II can now be defined as $[\rho] = \{\sigma \mid \sigma \simeq \rho\}$. In the case of pure state vectors $|\psi\rangle, |\psi'\rangle$, we must allow global phases and write $\psi \simeq \psi'$ if and only if there is some $U \in \mathcal{U}_{\text{sym}}^*$ such that $|\psi'\rangle = U|\psi\rangle$.

Observational equivalence can be characterized in terms of the projection into the invariant subalgebra:

Lemma 14. *Two states ρ and σ are observationally equivalent, i.e. $\rho \sim \sigma$, if and only if*

$$\Pi_{\text{inv}}(\rho) = \Pi_{\text{inv}}(\sigma).$$

Proof. This follows from the chain of equivalences

$$\begin{aligned}
\rho \sim \sigma &\Leftrightarrow \langle A, \rho\rangle_{\text{HS}} = \langle A, \sigma\rangle_{\text{HS}} \quad \forall A \in \mathcal{A}_{\text{inv}} \\
&\Leftrightarrow \langle \Pi_{\text{inv}}(B), \rho\rangle_{\text{HS}} = \langle \Pi_{\text{inv}}(B), \sigma\rangle_{\text{HS}} \quad \forall B \in \mathcal{L}(\mathcal{H}^{\otimes N}) \\
&\Leftrightarrow \langle B, \Pi_{\text{inv}}(\rho)\rangle_{\text{HS}} = \langle B, \Pi_{\text{inv}}(\sigma)\rangle_{\text{HS}} \quad \forall B \in \mathcal{L}(\mathcal{H}^{\otimes N}) \\
&\Leftrightarrow \Pi_{\text{inv}}(\rho) = \Pi_{\text{inv}}(\sigma).
\end{aligned}$$

□

D. Alignable states as states with a canonical representation

With these technical insights at hand, we are ready to return to the discussion of Section II. In the structural

approach to QRFs, we ask whether a given state has a natural representation, depending only on internal data, such that the communication task of Figure 5 can be successfully accomplished. In the following, let us focus on pure states $|\psi\rangle \in \mathcal{H}^{\otimes N}$ for simplicity. Our task is to find another state $|\psi'\rangle \in \mathcal{H}^{\otimes N}$ that is symmetry-equivalent to $|\psi\rangle$ and that is in some sense distinguished, i.e. yields a "canonical choice" for describing the set of symmetry-equivalent states, cf. Section II B 2.

In general, there may be many different possible ways to define such a "canonical choice". Let us pick one possible choice. Suppose that we fix one of the particles, say, particle i, where $1 \leq i \leq N$. Can we set the external reference frame such that this particle ends up at the "origin" — the unit element of the group? In other words, can we align our state "relative to particle i"? This is certainly possible in classical mechanics of N point particles in one dimension, given translation-invariance. Classically, it would indeed define us a unique representation. We will now see that a similar construction can be done for $\mathcal{G}$-systems, and that it leads to the notion of QRF of Refs. [38, 40].

Definition 15. *Let $i \in \{1, 2, \ldots, N\}$. A pure state $|\psi\rangle \in \mathcal{H}^{\otimes N}$ is called* i-alignable *if there exists some state $|\psi'\rangle \in \mathcal{H}^{\otimes N}$ with $\psi \simeq \psi'$ such that*

$$|\psi'\rangle \equiv |\psi'\rangle_{1,\ldots,N} = |e\rangle_i \otimes |\varphi\rangle_{1,\ldots,i-1,i+1,\ldots,N}.$$

In the following, we will also use the notation $|\varphi\rangle_{\overline{i}}$ for the vector $|\varphi\rangle_{1,\ldots,i-1,i+1,\ldots,N}$.

The state $|\varphi\rangle_{\overline{i}}$ in Definition 15 is exactly what is interpreted in Refs. [38–40] as "the state of the remaining $N-1$ particles as seen by particle i". Similarly, we will thus interpret $|e\rangle_i \otimes |\varphi\rangle_{\overline{i}}$ as the description of the N particle system relative to the QRF 'perspective' defined by particle i (which defines the origin).

Not all pure states are i-alignable. For example, the relational state $|\mathbf{h}; \mathbf{1}\rangle$ is an element of the subspace $\mathcal{H}_{\text{phys}}$, hence every $U \in \mathcal{U}_{\text{sym}}^*$ satisfies $U|\mathbf{h}; \mathbf{1}\rangle = e^{i\theta}|\mathbf{h}; \mathbf{1}\rangle$ for some $\theta \in \mathbb{R}$. Thus this state cannot be i-alignable for any i. While devoid of alignable states, we will see later that $\mathcal{H}_{\text{phys}}$ contains the complete relational information about all alignable states.

To analyze this notion further, the following lemma will be useful. Its proof is very simple and thus omitted.

Lemma 16. *For every $i \in \{1, \ldots, N\}$ and $\mathbf{h} \in \mathcal{G}^{N-1}$, there is a unique basis vector $|\mathbf{g}\rangle \in \mathcal{H}_{\mathbf{h}}$ with $g_i = g$, namely $|h_{i-1}^{-1} g, \mathbf{h} h_{i-1}^{-1} g\rangle$, with $h_0 := e$ the unit element of $\mathcal{G}$.*

This allows us to show that the state $|\psi'\rangle$ in Definition 15 is unique, and thus defines indeed a natural representation of the symmetry-equivalence class of $|\psi\rangle$:

Lemma 17. *If $|\psi\rangle \in \mathcal{H}^{\otimes N}$ is i-alignable, then the state $|\varphi\rangle_{\overline{i}}$ in Definition 15 is unique up to a global phase.*

Proof. Suppose that both ψ' and $\widetilde{\psi}'$ are states that satisfy the conditions of Definition 15. In particular, this means that $\psi \simeq \psi'$ and $\psi \simeq \widetilde{\psi}'$, and so there exists some $U \in \mathcal{U}_{\text{sym}}^*$ such that $|\widetilde{\psi}'\rangle = U|\psi'\rangle$, and $U = e^{i\theta} V$ for $V \in \mathcal{U}_{\text{sym}}$. In $|\psi'\rangle = |e\rangle_i \otimes |\varphi\rangle_{\overline{i}}$ and $|\widetilde{\psi}'\rangle = |e\rangle_i \otimes |\widetilde{\varphi}\rangle_{\overline{i}}$, we decompose φ and $\widetilde{\varphi}$ into product basis vectors: $|\varphi\rangle = \sum_{\mathbf{g} \in \mathcal{G}^{N-1}} \alpha_{\mathbf{g}} |g_1, \ldots, g_{N-1}\rangle$, and similarly for $|\widetilde{\varphi}\rangle$ with amplitudes $\widetilde{\alpha}_{\mathbf{g}}$. This implies that

$$\begin{aligned} U \sum_{\mathbf{g} \in \mathcal{G}^{N-1}} \alpha_{\mathbf{g}} |g_1, \ldots, g_{i-1}, e, g_i, \ldots, g_{N-1}\rangle \\ = \sum_{\mathbf{g} \in \mathcal{G}^{N-1}} \widetilde{\alpha}_{\mathbf{g}} |g_1, \ldots, g_{i-1}, e, g_i, \ldots, g_{N-1}\rangle. \end{aligned}$$

Now, according to Lemma 16, both of the decompositions $\sum_{\mathbf{g} \in \mathcal{G}^{N-1}} \ldots$ contain *at most one basis vector from every subspace* $\mathcal{H}_{\mathbf{h}}$ with non-zero amplitude, namely $|h_{i-1}^{-1}, \mathbf{h} h_{i-1}^{-1}\rangle$. But since U (and thus V) leaves the subspaces $\mathcal{H}_{\mathbf{h}}$ invariant, the last equation implies that $V|h_{i-1}^{-1}, \mathbf{h} h_{i-1}^{-1}\rangle = |h_{i-1}^{-1}, \mathbf{h} h_{i-1}^{-1}\rangle$ for every such vector that appears with non-zero amplitude $\alpha_{\mathbf{g}} \neq 0$ (and thus $\widetilde{\alpha}_{\mathbf{g}} \neq 0$). Hence $\widetilde{\alpha}_{\mathbf{g}} = e^{-i\theta} \alpha_{\mathbf{g}}$, and so $\widetilde{\varphi} = e^{-i\theta} \varphi$. □

E. QRF transformations as symmetry group elements

In the structural approach in Refs. [38–40], we can "jump" from one particle's reference frame into any other's. How is this idea expressed in our formalism? To see this, let us first show the following.

Theorem 18 (QRF state transformations)**.** *If there is some $i \in \{1, 2, \ldots, N\}$ such that $|\psi\rangle$ is i-alignable, then $|\psi\rangle$ is j-alignable for every $j \in \{1, 2, \ldots, N\}$. We will then simply call $|\psi\rangle$* alignable*. Moreover, for every $i, j \in \{1, \ldots, N\}$, there is a unique symmetry transformation $U \in \mathcal{U}_{\text{sym}}$ such that $U(|e\rangle_i \otimes |\varphi\rangle_{\overline{i}}) = |e\rangle_j \otimes |\varphi\rangle_{\overline{j}}$ for all $|\varphi\rangle_{\overline{i}}$. Furthermore, if $i \neq j$ then U is a proper* conditional *global translation, i.e. $U \bullet U^\dagger \in \mathcal{G}_{\text{sym}} \setminus \mathcal{G}$. Every such U induces a* unique *unitary ("QRF transformation") $V_{i \to j}$ such that $V_{i \to j} |\varphi\rangle_{\overline{i}} = |\varphi\rangle_{\overline{j}}$. This transformation can be written*

$$V_{i \to j} = \mathbb{F}_{i,j} \sum_{g \in \mathcal{G}} |g^{-1}\rangle\langle g|_j \otimes U_{g^{-1}}^{\otimes (N-2)},$$

where $\mathbb{F}_{i,j}$ flips (swaps) particles i and j. This is the discrete version of the form given in Refs. [38, 40].

Proof. Fix $i, j \in \{1, \ldots, N\}$. For every $\mathbf{h} \in \mathcal{G}^{N-1}$, let $g_{\mathbf{h}} := h_{j-1}^{-1} h_{i-1}$ (setting, as before, $h_0 := e$). Then the global translation by $g_{\mathbf{h}}$ satisfies

$$U_{g_{\mathbf{h}}}^{\otimes N} |h_{i-1}^{-1}, \mathbf{h} h_{i-1}^{-1}\rangle = |h_{j-1}^{-1}, \mathbf{h} h_{j-1}^{-1}\rangle.$$

Set $U := \bigoplus_{\mathbf{h} \in \mathcal{G}^{N-1}} U_{g_{\mathbf{h}}}^{\otimes N}$, then $U \in \mathcal{U}_{\text{sym}}$. According to Lemma 16, for every $\mathbf{h}$, U maps the unique basis vector $|\mathbf{g}\rangle \in \mathcal{H}_{\mathbf{h}}$ with $g_i = e$ to the unique basis vector $|\mathbf{g}'\rangle \in \mathcal{H}_{\mathbf{h}}$ with $g'_j = e$, and it is clear that U is the only symmetry transformation that does this. Thus, U maps all states of the form $|e\rangle_i \otimes |\varphi\rangle_{\overline{i}}$ to states of the form $|e\rangle_j \otimes |\widetilde{\varphi}\rangle_{\overline{j}}$.

Furthermore, if $i \neq j$ then there exist $\mathbf{h}, \mathbf{j}$ such that $g_{\mathbf{h}} \neq g_{\mathbf{j}}$. Thus, any such U is an $\mathbf{h}$-dependent transformation and thus cannot be a global translation.

Fix an arbitrary orthonormal basis $\{|\varphi\rangle_{\overline{i}}\}_\varphi$ of $\mathcal{H}^{\otimes(N-1)}$, then $U\left(|e\rangle_i \otimes |\varphi\rangle_{\overline{i}}\right) = |e\rangle_j \otimes |\varphi\rangle_{\overline{j}}$ yields another orthonormal basis. Thus, we can view this as a unitary $V_{i\to j}$ from $\mathcal{H}^{\otimes(N-1)}$ into another copy of $\mathcal{H}^{\otimes(N-1)}$. Since its action on basis vectors is fixed, there can be no more than one such map. To determine that it has the form as claimed, simply look at its action on the basis vectors $|g_1, \ldots, g_{N-1}\rangle$. □

So indeed, for every alignable state, and any particle $j \in \{1, \ldots, N\}$, there is a unique representation of that state "relative to the jth particle". Furthermore, the QRF transformation from i's to j's 'perspective' at the level of the *full* Hilbert space $\mathcal{H}^{\otimes N}$ corresponds to a symmetry transformation which lies in $\mathcal{G}_{\text{sym}}$, but *not* in $\mathcal{G}$ (if $i \neq j$). This observation highlights the physical significance of the symmetry group $\mathcal{G}_{\text{sym}}$. While we have seen that the set of invariant states $\mathcal{H}_{\text{phys}}$ is independent of whether one constructs it through coherently averaging over $\mathcal{G}_{\text{sym}}$ or its 'classical translation subgroup' $\mathcal{G}$, the symmetry group $\mathcal{G}_{\text{sym}}$ is key for understanding the meaning of the QRF transformations (which transform non-invariant descriptions): they are *conditional* symmetry transformations that depend on the interparticle relation $\mathbf{h}$.

To clarify the notation used in the definition of the QRF transformation $V_{i\to j}$, we give a simple example.

Example 19. *Suppose we have $N = 4$ particles, and an alignable state $|\psi\rangle$ such that*

$$|\psi\rangle \simeq |e\rangle_2 \otimes |g_1, g_3, g_4\rangle.$$

Thus, the state relative to the second particle is $|g_1, g_3, g_4\rangle$. To determine the state relative to the third particle, compute

$$\begin{aligned} V_{2\to 3}|g_1, g_3, g_4\rangle &= \mathbb{F}_{2,3} \sum_{g\in\mathcal{G}} |g^{-1}\rangle\langle g|_3 \otimes U_{g^{-1}}^{\otimes 2} |g_1, g_3, g_4\rangle \\ &= \mathbb{F}_{2,3} |g_3^{-1}\rangle\langle g_3| \otimes U_{g_3^{-1}}^{\otimes 2} |g_1, g_3, g_4\rangle \\ &= \mathbb{F}_{2,3} |\underbrace{g_3^{-1} g_1}_{1}, \underbrace{g_3^{-1}}_{3}, \underbrace{g_3^{-1} g_4}_{4}\rangle \\ &= |\underbrace{g_3^{-1} g_1}_{1}, \underbrace{g_3^{-1}}_{2}, \underbrace{g_3^{-1} g_4}_{4}\rangle, \end{aligned}$$

where the integers at the bottom denote the particle labels.

In Ref. [62], we will demonstrate equivalence of the above QRF transformations with the "quantum coordinate changes" of the perspective-neutral approach [39, 41–45].

As Theorem 18 has shown, the symmetry group $\mathcal{U}_{\text{sym}}$ contains the QRF transformations which switch from state descriptions relative to particle i to descriptions relative to particle j. But these do not exhaust the symmetry group. The following lemma gives another physically motivated example.

Example 20 (Center of mass)**.** *Consider again the cyclic group $\mathbb{Z}_n$ with addition modulo n, as shown in Figure 1. Let $m_1, \ldots, m_N$ be non-negative real numbers and $m := m_1 + \ldots + m_N$. For $\mathbf{h} \in \mathcal{G}^{N-1}$, define the group elements*

$$g(\mathbf{h}) := -\left\lfloor \frac{1}{m}\left(m_2 h_1 + \ldots + m_N h_{N-1}\right)\right\rfloor,$$

and set $U := \bigoplus_{\mathbf{h}\in\mathcal{G}^{N-1}} U_{g(\mathbf{h})}^{\otimes N}$. If we interpret the m_i as the masses *of the particles, with the origin as the position of particle 1 (i.e. $h_0 = 0$), then this symmetry transformation U describes a change of quantum coordinates such that the (integer part of) the "center of mass" becomes the origin.*

F. Characterization of alignable states

We have already seen that not all global states are alignable. The following lemma characterizes those that are.

Lemma 21. *A state $|\psi\rangle \in \mathcal{H}^{\otimes N}$ is alignable if and only if it can be written in the form*

$$|\psi\rangle = \sum_{\mathbf{h}\in\mathcal{G}^{N-1}} \alpha_{\mathbf{h}} |g_{\mathbf{h}}, \mathbf{h}\, g_{\mathbf{h}}\rangle \tag{8}$$

for some $g_{\mathbf{h}} \in \mathcal{G}$ and $\alpha_{\mathbf{h}} \in \mathbb{C}$. Moreover, the $\alpha_{\mathbf{h}}$ characterize the alignable state up to symmetry-equivalence. That is, for two alignable states $|\psi\rangle$ and $|\psi'\rangle$, we have $\psi \simeq \psi'$ if and only if their coefficients satisfy $\alpha_{\mathbf{h}} = e^{i\theta}\alpha'_{\mathbf{h}}$ for some $\theta \in \mathbb{R}$.

Proof. If $|\psi\rangle$ is alignable, then it is in particular 1-alignable. That is, there exists $U \in \mathcal{U}_{\text{sym}}^*$ such that

$$|\psi\rangle = U\left(|e\rangle_1 \otimes |\varphi\rangle_{\overline{1}}\right) = e^{i\theta} \sum_{\mathbf{h}\in\mathcal{G}^{N-1}} \alpha_{\mathbf{h}} U_{g_{\mathbf{h}}}^{\otimes N} |e, \mathbf{h}\rangle,$$

where the $\alpha_{\mathbf{h}}$ are the coefficients of $|\varphi\rangle_{\overline{1}}$ in terms of the product group basis, and the $U_{g_{\mathbf{h}}}^{\otimes N}$ translate the basis vectors of the subspaces $\mathcal{H}_{\mathbf{h}}$. Hence we get the claimed form for $|\psi\rangle$. The converse direction of the proof of the first part of this lemma is analogous. We also see that the $\alpha_{\mathbf{h}}$ only depend on the symmetry-equivalence class of the state $|\psi\rangle$ (up to a global phase); and, in the case of equality of those coefficients, two states must have the same $|\varphi\rangle_{\overline{1}}$ (up to a global phase), hence they must be symmetry-equivalent. □

The above form shows that any maximal subspace of $\mathcal{H}^{\otimes N}$ which is contained in the set of alignable states has dimension $|\mathcal{G}|^{N-1}$, i.e. is isomorphic to $\mathcal{H}^{\otimes(N-1)}$. This shows that the QRFs as defined above are indeed "good" QRFs as explained in Subsection II B 2. Due to Lemma 9, it also shows that the relational subspace $\mathcal{H}_{\text{phys}}$ has the right dimension for its states to contain "all the particles' internal QRF perspectives at once". This observation is corroborated by the following useful Lemma and will be further discussed in Ref. [62].

Lemma 22. *Given any alignable state $|\psi\rangle$ (which is hence of the form (8)), its projection onto the invariant subalgebra is*

$$\Pi_{\text{inv}}(|\psi\rangle\langle\psi|) = \sum_{\mathbf{h},\mathbf{j}\in\mathcal{G}^{N-1}} \frac{\alpha_{\mathbf{h}}\overline{\alpha}_{\mathbf{j}}}{|\mathcal{G}|}|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}| + \sum_{\mathbf{h}\in\mathcal{G}^{N-1}} \frac{|\alpha_{\mathbf{h}}|^2}{|\mathcal{G}|}\Pi_{\mathbf{h};\chi\neq\mathbf{1}},$$

where $\Pi_{\mathbf{h};\chi\neq\mathbf{1}} := \sum_{\chi\neq\mathbf{1}} |\mathbf{h};\chi\rangle\langle\mathbf{h};\chi|$. Thus, for any two alignable states $|\psi\rangle, |\psi'\rangle$, we have $\psi \sim \psi'$ if and only if $\psi \simeq \psi'$: such states are symmetry-equivalent if and only if they are observationally equivalent. Furthermore, they are equivalent if and only if $\langle\psi|A|\psi\rangle = \langle\psi'|A|\psi'\rangle$ for all $A \in \mathcal{A}_{\text{phys}}$; that is, all invariant information of alignable states is fully contained in their projection $\Pi_{\text{phys}}|\psi\rangle$.

Proof. Using (7), we obtain

$$\Pi_{\text{phys}}|\psi\rangle = \frac{1}{\sqrt{|\mathcal{G}|}} \sum_{\mathbf{h}\in\mathcal{G}^{N-1}} \alpha_{\mathbf{h}}|\mathbf{h};\mathbf{1}\rangle.$$

Direct calculation shows that $|\langle\psi|\mathbf{h};\chi\rangle|^2 = |\alpha_{\mathbf{h}}|^2/|\mathcal{G}|$. The result then follows by using the form of Π_{inv} as given in Theorem 12. Both notions of equivalence boil down to the fact that the states have the same amplitudes $\alpha_{\mathbf{h}}$ up to a global phase. Finally, for ρ any alignable state, the above form of $\Pi_{\text{inv}}(\rho)$ implies

$$\text{tr}(\rho A) = \text{tr}(\rho_{\text{phys}}A) + \sum_{\mathbf{h}\in\mathcal{G}^{N-1}} \langle\mathbf{h};\mathbf{1}|\rho_{\text{phys}}|\mathbf{h};\mathbf{1}\rangle \text{tr}(\Pi_{\mathbf{h};\chi\neq\mathbf{1}}A) \tag{9}$$

for all $A \in \mathcal{A}_{\text{inv}}$, where $\rho_{\text{phys}} := \Pi_{\text{phys}}\rho\Pi_{\text{phys}}$. That is, the expectation values of all invariant observables, and thus the notion of observational equivalence, depends only on the state's projection into $\mathcal{A}_{\text{phys}}$. □

In other words, the algebra $\mathcal{A}_{\text{phys}}$ is tomographically complete for the invariant information in alignable states. In this sense, it can be said that the external relatum independent states of the structural approach are what we called the relational states, namely the ones in $\mathcal{H}_{\text{phys}}$. That is, the external relatum independent states of the structural approach are the *coherently* group-averaged states, while their counterparts in the QI approach are the *incoherently* group-averaged states (cf. Subsection II B 1).[7]

G. Alignable and relational observables

Given the notion of alignable states and the duality between states and observables, it is natural to also define alignable observables in the obvious way.

[7] We thank A. R. H. Smith for suggesting us to emphasize this technical distinction.

Definition 23 (Alignable observables). *An operator $A \in \mathcal{L}(\mathcal{H}^{\otimes N})$ is called i-*alignable *for $i \in \{1,\dots,N\}$ if there exists $U \in \mathcal{U}_{\text{sym}}$ such that*

$$UAU^\dagger = |e\rangle\langle e|_i \otimes A_{\bar{i}}$$

for some $A_{\bar{i}} \in \mathcal{L}(\mathcal{H}^{\otimes(N-1)})$. If A is an observable, it is called an i-alignable observable.

This leads to the following extension of Theorem 18 from QRF state to observable transformations. The proof is analogous and thus omitted.

Theorem 24 (QRF observable transformations). *If $A \in \mathcal{L}(\mathcal{H}^{\otimes N})$ is i-alignable, then it is also j-alignable for every $j \in \{1,\dots,N\}$, and so we will call A simply* alignable. *In particular, there is a unique $U \in \mathcal{U}_{\text{sym}}$ such that*

$$U\left(|e\rangle\langle e|_i \otimes A_{\bar{i}}\right)U^\dagger = |e\rangle\langle e|_j \otimes A_{\bar{j}},$$

where $A_{\bar{j}} = V_{i\to j}\, A_{\bar{i}}\, V_{j\to i}$. Here, U is the unique symmetry transformation from Theorem 18 such that $U\left(|e\rangle_i \otimes |\varphi\rangle_{\bar{i}}\right) = |e\rangle_j \otimes |\varphi\rangle_{\bar{j}}$ for all $|\varphi\rangle_{\bar{i}}$ and $V_{i\to j}$ is the unitary "QRF transformation" induced by it.

These are the discrete versions of the observable transformations in Refs. [38, 39].

Finally, note that if $A_{\bar{1}} = \sum_{\mathbf{h},\mathbf{j}\in\mathcal{G}^{N-1}} a_{\mathbf{h},\mathbf{j}}|\mathbf{h}\rangle\langle\mathbf{j}|$, then

$$\widehat{\Pi}_{\text{phys}}\left(|e\rangle\langle e|_1 \otimes A_{\bar{1}}\right) = \sum_{\mathbf{h},\mathbf{j}\in\mathcal{G}^{N-1}} \frac{a_{\mathbf{h},\mathbf{j}}}{|\mathcal{G}|}|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|,$$

where $\widehat{\Pi}_{\text{phys}}$ henceforth denotes the superoperator that acts as $\widehat{\Pi}_{\text{phys}}(A) = \Pi_{\text{phys}}\, A\, \Pi_{\text{phys}}$, as obvious from Lemma 22. Up to a factor of $|\mathcal{G}|$, this is identical to the original representation of $A_{\bar{1}}$, but in another basis. This proves the following theorem, extending a result from [44, 45]:

Theorem 25. *Consider the map*

$$A_{\bar{i}} \mapsto F_{A_{\bar{i}},i} := |\mathcal{G}| \cdot \widehat{\Pi}_{\text{phys}}\left(|e\rangle\langle e|_i \otimes A_{\bar{i}}\right), \tag{10}$$

where $A_{\bar{i}} \in \mathcal{L}(\mathcal{H}^{\otimes(N-1)})$ is any operator. This defines an isomorphism between the operators $A_{\bar{i}}$ and $\mathcal{A}_{\text{phys}}$, preserving products, linear combinations, and adjoints.

This gives us two independent motivations to focus on $\mathcal{A}_{\text{phys}}$: for alignable states, the projection into this subalgebra contains all invariant information; and it does so in a way that preserves the natural structure of the alignable observables.

Furthermore, returning to the comparison with constraint quantization, it follows from Refs. [44, 45] that $F_{A_{\bar{i}},i}$ is (the finite group analog of) the *relational Dirac observable* which encodes in an invariant manner the question "what is the value of $A_{\bar{i}}$ given that particle i sits at the origin?" Such relational observables are a standard tool in canonical quantum gravity, e.g. see Refs. [18–26, 42, 43]. Theorem 25 is the reason why we refer to $\mathcal{A}_{\text{phys}}$ as the algebra generated by relational observables.

Due to $\mathcal{A}_{\text{phys}} \subset \mathcal{A}_{\text{inv}}$, we have

$$\widehat{\Pi}_{\text{phys}}\left(\Pi_{\text{inv}}(|e\rangle\langle e|_i \otimes A_{\bar{i}})\right) = \widehat{\Pi}_{\text{phys}}\left(|e\rangle\langle e|_i \otimes A_{\bar{i}}\right).$$

Since Π_{inv} is the incoherent $\mathcal{G}$-twirl over $\mathcal{U}_{\text{sym}}$, it is clear that the image $\Pi_{\text{inv}}(|e\rangle\langle e|_i \otimes A_{\bar{i}}) \in \mathcal{A}_{\text{inv}}$ only depends on the symmetry equivalence class of the alignable observable A. Hence, in particular we have

$$F_{A_{\bar{i}},i} = F_{A_{\bar{j}},j} \qquad \text{for all } i,j \in \{1,\dots,N\},$$

and so the relational observable in Eq. (10) does not depend on the choice of particle i.

This systematic equivalence of relational observables is once more made possible by studying the larger symmetry group $\mathcal{G}_{\text{sym}}$ rather than its subgroup $\mathcal{G}$ as usual in the literature. Indeed, Theorems 18 and 24 demonstrate that any two particle alignments of an observable are related by a unique symmetry group element which generically lies in $\mathcal{G}_{\text{sym}} \setminus \mathcal{G}$.

Lemma 22 (and its obvious generalization to alignable observables) yields an interesting insight that will become relevant in Section IV: if we look at the image of all alignable states and observables under Π_{inv}, then we do not obtain the full invariant algebra $\mathcal{A}_{\text{inv}}$. Instead, we always obtain an operator in the smaller subalgebra

$$\mathcal{A}_{\text{alg}} := \left\{ A_{\text{phys}} + \sum_{\mathbf{h}\in\mathcal{G}^{N-1}} a_{\mathbf{h}} \Pi_{\mathbf{h};\chi\neq\mathbf{1}} \right\},$$

where $A_{\text{phys}} \in \mathcal{A}_{\text{phys}}$, and $a_{\mathbf{h}} \in \mathbb{C}$ are arbitrary complex numbers. According to the definitions in Subsection III C, this gives us the subalgebra inclusions

$$\mathcal{A}_{\text{phys}} \subset \mathcal{A}_{\text{alg}} \subset \mathcal{A}_{\text{inv}} \subset \mathcal{A}'_{\text{inv}} \subset \mathcal{L}(\mathcal{H}^{\otimes N}).$$

Hence, if we denote the orthogonal projection into $\mathcal{A}_{\text{alg}}$ by Π_{alg}, we have $\Pi_{\text{alg}} \circ \Pi_{\text{inv}} = \Pi_{\text{inv}} \circ \Pi_{\text{alg}} = \Pi_{\text{alg}}$. More specifically, the following holds.

Lemma 26. *$\mathcal{A}_{\text{alg}}$ is the smallest subalgebra of $\mathcal{L}(\mathcal{H}^{\otimes N})$ that contains $\Pi_{\text{inv}}(|\psi\rangle\langle\psi|)$ for all alignable states $|\psi\rangle$.*

Proof. It is clear that the rank-one projectors $|\psi\rangle\langle\psi|$ for $|\psi\rangle = |e\rangle_1 \otimes |\varphi\rangle_{\bar{1}}$ linearly span all of $|e\rangle\langle e|_1 \otimes \mathcal{L}(\mathcal{H}^{\otimes(N-1)})$, thus we are looking for the subalgebra $\mathcal{A}$ that is generated by the image of these operators under Π_{inv}. Lemma 22 shows that $\Pi_{\text{inv}}(|\psi\rangle\langle\psi|)$ is contained in $\mathcal{A}_{\text{alg}}$ for every alignable state $|\psi\rangle$, hence $\mathcal{A} \subseteq \mathcal{A}_{\text{alg}}$. Conversely, if $A = \sum_{\mathbf{h},\mathbf{j}\in\mathcal{G}^{N-1}} a_{\mathbf{h},\mathbf{j}} |\mathbf{h}\rangle\langle\mathbf{j}|$, then

$$\Pi_{\text{inv}}(|e\rangle\langle e| \otimes A) = \sum_{\mathbf{h},\mathbf{j}\in\mathcal{G}^{N-1}} \frac{a_{\mathbf{h},\mathbf{j}}}{|\mathcal{G}|} |\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}| + \sum_{\mathbf{h}\in\mathcal{G}^{N-1}} \frac{a_{\mathbf{h},\mathbf{h}}}{|\mathcal{G}|} \Pi_{\mathbf{h};\chi\neq\mathbf{1}}.$$

Setting $A := |\mathbf{h}\rangle\langle\mathbf{j}|$ for $\mathbf{h} \neq \mathbf{j}$ shows that $|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}| \in \mathcal{A}$. But then, we also have $|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{h};\mathbf{1}| = |\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|(|\mathbf{j};\mathbf{1}\rangle\langle\mathbf{h};\mathbf{1}|) \in \mathcal{A}$, and so all operators A_{phys} fully supported on $\mathcal{H}_{\text{phys}}$ are in $\mathcal{A}$. Finally, considering the image of $A = |\mathbf{h};\mathbf{1}\rangle\langle\mathbf{h};\mathbf{1}|$ shows that $\Pi_{\mathbf{h};\chi\neq\mathbf{1}} \in \mathcal{A}$, and so $\mathcal{A} \supseteq \mathcal{A}_{\text{alg}}$. □

H. Communication scenario of the structural approach revisited

We are now in a position to revisit the communication scenario elucidating the operational essence of the structural approach to QRFs in section II B 2 (see also Figure 5). It is clear how Alice and Bob can win the game proposed by Refaella in the case of N particles on the configuration space $\mathcal{G}$ in the absence of a shared external relatum, given that the class of states $\mathcal{C}$ that they are interested in is the set of alignable states. For example, before the game begins, they can agree to always use particle 1 as the internal reference system relative to which the remaining particles are described. This yields the "canonical choice" to represent any equivalence class of (pure) states in the form $|e\rangle_1 \otimes |\varphi\rangle_{\bar{1}}$ and Alice's and Bob's return to Refaella will always agree.

Alice and Bob have, of course, the option to choose any of the N particles as a reference system, each likewise defining a "canonical choice". It is clear that all these different possible conventions by Alice and Bob are precisely related by the QRF transformations of Lemma 18 and that each such transformation is an element of the symmetry group $\mathcal{U}_{\text{sym}}$.

IV. THE 'PARADOX OF THE THIRD PARTICLE' AND THE RELATIONAL TRACE

Ref. [49] presents a "paradox of the third particle" in the context of QRFs. We will now formulate this apparent paradox in our formalism, and see that the structure of observables, as elaborated in Section III, helps to clarify its physical background and to resolve it in terms of relational observables. As we will see, the core of the problem is how to embed the two-particle observables into the set of three-particle observables, and the key will be to do so in a relational manner. This bears some resemblance to the issue of boundaries and edge modes in gauge theory and gravity [51–56], which is related to the question of how to embed the gauge-invariant observables of neighbouring subregions in spacetime into the set of gauge-invariant observables associated with the union ('gluing') of these subregions.

The setup of Ref. [49] consists of three particles in one dimension, i.e. on the real line $\mathbb{R}$. But since only a finite number of positions is relevant for the paradox, we can discretize space and its translations. Thus, we consider a $\mathcal{G}$-system with $\mathcal{G} = \mathbb{Z}_n$, the cyclic group of order n, as in Fig. 1. The group operation is addition modulo n; in the following, whenever we write $a+b$, we actually mean $a+b \mod n$. As described earlier, this is a discrete model of the translation group acting on the real line. We start with two particles that have been prepared (by an external observer with access to the reference frame) in the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|-a\rangle_1|b\rangle_2 + e^{i\theta}|a\rangle_1|-b\rangle_2\right), \tag{11}$$

where $a, b \in \{0, 1, 2, \ldots, n-1\}$, and $\theta \in \mathbb{R}$. This state is symmetry-equivalent to[8]

$$|\psi\rangle \simeq |0\rangle_1 \otimes \frac{1}{\sqrt{2}} \left(|a+b\rangle_2 + e^{i\theta}|-a-b\rangle_2\right). \tag{12}$$

In our terminology, this means that the state is alignable. In Ref. [49], this form is used as a motivation to declare: *"Therefore, we conclude that particle 1 sees particle 2 in a pure state. Importantly this implies that particle 1 can get access to the phase θ by interacting with particle 2 alone, i.e. without access to the external reference frame."* In our conceptual framework, we would rather describe the situation as follows: consider an external observer who has access to particles 1 and 2, but has no access to the external reference frame. There are some observables that this observer can measure for which the phase θ is relevant.

This is because the state that is effectively seen by this observer is the projection of ψ into the invariant subalgebra, which we can determine via Lemma 22. The coefficients of this state are $\alpha_h = 1/\sqrt{2}$ for $h = a + b$ and $\alpha_j = e^{i\theta}/\sqrt{2}$ for $j = -a - b$, thus

$$\begin{aligned}\Pi_{\text{inv}}(|\psi\rangle\langle\psi|) = &\frac{1}{2n}|h;\mathbf{1}\rangle\langle h;\mathbf{1}| + \frac{e^{-i\theta}}{2n}|h;\mathbf{1}\rangle\langle j;\mathbf{1}| \\ &+\frac{e^{i\theta}}{2n}|j;\mathbf{1}\rangle\langle h;\mathbf{1}| + \frac{1}{2n}|j;\mathbf{1}\rangle\langle j;\mathbf{1}| \\ &+\frac{1}{2n}\Pi_{h;\chi\neq\mathbf{1}} + \frac{1}{2n}\Pi_{j;\chi\neq\mathbf{1}}\end{aligned} \tag{13}$$

and this state depends on θ in a nontrivial way.

Now a third particle is introduced. From the external perspective, it is prepared in a pure state $|c\rangle$ independently of the other two particles, where $c \in \{0, 1, \ldots, n-1\}$. From that perspective, the global state thus reads

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|-a\rangle_1|b\rangle_2 + e^{i\theta}|a\rangle_1|-b\rangle_2\right)|c\rangle_3. \tag{14}$$

This state is still alignable. Relative to particle 1, it becomes

$$|\Psi'\rangle = |0\rangle_1 \frac{1}{\sqrt{2}} \left(|a+b\rangle_2|a+c\rangle_3 + e^{i\theta}|-a-b\rangle_2|-a+c\rangle_3\right) \tag{15}$$

with $\Psi' \simeq \Psi$. This is a state for which particles 2 and 3 are formally entangled. Now suppose that particle 3 is very far away, such that our external observer (or, as the authors of Ref. [49] would say, such that particle 1) has no access to particle 3. If one now formally takes the partial trace over particle 3, one obtains a reduced state of particles 1 and 2 that is *independent of the phase θ*. This seems to contradict our earlier claim — now it looks as if an external observer without access to the external reference frame or to particle 3 *cannot* see any observable consequences of the phase θ.

We arrive at an apparent paradox: computing the partial trace via Ψ or via Ψ' gives different predictions, even though both states are symmetry-equivalent. Moreover, the result from tracing out the third particle via Ψ' seems absurd, given the seemingly innocuous role that the third particle plays in state Ψ. **Can the phase θ be accessed by an observer without access to the external relatum and *with restricted access to only particles* 1 *and* 2?** It seems like there should be an objective answer to this question which does not depend on whether it is asked in the context of state Ψ or Ψ'.

Clearly, since the usual partial trace yields differing results for Ψ and Ψ', it cannot represent the correct rule to compute reduced states in the setting of QRFs. To shed light on the reason for why this is the case, let us reconsider how the standard partial trace can be motivated. Consider three distinguishable particles as one usually does in quantum information theory — in our setting, this implies that the state of the particles is defined relative to an accessible reference frame. Denote by A_{12} some observable that is measurable if one has access to particles 1 and 2 only (and to the reference frame). Then one can equivalently describe this as an observable on the three particles, such that the third particle is ignored. Formally, this can be done via a map

$$A_{12} \mapsto \Phi(A_{12}) = A_{12} \otimes \mathbf{1}_3. \tag{16}$$

That is, the 2-particle observables are naturally embedded into the 3-particle observables via some map Φ, which takes the tensor product with the identity observable. This map preserves all relevant structure (as it must): it takes linear combinations to linear combinations, products to products, the adjoint to the adjoint, and the identity to the identity. Formally, this is summarized by saying that Φ is a *unital *-homomorphism*. It defines what we mean when we talk about "observables pertaining only to particles 1 and 2" within the set of all 3-particle observables. In the following, we will refer to *-homomorphisms simply as *embeddings* (which may or may not be unital).

Now consider any quantum state ρ_{123} on the three particles. We would like to determine the state ρ_{12} that results if one has only access to particles 1 and 2. By this, we mean the state that gives the same expectation values as ρ_{123} on all local observables A_{12}. Thus, we demand

$$\text{tr}(\rho_{12}A_{12}) = \text{tr}(\rho_{123}\Phi(A_{12})) \quad \text{for all } A_{12}.$$

If we write this in terms of the Hilbert-Schmidt inner product, then

$$\langle \rho_{12}, A_{12}\rangle_{\text{HS}} = \langle \rho_{123}, \Phi(A_{12})\rangle_{\text{HS}} = \langle \Phi^\dagger(\rho_{123}), A_{12}\rangle_{\text{HS}}.$$

That is, we must have $\rho_{12} = \Phi^\dagger(\rho_{123})$, with $\Phi^\dagger$ the Hilbert-Schmidt adjoint of Φ. But given Eq. (16), the form of it is easy to see that $\Phi^\dagger = \text{Tr}_3$, and this recovers the partial trace.

[8] In fact, in the sequel we will assume that $a \neq 0$, for otherwise the states in Eqs. (11) and (12) coincide.

In the context of QRFs and $\mathcal{G}$-systems, we have a different structure of observables: what is measurable without access to the external reference frame corresponds to the *invariant observables*. We therefore need an analog of the above construction for the invariant subalgebra $\mathcal{A}_{\text{inv}}$. In fact, since we are only interested in alignable states, Lemma 26 tells us that we can focus on the subalgebra $\mathcal{A}_{\text{alg}}$.

A. The non-uniqueness of invariant embeddings

In light of the above considerations, let us try to construct an embedding of the relevant 2-particle observables (or more generally of $\mathcal{A}_{\text{alg}}^{(N)}$) into the 3-particle observables (more generally into $\mathcal{A}_{\text{alg}}^{(N+M)}$, where the superscript denotes the number of particles). To obtain some crucial physical intuition, we will now define one such embedding (called $\widetilde{\Phi}^{(1)}$) in an intuitive manner, before we turn to a more systematic treatment below.

We start our construction with the orthogonal projector $\Pi_{\mathbf{h}}^{(N)} := \sum_{\chi\in\widehat{\mathcal{G}}} |\mathbf{h};\chi\rangle\langle\mathbf{h};\chi|$. Its embedding must be on orthogonal projector in $\mathcal{A}_{\text{alg}}^{(N+M)}$; measuring this projector amounts to asking whether the first N particles have pairwise relations described by $\mathbf{h}$. Clearly, the answer must be "yes" whenever the $M+N$ particles are in some joint relation $(\mathbf{h},\mathbf{g})$ for arbitrary $\mathbf{g}\in\mathcal{G}^M$, and the answer must be "no" for all other relations $(\mathbf{h}',\mathbf{g})$ whenever $\mathbf{h}'\neq\mathbf{h}$. This suggests to choose the embedding as

$$\widetilde{\Phi}^{(1)}(\Pi_{\mathbf{h}}^{(N)}) = \sum_{\mathbf{g}\in\mathcal{G}^M} \Pi_{\mathbf{h},\mathbf{g}}^{(N+M)}. \tag{17}$$

The reason for the superscript '(1)' will become clear shortly. Now consider the orthogonal projector $\Pi_{\text{phys}}^{(N)} := \sum_{\mathbf{h}\in\mathcal{G}^{N-1}} |\mathbf{h};\mathbf{1}\rangle\langle\mathbf{h};\mathbf{1}|$. A state $|\psi\rangle$ is in the image of this projector if and only if it is translation-invariant, i.e. $U_g^{\otimes N}|\psi\rangle = |\psi\rangle$ for all $g\in\mathcal{G}$. Suppose that a state of $N+M$ particles is translation-invariant; then we would like to be able to say that the state of the first N particles is translation-invariant, too. This motivates us to demand $\widetilde{\Phi}^{(1)}(\Pi_{\text{phys}}^{(N)}) = \Pi_{\text{phys}}^{(N+M)}$: in other words, we embed the translation-invariant observables of N particles into the translation-invariant observables of $N+M$ particles.

We have $|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{h};\mathbf{1}| = \Pi_{\mathbf{h}}^{(N)}\Pi_{\text{phys}}^{(N)}$. Since $\widetilde{\Phi}^{(1)}$ is supposed to be an embedding and thus multiplicative, this implies

$$\widetilde{\Phi}^{(1)}(|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{h};\mathbf{1}|) = \sum_{\mathbf{g}\in\mathcal{G}^M} |\mathbf{h},\mathbf{g};\mathbf{1}\rangle\langle\mathbf{h};\mathbf{g};\mathbf{1}|. \tag{18}$$

To exhaust all of $\mathcal{A}_{\text{alg}}^{(N)}$, we still need to embed the operators $|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|$ for $\mathbf{h}\neq\mathbf{j}$. Given Eq. (18), it seems formally natural to define

$$\widetilde{\Phi}^{(1)}(|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|) := \sum_{\mathbf{g}\in\mathcal{G}^M} |\mathbf{h},\mathbf{g};\mathbf{1}\rangle\langle\mathbf{j},\mathbf{g};\mathbf{1}|. \tag{19}$$

Natural as this definition may seem, we will soon see that it relies on quite subtle physical assumptions. For now, let us work with this definition and explore its consequences. First, linearity and Eqs. (17) and (18) imply

$$\widetilde{\Phi}^{(1)}\left(\Pi_{\mathbf{h},\chi\neq\mathbf{1}}\right) = \sum_{\mathbf{g}\in\mathcal{G}^M} \Pi_{\mathbf{h},\mathbf{g};\chi\neq\mathbf{1}}. \tag{20}$$

These demands yield an embedding that can equivalently be defined as follows:

Lemma 27. *There is a unique unital embedding $\widetilde{\Phi}^{(1)}$ of $\mathcal{A}_{\text{alg}}^{(N)}$ into $\mathcal{A}_{\text{alg}}^{(N+M)}$ that satisfies*

$$\widetilde{\Phi}^{(1)}\left(\Pi_{\text{alg}}^{(N)}\left(\widehat{e}_1\otimes A_{\overline{1}}\right)\right) = \Pi_{\text{alg}}^{(N+M)}\left(\widehat{e}_1\otimes A_{\overline{1}}\otimes\mathbf{1}^{(M)}\right) \tag{21}$$

for all $A_{\overline{1}}\in\mathcal{L}(\mathcal{H}^{\otimes(N-1)})$, where $\widehat{e}_1 := |e\rangle\langle e|_1$ and $\overline{1} := \{2,3,\ldots,N\}$. It is given by the linear extension of Eqs. (19) and (20).

Proof. Write $A_{\overline{1}} = \sum_{\mathbf{h},\mathbf{j}\in\mathcal{G}^{N-1}} a_{\mathbf{h},\mathbf{j}}|\mathbf{h}\rangle\langle\mathbf{j}|$ and use Lemma 22 to obtain

$$\Pi_{\text{alg}}^{(N)}\left(|e\rangle\langle e|_1\otimes A_{\overline{1}}\right) = \sum_{\mathbf{h},\mathbf{j}\in\mathcal{G}^{N-1}} \frac{a_{\mathbf{h},\mathbf{j}}}{|\mathcal{G}|}|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}| + \sum_{\mathbf{h}\in\mathcal{G}^{N-1}} \frac{a_{\mathbf{h},\mathbf{h}}}{|\mathcal{G}|}\Pi_{\mathbf{h};\chi\neq\mathbf{1}}. \tag{22}$$

A similar representation can be obtained for the right-hand side of Eq. (21). Thus, it is clear that the linear extension of Eqs. (19) and (20) satisfies Eq. (21). In particular, the unit is preserved.

Now let $\widetilde{\Phi}$ be any embedding of $\mathcal{A}_{\text{alg}}^{(N)}$ into $\mathcal{A}_{\text{alg}}^{(N+M)}$ which satisfies Eq. (21). Choose $\mathbf{h},\mathbf{j}\in\mathcal{G}^{N-1}$ arbitrary, and let $A_{\overline{1}}$ be the operator with $a_{\mathbf{h},\mathbf{j}} = |\mathcal{G}|$ and all other coefficients zero. Then Eq. (22) becomes $|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|$, and $\Pi_{\text{alg}}^{(N+M)}\left(|e\rangle\langle e|_1\otimes A_{\overline{1}}\otimes\mathbf{1}^{(M)}\right)$ becomes the right-hand side of Eq. (19). Now, still for $\mathbf{h}\neq\mathbf{j}$,

$$\left(|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|\right)^\dagger\left(|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|\right) = |\mathbf{j};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|.$$

Since $\widetilde{\Phi}$ preserves adjoints and products, this proves Eq. (19) also in the case $\mathbf{h}=\mathbf{j}$. Finally, choosing $A_{\overline{1}}$ such that $a_{\mathbf{h},\mathbf{h}} = |\mathcal{G}|$ and all other coefficients zero proves Eq. (20), which implies that $\widetilde{\Phi}$ is the linear extension of Eqs. (19) and (20). □

What this lemma demonstrates is that our construction of $\widetilde{\Phi}^{(1)}$ can be interpreted in an alternative way — at least on "alignable" observables. What our embedding map does to those observables is as follows: *write them in the form $|e\rangle\langle e|_1\otimes A_{\overline{1}}$, and embed them into the total Hilbert space according to its defining tensor product structure. Then*

demand that $\widetilde{\Phi}^{(1)}$ *maps the invariant part of* $|e\rangle\langle e|_1 \otimes A_{\overline{1}}$ *to the invariant part of its embedding* $|e\rangle\langle e|_1 \otimes A_{\overline{1}} \otimes \mathbf{1}^{(M)}$. This defines a particular, natural embedding of $\mathcal{A}_{\text{alg}}^{(N)}$ into $\mathcal{A}_{\text{alg}}^{(N+M)}$.

But this suggests directly that our candidate relational trace is defective: *its definition is implicitly based on the choice of particle* 1 *as our reference.* Indeed, the following lemma shows that there is a large *class of invariant embeddings*. In particular, choosing another particle as the reference particle will in general lead to inequivalent embeddings. The proof is given by a straightforward calculation and thus omitted.

Lemma 28. *Let* $U \in \mathcal{U}_{\text{sym}}$ *be any symmetry transformation (for example, a QRF transformation). Then there is a unique unital embedding* $\widetilde{\Phi}^U$ *of* $\mathcal{A}_{\text{alg}}^{(N)}$ *into* $\mathcal{A}_{\text{alg}}^{(N+M)}$ *which satisfies*

$$\widetilde{\Phi}^U\left(\Pi_{\text{alg}}^{(N)}(\widehat{e}_1 \otimes A_{\overline{1}})\right) = \Pi_{\text{alg}}^{(N+M)}\left(U(\widehat{e}_1 \otimes A_{\overline{1}})U^\dagger \otimes \mathbf{1}^{(M)}\right).$$

Writing $U = \bigoplus_{\mathbf{h}\in\mathcal{G}^{N-1}} U_{g(\mathbf{h})}^{\otimes N}$, *it acts as*

$$\widetilde{\Phi}^U\left(|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|\right) = \sum_{\mathbf{g}\in\mathcal{G}^M} |\mathbf{h}, g(\mathbf{h})^{-1}\mathbf{g};\mathbf{1}\rangle\langle\mathbf{j}, g(\mathbf{j})^{-1}\mathbf{g};\mathbf{1}|, \quad (23)$$

$$\widetilde{\Phi}^U\left(\Pi_{\mathbf{h};\chi\neq\mathbf{1}}\right) = \sum_{\mathbf{g}\in\mathcal{G}^N} \Pi_{\mathbf{h},\mathbf{g};\chi\neq\mathbf{1}}.$$

Via $\Phi^U := \widetilde{\Phi}^U \circ \Pi_{\text{alg}}^{(N)}$, *this extends to a completely positive unital map* $\Phi^U : \mathcal{L}(\mathcal{H}^{\otimes N}) \to \mathcal{L}(\mathcal{H}^{\otimes(N+M)})$.

Since $\Pi_{\text{alg}}^{(N)} = \Pi_{\text{alg}}^{(N)} \circ \Pi_{\text{inv}}^{(N)}$, we can write the left-hand side of the first equation above as

$$\widetilde{\Phi}^U\left(\Pi_{\text{alg}}^{(N)}(\widehat{e}_1 \otimes A_{\overline{1}})\right) = \widetilde{\Phi}^U\left(\Pi_{\text{alg}}^{(N)}(U(\widehat{e}_1 \otimes A_{\overline{1}})U^\dagger)\right).$$

In particular, if we choose U as the QRF transformation of Theorem 18 that "changes from the perspective of particle 1 to particle i", we obtain a natural invariant embedding $\widetilde{\Phi}^{(i)}$ "relative to particle i". It satisfies

$$\widetilde{\Phi}^{(i)}\left(\Pi_{\text{alg}}^{(N)}(\widehat{e}_i \otimes A_{\overline{i}})\right) = \Pi_{\text{alg}}^{(N+M)}\left(\widehat{e}_i \otimes A_{\overline{i}} \otimes \mathbf{1}^{(M)}\right), \quad (24)$$

and acts on the basis elements of $\mathcal{A}_{\text{phys}}$ as

$$\widetilde{\Phi}^{(i)}\left(|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|\right) = \sum_{\mathbf{g}\in\mathcal{G}^M} |\mathbf{h},\mathbf{g};\mathbf{1}\rangle\langle\mathbf{j}, h_{i-1}^{-1}j_{i-1}\mathbf{g};\mathbf{1}|.$$

For a better understanding of the *physical* reason of this non-uniqueness of embedding, let us reconsider our intuitive construction of $\widetilde{\Phi}^{(1)}$ above. First, note that all the $\widetilde{\Phi}^U$ satisfy Eq. (18), which had a clear physical motivation. However, the $\Phi^{(i)}$ violate Eq. (19), which we had motivated purely by formal analogy. To shed light on Eq. (19) and its generalization, Eq. (23), suppose for concreteness that we are interested in embedding $N = 2$ particles into $N + M = 3$ particles, and our group $\mathcal{G}$ is the cyclic group $\mathbb{Z}_n$ with addition modulo n. Consider the orthogonal projector $|\psi\rangle\langle\psi|$, where

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|1;\mathbf{1}\rangle + |2;\mathbf{1}\rangle\right). \quad (25)$$

This is an element of $\mathcal{H}_{\text{phys}}^{(2)}$, the 2-particle subspace of invariant states, and it describes a superposition of two particles either being one or two places apart, see Figure 6. Note that there is no origin that would locate the particles absolutely; all we have is their relations.

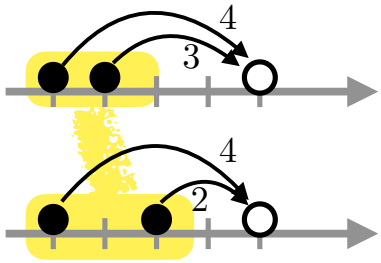


FIG. 6. The state $|\psi\rangle$ in (25) describes a superposition (indicated in yellow) of two particles (black dots) being either one or two places apart. If there is a third particle (white dot) such that the resulting state is still identical to the pure state $|\psi\rangle$ on the first two particles, then that particle should carry no information as to "which branch" is actualized. That is, its relation to the first two particles should be the same in both branches. But whether this is the case depends on our convention of how we *define* its relation to the first two particles. This results in different embedding maps Φ^U.

Now suppose we would like to embed the corresponding observable $|\psi\rangle\langle\psi|$ into the three-particle observables. The essence of the problem lies in embedding $|h;\mathbf{1}\rangle\langle j;\mathbf{1}|$, where $h = 1$ and $j = 2$. The local state of the two particles will remain coherent if the third particle carries no information on whether configuration h or j is actualized — that is, if the properties of the third particle are the same in both branches of the superposition. But the *only* properties of the third particle are its *relations* to the other two particles. Thus, if we complete $|\psi\rangle$ to a state on three particles

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|h,g;\mathbf{1}\rangle + |j,g';\mathbf{1}\rangle\right), \quad (26)$$

then the local state of the first two particles will remain coherent if and only if the third particle inside the configuration (h, g) has *the same relation to the first two particles* as the third particle in the configuration (j, g'). But the crucial insight is that this will *depend on what we mean by "relation to the first two particles"*.

For instance, suppose that $g = 4$. If we choose the convention to say that the relation is identical if *the relation to the first particle is the same*, then this will be the case if $g' = g = 4$. But if we demand instead that the *relation to*

the second particle is the same, then we need $g' = 5$. Our definition of $\Phi^{(1)}$ is implicitly relying on the former convention, while $\Phi^{(2)}$ would rely on the latter. The reason why Eq. (19) (for $\Phi = \Phi^{(1)}$) looks so simple is that we have implicitly labelled the relations $\mathbf{h} \in \mathcal{G}^{N-1}$ as *relative to the first particle* in all of this work, recall Eq. (1). This is no loss of generality, and it had no implications whatsoever for Section III, but here it becomes relevant. The factor of $h_{i-1}^{-1} j_{i-1}$ in Eq. (23) adapts the convention. Note that it does not alter the pairwise relations among the first N particles, or the pairwise relations among the last M particles, but only the relation between the two groups of particles.

Is there a way to escape the non-uniqueness of embeddings via some formal construction that is manifestly relational, but does not depend on a choice of reference within the N-particle subsystem relative to which the relation of the new M particles is defined? The right-hand side of Eq. (24) shows that the maps $\widetilde{\Phi}^{(i)}$ embed the *invariant* operator $\Pi_{\text{alg}}^{(N)}(\widehat{e}_i \otimes A_{\bar{i}})$ by embedding the original, *non-invariant* operator $\widehat{e}_i \otimes A_{\bar{i}}$ into the total Hilbert space, followed by the projection into the global subalgebra $\mathcal{A}_{\text{alg}}$. While the resulting map $\widetilde{\Phi}^{(i)}$ is invariant, its *definition* is therefore not. Can we perhaps make the definition invariant by embedding not $\widehat{e}_i \otimes A_{\bar{i}}$ directly, but its invariant part? The following lemma answers this question in the negative:

Lemma 29. *Define the map* $\widetilde{\Phi} : \mathcal{A}_{\text{alg}}^{(N)} \to \mathcal{A}_{\text{alg}}^{(N+M)}$ *as*

$$\widetilde{\Phi}(A^{(N)}) := \Pi_{\text{alg}}^{(N+M)} \left(A^{(N)} \otimes \mathbf{1}^{(M)} \right).$$

Then this map is not a valid embedding. Namely, it is not in general multiplicative, i.e. there exist $A, B \in \mathcal{A}_{\text{alg}}^{(N)}$ *with* $\widetilde{\Phi}(AB) \neq \widetilde{\Phi}(A)\widetilde{\Phi}(B)$.

Proof. A tedious but straightforward calculation yields

$$\begin{aligned}\widetilde{\Phi}\left(|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|\right) = \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\sum_{\mathbf{g}\in\mathcal{G}^M} |\mathbf{h},\mathbf{g};\mathbf{1}\rangle\langle\mathbf{j},g\mathbf{g};\mathbf{1}| \\ + \frac{1}{|\mathcal{G}|}\delta_{\mathbf{h},\mathbf{j}} \sum_{\mathbf{g}\in\mathcal{G}^M} \Pi_{\mathbf{h},\mathbf{g};\chi\neq\mathbf{1}}.\end{aligned}$$

But then, for $\mathbf{h} \neq \mathbf{j}$, we obtain

$$\widetilde{\Phi}(|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|)\widetilde{\Phi}(|\mathbf{j};\mathbf{1}\rangle\langle\mathbf{h};\mathbf{1}|) \neq \widetilde{\Phi}(|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{h};\mathbf{1}|). \qquad \square$$

A similar argument applies if we try to embed $\mathcal{A}_{\text{inv}}^{(N)}$ into $\mathcal{A}_{\text{inv}}^{(N+M)}$: the analog of the above construction, with $\Pi_{\text{alg}}^{(N+M)}$ replaced by $\Pi_{\text{inv}}^{(N+M)}$, does not yield a valid embedding.

B. A class of invariant traces

For every embedding Φ^U of Lemma 28, we obtain a corresponding "invariant trace":

Lemma 30. *For every* $U = \bigoplus_{\mathbf{h}\in\mathcal{G}^{N-1}} U_{g(\mathbf{h})}^{\otimes N} \in \mathcal{U}_{\text{sym}}$, *define a corresponding "invariant trace"* $\text{Trinv}^U := \left(\Phi^U\right)^\dagger$. *It is trace-preserving, and it maps invariant operators to invariant operators. In particular,* $\text{Trinv}_{(M)}^U \left(\mathcal{A}_{\text{alg}}^{(N+M)}\right) = \mathcal{A}_{\text{alg}}^{(N)}$, *and can be explicitly written in the following form:*

$$\text{Trinv}_{(M)}^U \rho = \sum_{\mathbf{h},\mathbf{j}\in\mathcal{G}^{N-1}} |\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}| \sum_{\mathbf{g}\in\mathcal{G}^M} \langle \mathbf{h}, g(\mathbf{h})^{-1}\mathbf{g};\mathbf{1}|\rho|\mathbf{j}, g(\mathbf{j})^{-1}\mathbf{g};\mathbf{1}\rangle + \sum_{\mathbf{h}\in\mathcal{G}^{N-1}} \Pi_{\mathbf{h};\chi\neq\mathbf{1}} \sum_{\mathbf{g}\in\mathcal{G}^M} \frac{\text{tr}(\Pi_{\mathbf{h},\mathbf{g};\chi\neq\mathbf{1}}\rho)}{|\mathcal{G}|-1}.$$

We omit the straightforward proof.

Since the different invariant traces formalize different ways to embed the two-particle observables into the three-particle observables, it is clear that the answer to the question raised at the beginning of this section will depend on the embedding. In other words, the question of whether the phase θ is accessible on the first two particles depends on the operational definition of how to access the first two particles within the *total* three-particle Hilbert space.

To illustrate this fact, let us apply two different invariant traces to the paradox. In contrast to the usual partial trace, every invariant trace yields identical results when applied to the equivalent states Ψ and Ψ' in Eqs. (14) and (15) (hence the name "invariant"). That is,

$$\text{Trinv}_3^U |\Psi\rangle\langle\Psi| = \text{Trinv}_3^U |\Psi'\rangle\langle\Psi'|$$

for all $U \in \mathcal{U}_{\text{sym}}$. However, different U yield different results. For example, consider the invariant trace $\text{Trinv}^{\mathbf{1}}$ which is the adjoint of the embedding $\Phi^{(1)}$ that we have constructed in Subsection IV A. A straightforward calculation gives

$$\begin{aligned}\text{Trinv}_3^{\mathbf{1}}|\Psi\rangle\langle\Psi| &= \frac{1}{2n}|h;\mathbf{1}\rangle\langle h;\mathbf{1}| + \frac{1}{2n}|j;\mathbf{1}\rangle\langle j;\mathbf{1}| \\ &\quad + \frac{1}{2n}\Pi_{h;\chi\neq\mathbf{1}} + \frac{1}{2n}\Pi_{j;\chi\neq\mathbf{1}} \\ &= \Pi_{\text{inv}}^{(2)}\left[|e\rangle\langle e|_1 \otimes \left(\frac{1}{2}|h\rangle\langle h| + \frac{1}{2}|j\rangle\langle j|\right)_2\right],\end{aligned}$$

Up to observational equivalence, we hence get a mixed state

$$\text{Trinv}_3^{\mathbf{1}}|\Psi\rangle\langle\Psi| \sim |e\rangle\langle e|_1 \otimes \left(\frac{1}{2}|h\rangle\langle h| + \frac{1}{2}|j\rangle\langle j|\right)_2.$$

In particular, the phase θ has disappeared. This is not surprising: ultimately, $\mathrm{Trinv}^{\mathbf{1}}$ amounts to taking the usual partial trace in the representation of the state relative to particle 1, i.e. of the state $|\Psi'\rangle$ of Eq. (15).

On the other hand, consider the symmetry transformation U from Example 20 which transforms to the center of mass. Like in Ref. [49], let us choose m_1 and m_2 such that $m_1 a = m_2 b$. Using Lemma 30, It is clear that

$$\mathrm{Trinv}_3^U\left(\Pi_{\mathbf{h};\chi\neq\mathbf{1}}\right) = \Pi_{h_1;\chi\neq\mathbf{1}}$$

and similarly for $\mathbf{h}$ replaced by $\mathbf{j}$. Furthermore,

$$\mathrm{Trinv}_3^U\left(|\mathbf{l};\mathbf{1}\rangle\langle\mathbf{p};\mathbf{1}|\right) = \sum_{h,j\in\mathbb{Z}_n} |h;\mathbf{1}\rangle\langle j;\mathbf{1}| \sum_{g\in\mathbb{Z}_n} \langle h, g+\left\lfloor\frac{m_2}{m}h\right\rfloor;\mathbf{1}|\mathbf{l};\mathbf{1}\rangle\langle\mathbf{p};\mathbf{1}|j, g+\left\lfloor\frac{m_2}{m}j\right\rfloor;\mathbf{1}\rangle = |l_1;\mathbf{1}\rangle\langle p_1;\mathbf{1}|\delta_{l_2-\lfloor\frac{m_2}{m}l_1\rfloor,p_2-\lfloor\frac{m_2}{m}p_1\rfloor}$$

for all $\mathbf{l},\mathbf{p} \in \mathbb{Z}_n^2$. Since $\mathrm{Trinv}_3^U|\Psi\rangle\langle\Psi| = \mathrm{Trinv}_3^U\left(\Pi_{\mathrm{alg}}(|\Psi\rangle\langle\Psi|)\right)$, we can expand $\Pi_{\mathrm{alg}}(|\Psi\rangle\langle\Psi|)$ into basis elements and apply the above equations. As a result, this yields exactly Eq. (13) (since a is an integer, it turns out that we can ignore all $\lfloor\cdot\rfloor$). That is,

$$\mathrm{Trinv}_3^U\left(|\Psi\rangle\langle\Psi|\right) = \Pi_{\mathrm{inv}}\left(|\psi\rangle\langle\psi|\right) \sim |\psi\rangle\langle\psi|.$$

That is, up to observational equivalence, we obtain the original pure alignable state $|\psi\rangle$ of Eq. (11). Moreover, in this observational equivalence class, $|\psi\rangle$ is *unique up to symmetry equivalence*: namely, if there is another pure alignable state $|\psi'\rangle$ with $\mathrm{Trinv}_3^U(|\Psi\rangle\langle\Psi|) \sim |\psi'\rangle\langle\psi'|$, then $\psi \sim \psi'$. But due to Lemma 22, this implies that $\psi \simeq \psi'$. Thus, in particular, there is (up to a global phase) a unique representation of this state relative to the i-th particle, $|\psi\rangle \simeq |e\rangle\langle e|_i \otimes |\varphi\rangle_{\bar{i}}$. According to Eq. (12), it is

$$|\varphi\rangle_{\bar{1}} = \frac{1}{\sqrt{2}}\left(|a+b\rangle_2 + e^{i\theta}|-a-b\rangle_2\right)$$

relative to the first particle. Thus, under the "center of mass" relational trace Trinv^U, the phase θ survives, in contrast to the result for $\mathrm{Trinv}^{\mathbf{1}}$.

C. Definition of the relational trace

Recalling the invariant algebra inclusions $\mathcal{A}_{\mathrm{phys}} \subset \mathcal{A}_{\mathrm{alg}} \subset \mathcal{A}_{\mathrm{inv}}$, we have thus far focused on constructing an invariant trace for $\mathcal{A}_{\mathrm{alg}}$, since by Lemma 26, it is the smallest algebra containing the invariant part of all alignable states and observables. However, in the previous subsection, we have seen that there does not exist a unique invariant trace on it, and the same conclusion applies to $\mathcal{A}_{\mathrm{inv}}$. We will now show that, by contrast, there *does* exists a natural embedding of the algebra $\mathcal{A}_{\mathrm{phys}}^{(N)}$ of relational N-particle operators into the algebra $\mathcal{A}_{\mathrm{phys}}^{(N+M)}$ of relational $(N+M)$-particle operators. This will also lead to a natural definition of an invariant trace in terms of relational states.

This trace, which we hence call the *relational trace*, has a natural and simple definition that is manifestly invariant under relative translations between the two particle groups. It is therefore independent of the various physically distinct conventions discussed in the previous subsections. On the relational subspace $\mathcal{H}_{\mathrm{phys}}$, the paradox is therefore unambiguously resolved. It is sufficient to focus on this subspace because Lemma 22 tells us that the relational observables are tomographically complete for the invariant information in all alignable states. Thus, $\mathcal{H}_{\mathrm{phys}}$ contains all the information we are interested in.

Lemma 31. *The map* $\widetilde{\Phi}_{\mathrm{phys}} : \mathcal{A}_{\mathrm{phys}}^{(N)} \to \mathcal{A}_{\mathrm{phys}}^{(N+M)}$, *defined by*

$$\widetilde{\Phi}_{\mathrm{phys}}(A_{\mathrm{phys}}^{(N)}) := \widehat{\Pi}_{\mathrm{phys}}^{(N+M)}\left(A_{\mathrm{phys}}^{(N)} \otimes \mathbf{1}^{(M)}\right)$$

is an embedding. It simplifies to

$$\widetilde{\Phi}_{\mathrm{phys}}(A_{\mathrm{phys}}^{(N)}) = A_{\mathrm{phys}}^{(N)} \otimes \Pi_{\mathrm{phys}}^{(M)},$$

but it is not *unital.*

Recall that this construction does not yield an embedding of all invariant observables $\mathcal{A}_{\mathrm{inv}}^{(N)}$, or even of $\mathcal{A}_{\mathrm{alg}}^{(N)}$, as we have seen in Lemma 29. Thus, it is remarkable that it works for the subalgebra of relational observables.

Proof. Noting that every $A_{\mathrm{phys}}^{(N)} \in \mathcal{A}_{\mathrm{phys}}^{(N)}$ satisfies $A_{\mathrm{phys}}^{(N)} = \widehat{\Pi}_{\mathrm{phys}}^{(N)}(A_{\mathrm{phys}}^{(N)})$, we can recast $\widetilde{\Phi}_{\mathrm{phys}}$ in the form

$$\widetilde{\Phi}_{\mathrm{phys}}(A_{\mathrm{phys}}^{(N)}) = \Pi_{\mathrm{phys}}^{(N+M)}\left(\Pi_{\mathrm{phys}}^{(N)} \otimes \mathbf{1}^{(M)}\right)\left(A_{\mathrm{phys}}^{(N)} \otimes \mathbf{1}^{(M)}\right) \times \left(\Pi_{\mathrm{phys}}^{(N)} \otimes \mathbf{1}^{(M)}\right)\Pi_{\mathrm{phys}}^{(N+M)}.$$

By considering the action on an arbitrary basis state $|g_1,\ldots,g_{N+M}\rangle$, it is easy to verify that

$$\left(\Pi_{\mathrm{phys}}^{(N)} \otimes \mathbf{1}^{(M)}\right)\Pi_{\mathrm{phys}}^{(N+M)} = \Pi_{\mathrm{phys}}^{(N)} \otimes \Pi_{\mathrm{phys}}^{(M)}. \quad (27)$$

Hence, using idempotence of the projector,

$$\widetilde{\Phi}_{\mathrm{phys}}(A_{\mathrm{phys}}^{(N)}) = A_{\mathrm{phys}}^{(N)} \otimes \Pi_{\mathrm{phys}}^{(M)}$$

and the image commutes with $\Pi_{\mathrm{phys}}^{(N+M)}$. Checking the embedding properties is now trivial. Furthermore, note that the unit element of $\mathcal{A}_{\mathrm{phys}}^{(N)}$ is $\Pi_{\mathrm{phys}}^{(N)}$,

while the unit element of $\mathcal{A}^{(N+M)}_{\text{phys}}$ is $\Pi^{(N+M)}_{\text{phys}}$. However, $\widetilde{\Phi}_{\text{phys}}(\Pi^{(N)}_{\text{phys}}) = \Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}}$. To see that $\widetilde{\Phi}_{\text{phys}}(\Pi^{(N)}_{\text{phys}})$ does not act as the identity on $\mathcal{H}^{(N+M)}_{\text{phys}}$, observe that $\operatorname{tr}\left(\Pi^{(N)}_{\text{phys}}\right) = \dim \mathcal{H}^{(N)}_{\text{phys}} = |\mathcal{G}|^{N-1}$, hence $\operatorname{tr}\left(\Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}}\right) = |\mathcal{G}|^{N+M-2}$, but $\operatorname{tr}\left(\Pi^{(N+M)}_{\text{phys}}\right) = |\mathcal{G}|^{N+M-1}$. □

Returning to the discussion of how to define the relations of the additional M particles to the first group of N particles (cf. Fig. 6), we now have a unique answer: the projector $\Pi^{(M)}_{\text{phys}}$ in $\widetilde{\Phi}_{\text{phys}}$ takes the *coherent average* over all possible such relations. This can also be seen by inspecting the action of $\widetilde{\Phi}_{\text{phys}}$ on the basis elements of $\mathcal{A}_{\text{phys}}$:

$$\widetilde{\Phi}_{\text{phys}}\left(|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|\right) = \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G},\mathbf{g}\in\mathcal{G}^M} |\mathbf{h},\mathbf{g};\mathbf{1}\rangle\langle\mathbf{j},g\mathbf{g};\mathbf{1}|.$$

In contrast to Eq. (26) and the embeddings of Subsection IV A, this embedding does *not* assign to every M-particle configuration $\mathbf{g}$ another one, $\mathbf{g}'$, which has "the same relation" to the first N particles in branches $\mathbf{h}$ and $\mathbf{j}$. Instead, it generates the *uniform superposition* of all the possibilities.

This averaging is also the reason for the failure of the unitality property. However, as we will see shortly, the absence of unitality is precisely the reason why the relational trace defined below maps relational $(N+M)$-particle states into relational N-particle states. It will thus be rather a feature than a failure.

While $\widetilde{\Phi}_{\text{phys}}$ is not unital, note that the embedding of the relational N-particle unit $\widetilde{\Phi}_{\text{phys}}(\Pi^{(N)}_{\text{phys}}) = \Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}}$ certainly *does* act as the identity on its image,

$$\mathcal{H}^{(N\otimes M)}_{\text{phys}} := \left\{\Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}}\,|\psi\rangle \;\middle|\; |\psi\rangle \in \mathcal{H}^{\otimes(N+M)}\right\}.$$

This subspace of the space of relational $(N+M)$-particle states $\mathcal{H}^{(N+M)}_{\text{phys}}$ will be essential below when resolving the paradox of the third particle.

The natural embedding induces a natural trace.

Definition 32 (Relational trace). *The relational trace is defined to be the Hilbert-Schmidt adjoint* $\text{Trel} := \Phi^{\dagger}_{\text{phys}}$ *of the extended embedding map* $\Phi_{\text{phys}} : \mathcal{L}(\mathcal{H}^{\otimes N}) \to \mathcal{L}(\mathcal{H}^{\otimes(N+M)})$ *defined by* $\Phi_{\text{phys}} := \widetilde{\Phi}_{\text{phys}} \circ \widehat{\Pi}^{(N)}_{\text{phys}}$. *That is, the relational trace is the unique map with the property*

$$\operatorname{tr}\left(\Phi_{\text{phys}}(A^{(N)})\,\rho\right) = \operatorname{tr}\left(A^{(N)}\,\text{Trel}_{(M)}\rho\right)$$

for all $A^{(N)} \in \mathcal{L}(\mathcal{H}^{\otimes N})$ *and all* $\rho \in \mathcal{L}(\mathcal{H}^{\otimes(N+M)})$ *(in particular for all states).*

Specifically, note that

$$\begin{aligned}\operatorname{tr}\left(\Phi_{\text{phys}}(A^{(N)})\,\rho\right) &= \operatorname{tr}\left(\Phi_{\text{phys}}(A^{(N)})\,\widehat{\Pi}^{(N+M)}_{\text{phys}}(\rho)\right)\\ &= \operatorname{tr}\left(\Phi_{\text{phys}}(A^{(N)})\,\rho_{\text{phys}}\right). \qquad (28)\end{aligned}$$

This is precisely the expectation value of the relational observable $\Phi_{\text{phys}}(A^{(N)}) \in \mathcal{A}^{(N+M)}_{\text{phys}}$ in the relational state $\rho_{\text{phys}} \in \mathcal{S}(\mathcal{H}^{(N+M)}_{\text{phys}})$, evaluated in the manifestly invariant inner product on $\mathcal{S}(\mathcal{H}^{(N+M)}_{\text{phys}})$. The relational trace is thus unambiguously defined in terms of the so-called physical inner product of constraint quantization [19, 66, 67], i.e. the inner product on $\mathcal{H}_{\text{phys}}$. We will analyze this in more detail in our upcoming work [62].

Theorem 33. *The relational trace takes the explicit form*

$$\text{Trel}_{(M)}\rho = \text{Tr}_{(M)}\left[\Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}}\,\rho\,\Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}}\right],$$

where $\text{Tr}_{(M)}$ *is the standard partial trace over particles* $N+1,\ldots,N+M$. *It maps relational operators onto relational operators, i.e.* $\text{Trel}_{(M)}\left(\mathcal{A}^{(N+M)}_{\text{phys}}\right) = \mathcal{A}^{(N)}_{\text{phys}}$, *and is trace-preserving for states in* $\mathcal{S}(\mathcal{H}^{(N\otimes M)}_{\text{phys}})$, *but trace-decreasing outside of it. Furthermore, it preserves observational equivalence, i.e.* $\rho \sim \sigma$ *implies* $\text{Trel}_{(M)}\rho = \text{Trel}_{(M)}\sigma$.

Proof. The first statement follows from

$$\begin{aligned}\operatorname{tr}\left(\Phi_{\text{phys}}(A^{(N)})\,\rho\right) &= \operatorname{tr}\left[\left(\Pi^{(N)}_{\text{phys}}A^{(N)}\Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}}\right)\rho\right]\\ &= \operatorname{tr}\Big[\left(A^{(N)} \otimes \mathbf{1}^{(M)}\right)\\ &\quad\times\left(\Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}}\,\rho\,\Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}}\right)\Big],\end{aligned}$$

which holds for any $A^{(N)} \in \mathcal{L}(\mathcal{H}^{\otimes N})$ and any $\rho \in \mathcal{L}(\mathcal{H}^{\otimes(N+M)})$.

Given the conjugation of its input with the projector $\Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}}$, it is clear that $\text{Trel}_{(M)}$ is trace-preserving for states in $\mathcal{S}(\mathcal{H}^{(N\otimes M)}_{\text{phys}})$, but not for states outside of it. It is also clear that $\text{Trel}_{(M)}$ maps operators from $\mathcal{A}^{(N+M)}_{\text{phys}}$ into operators in $\mathcal{A}^{(N)}_{\text{phys}}$, since the projectors $\Pi^{(N)}_{\text{phys}}$ can be taken outside of the trace over particles $N+1,\ldots,N+M$. To see that it is surjective, it is straightforward to check that $|\mathcal{G}|^{1-M}\text{Trel}_{(M)}\left(A^{(N)}_{\text{phys}} \otimes \mathbf{1}^{(M)}\right) = A^{(N)}_{\text{phys}}$ for all $A^{(N)}_{\text{phys}} \in \mathcal{A}^{(N)}_{\text{phys}}$.

Finally, note that the image of Φ_{phys} is contained in $\mathcal{A}^{(N+M)}_{\text{phys}}$, thus $\Phi_{\text{phys}} = \widehat{\Pi}^{(N+M)}_{\text{phys}} \circ \Phi_{\text{phys}}$. Taking the Hilbert-Schmidt adjoint of this equation yields $\text{Trel}_{(M)} \circ \widehat{\Pi}^{(N+M)}_{\text{phys}} = \text{Trel}_{(M)}$. Now suppose we have $\rho \sim \sigma$, then Lemma 14 and $\widehat{\Pi}^{(N+M)}_{\text{phys}} \circ \Pi^{(N+M)}_{\text{inv}} = \widehat{\Pi}^{(N+M)}_{\text{phys}}$ imply $\widehat{\Pi}^{(N+M)}_{\text{phys}}(\rho) = \widehat{\Pi}^{(N+M)}_{\text{phys}}(\sigma)$. Altogether this implies that $\text{Trel}_{(M)}$ preserves observational equivalence. □

We can write

$$\text{Trel}_{(M)} = \widehat{\Pi}^{(N)}_{\text{phys}} \circ \text{Tr}_{(M)} \circ \widehat{\Pi}^{(N+M)}_{\text{phys}}. \qquad (29)$$

We can thus view the relational partial trace $\text{Trel}_{(M)}$ as an invariant extension of the standard partial trace

$\mathrm{Tr}_{(M)}$. The non-unitality of $\widetilde{\Phi}_{\mathrm{phys}}$ is reflected in the final application of $\widehat{\Pi}^{(N)}_{\mathrm{phys}}$. Without this projection, the image of $\mathrm{Trel}_{(M)}$ would not in general be contained in $\mathcal{A}^{(N)}_{\mathrm{phys}}$. For example, the uniform mixture $\rho^{(N+M)} := \Pi^{(N+M)}_{\mathrm{phys}}/|\mathcal{G}|^{N+M-1}$ on the physical subspace of $N+M$ particles yields $\mathrm{Tr}_{(M)}\rho^{(N+M)} = \mathbf{1}^{(N)}/|\mathcal{G}|^N$, whose decomposition according to Theorem 12 contains operators outside of $\mathcal{A}^{(N)}_{\mathrm{phys}}$.[9] *Thus, non-unitality is the price to pay for remaining relational.*

This leads to $\mathrm{Trel}_{(M)}$ being trace-*decreasing*, unless the initial state is fully supported on the subspace $\mathcal{H}^{(N\otimes M)}_{\mathrm{phys}}$. Should we be worried about this fact — shouldn't marginals of normalized quantum states be normalized? Not in this case. In contrast to the standard partial trace, the relational trace is *not* supposed to tell us what the reduced quantum state on a subsystem is. Instead, it is constructed to tell us precisely the following:

Theorem 34. *Given some $(N+M)$-particle state $\rho^{(N+M)} \in \mathcal{S}(\mathcal{H}^{\otimes(N+M)})$, the following* conditional state of N particles *is normalized or subnormalized:*

$$\rho^{(N)} := \frac{\mathrm{Trel}_{(M)}\rho^{(N+M)}}{\mathrm{tr}\left(\rho^{(N+M)}\Pi^{(N+M)}_{\mathrm{phys}}\right)}.$$

Consider any relational projector $0 \leq E^{(N)}_{\mathrm{phys}} \leq \Pi^{(N)}_{\mathrm{phys}}$ which we interpret as a "relational event". Then the state $\rho^{(N)}$ tells us the probabilities of this N-particle event, conditioned on the $(N+M)$-particle system being relational*:*

$$\mathrm{tr}\left(E^{(N)}_{\mathrm{phys}}\rho^{(N)}\right) = \mathrm{Prob}\left(E^{(N)}_{\mathrm{phys}}|\ \Pi^{(N+M)}_{\mathrm{phys}}\right).$$

That is, the renormalized result $\rho^{(N)}$ of the relational trace gives us **the expectation values of all N-particle relational observables, conditioned on the global $(N+M)$-particle state being fully relational.**

Proof. It follows from Theorem 33 and Eq. (27) that

$$\begin{aligned}\mathrm{tr}\left(\mathrm{Trel}_{(M)}\rho^{(N+M)}\right) &= \mathrm{tr}\left(\Pi^{(N)}_{\mathrm{phys}}\otimes\Pi^{(M)}_{\mathrm{phys}}\rho^{(N+M)}\right)\\ &\leq \mathrm{tr}\left(\Pi^{(N+M)}_{\mathrm{phys}}\rho^{(N+M)}\right),\end{aligned}$$

hence $\rho^{(N)}$ is not supernormalized. Now, the probability that an initial global measurement of the projector $\Pi^{(N+M)}_{\mathrm{phys}}$ yields outcome "yes", and then a subsequent local measurement of $E^{(N)}_{\mathrm{phys}}$ yields "yes" too, is

$$\begin{aligned}&\mathrm{Prob}\left(E^{(N)}_{\mathrm{phys}}, \Pi^{(N+M)}_{\mathrm{phys}}\right)\\ &\quad= \mathrm{tr}\left(\Phi_{\mathrm{phys}}(E^{(N)}_{\mathrm{phys}})\widehat{\Pi}^{(N+M)}_{\mathrm{phys}}(\rho^{(N+M)})\right)\\ &\quad= \mathrm{tr}\left(E^{(N)}_{\mathrm{phys}}\mathrm{Trel}_{(M)}\circ\widehat{\Pi}^{(N+M)}_{\mathrm{phys}}(\rho^{(N+M)})\right)\\ &\quad= \mathrm{tr}\left(E^{(N)}_{\mathrm{phys}}\mathrm{Trel}_{(M)}\rho^{(N+M)}\right),\end{aligned}$$

where we have used Definition 32 and Eq. (29). The rest of the claim follows from the definition of conditional probability. □

Recall that in Lemma 9, we have seen that the projection $\rho_{\mathrm{phys}} := \widehat{\Pi}^{(N+M)}_{\mathrm{phys}}(\rho)$ for alignable states ρ is subnormalized, but is sufficient to determine the expectation values of all invariant observables. Thus, ρ_{phys} should not be seen as the marginal of ρ on some subsystem, but as the "relational part" of ρ. The "relational weight" $\mathrm{tr}\,\rho_{\mathrm{phys}}$ is in general less than one,[10] and it can decrease when disregarding some of the particles.

[9] More generally, the image of $\mathcal{H}^{(N+M)}_{\mathrm{phys}}\setminus\mathcal{H}^{(N\otimes M)}_{\mathrm{phys}}$ under $\mathrm{Tr}_{(\mathrm{M})}$ does not lie in the N-particle relational subspace $\mathcal{H}^{(N)}_{\mathrm{phys}}$.

[10] This assumes that $\rho \in \mathcal{L}(\mathcal{H}^{\otimes(N+M)})$ in $\rho_{\mathrm{phys}} = \widehat{\Pi}^{(N+M)}_{\mathrm{phys}}(\rho)$ is normalized, as appropriate in the context of our manuscript where there is an external observer who could measure also non-invariant observables in the presence of an external frame. By contrast, in the perspective-neutral approach [39, 41–45] (and more generally in constraint quantization), one disregards any external structure and works directly with normalized relational states. That is, one would *define* the normalization of the $(N+M)$-particle relational state as $\mathrm{tr}\rho_{\mathrm{phys}} = 1$. Indeed, note that the *full* standard trace $\mathrm{tr}\rho_{\mathrm{phys}} = \mathrm{tr}\left(\Pi^{(\mathrm{N+M})}_{\mathrm{phys}}\rho\right)$ is precisely the extension of the so-called physical inner product [19, 66, 67] to density matrices. However, note also from the previous footnote that the standard *partial* trace $\mathrm{Tr}_{(\mathrm{M})}$ is not in general appropriate for relational states ρ_{phys}. The additional projection with $\widehat{\Pi}^{(N)}_{\mathrm{phys}}$ in the relational trace fixes this issue, but reduces the norm of states with support outside of $\mathcal{H}^{(N\otimes M)}_{\mathrm{phys}}$. This has a transparent physical interpretation which is best seen through the norm reduction of invariant basis states: $\mathrm{Trel}_{(\mathrm{M})}(|\mathbf{h}_{\mathrm{N}},\mathbf{h}_{\mathrm{M}};\mathbf{1}\rangle\langle\mathbf{j}_{\mathrm{N}},\mathbf{j}_{\mathrm{M}};\mathbf{1}|)$, where $\mathbf{h}_N,\mathbf{j}_N \in \mathcal{G}^{N-1}$ and $\mathbf{h}_M,\mathbf{j}_M \in \mathcal{G}^M$, is equal to $\frac{1}{|\mathcal{G}|}\,|\mathbf{h}_N;\mathbf{1}\rangle\langle\mathbf{j}_N;\mathbf{1}|$ if $\mathbf{h}_M = g\mathbf{j}_M$ for some $g \in \mathcal{G}$, and is zero otherwise. The variables $\mathbf{h}_M,\mathbf{j}_M$ lying in $\mathcal{G}^M$ rather than $\mathcal{G}^{M-1}$ reflects the fact that they encode not only the $M-1$ internal relations of the M particles, but also a definite relation between the two particle groups, which is a property of *both* groups together. The normalization reduction factor $1/|\mathcal{G}|$ comes from the coherent averaging over the relations between the two particle groups (which is partially a property of the N particles), and quantifies the corresponding ignorance of the relational N-particle state obtained via $\mathrm{Trel}_{(\mathrm{M})}$. By construction, the latter contains all information about the relational N-particle observables which are independent of the relation between the particle groups. However, the 'ignorance factor' $1/|\mathcal{G}|$ has to be taken into account.

D. Relational resolution of the paradox

Let us now apply these insights to the paradox of the third particle. Suppose $\rho^{(N)}$ is a state of the N particles before taking the additional M particles into account, prepared by an observer with access to the external reference frame. The corresponding relational state is $\rho^{(N)}_{\text{phys}} = \widehat{\Pi}^{(N)}_{\text{phys}}(\rho^{(N)})$. Next, suppose the M additional particles are prepared in a normalized state $\rho^{(M)}$, and the composite state of all particles is of the product form $\rho^{(N+M)} = \rho^{(N)} \otimes \rho^{(M)}$. We can then construct the relational state corresponding to this composition. Here it is important to note that $\rho^{(N+M)}_{\text{phys}} = \widehat{\Pi}^{(N+M)}_{\text{phys}}\left(\rho^{(N+M)}\right) = \widehat{\Pi}^{(N+M)}_{\text{phys}}\left(\widetilde{\rho}^{(N+M)}\right)$, for any $\widetilde{\rho}^{(N+M)} \simeq \rho^{(N+M)}$. That is, all members of the symmetry equivalence class of $\rho^{(N+M)}$ (incl. states featuring entanglement between the two particle groups) yield the same relational $(N+M)$-particle state. However, due to Theorem 33, it is only the projection into $\mathcal{H}^{(N\otimes M)}_{\text{phys}}$ that matters for the relational trace. But Eq. (27) implies that

$$\begin{aligned}
&\Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}}\, \rho^{(N+M)}_{\text{phys}}\, \Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}} \\
&= \Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}} \left(\rho^{(N)} \otimes \rho^{(M)}\right) \Pi^{(N)}_{\text{phys}} \otimes \Pi^{(M)}_{\text{phys}} \\
&= \rho^{(N)}_{\text{phys}} \otimes \rho^{(M)}_{\text{phys}},
\end{aligned}$$

and so we have $\text{Trel}_{(M)}\rho^{(N+M)}_{\text{phys}} = \rho^{(N)}_{\text{phys}} \cdot \text{tr}_{(M)}\rho^{(M)}_{\text{phys}}$. Up to a constant factor, this is precisely the initial relational N-particle state that we had before taking the additional M particles into account.

Let us see what this implies for the paradox of the third particle. Thus, let us concretely compute the relational trace Trel_3 of the state $|\Psi\rangle$ of Eq. (14). Due to invariance, the result will be identical if we apply it to the state $|\Psi'\rangle$ of Eq. (15). As a first step, we find

$$\begin{aligned}
\Pi^{(3)}_{\text{phys}}(|\Psi\rangle\langle\Psi|)\Pi^{(3)}_{\text{phys}} = \frac{1}{2n}|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{h};\mathbf{1}| + \frac{e^{-i\theta}}{2n}|\mathbf{h};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}| \\
+\frac{e^{i\theta}}{2n}|\mathbf{j};\mathbf{1}\rangle\langle\mathbf{h};\mathbf{1}| + \frac{1}{2n}|\mathbf{j};\mathbf{1}\rangle\langle\mathbf{j};\mathbf{1}|.
\end{aligned}$$

Using ${}_3\langle g_3|\mathbf{h};\mathbf{1}\rangle = \frac{1}{\sqrt{|\mathcal{G}|}}\,|g_3 h_2^{-1}, g_3 h_2^{-1} h_1\rangle$ yields

$$\begin{aligned}
\text{Trel}_3(|\Psi\rangle\langle\Psi|) &= \Pi^{(2)}_{\text{phys}}\text{Tr}_3\left[\Pi^{(3)}_{\text{phys}}(|\Psi\rangle\langle\Psi|)\Pi^{(3)}_{\text{phys}}\right]\Pi^{(2)}_{\text{phys}} \\
&= \Pi^{(2)}_{\text{phys}}|\psi\rangle\langle\psi|\,\Pi^{(2)}_{\text{phys}} \\
&= \frac{1}{2n^2}|h;\mathbf{1}\rangle\langle h;\mathbf{1}| + \frac{e^{-i\theta}}{2n^2}|h;\mathbf{1}\rangle\langle\, j;\mathbf{1}| \\
&\quad + \frac{e^{i\theta}}{2n^2}|j;\mathbf{1}\rangle\langle h;\mathbf{1}| + \frac{1}{2n^2}|j;\mathbf{1}\rangle\langle j;\mathbf{1}|.
\end{aligned}$$

Since $|\Psi\rangle$ is alignable, Lemma 22 tells us that $\langle\Psi|\Pi^{(N+M)}_{\text{phys}}|\Psi\rangle = 1/n$. Thus, computing the conditional state of Theorem 34, we obtain the projection of the state in Eq. (13) into the relational subalgebra. Hence, we recover exactly the relational state of the first two particles which we had before adding the third. In particular, the phase θ is preserved: as expected, it remains accessible on the first two particles.

The algebra $\mathcal{A}_{\text{phys}}$ generated by relational observables and the subspace $\mathcal{H}_{\text{phys}}$ of relational states is the arena of the perspective-neutral approach to QRFs [39, 41–45]. As such, there is no paradox of the third particle in this approach. Furthermore, it provides a compelling conceptual interpretation of this resolution: the relational states are the perspective-neutral states, i.e. they correspond to a description of the composite particle system prior to choosing an internal reference relative to which the state is described. The perspective-neutral states contain the entire information about all internal QRF perspectives at once. The relational trace is performed at the perspective-neutral level, and consistency at that level implies consistency in all internal perspectives. We will further elaborate on this in Ref. [62].

E. Comparison to the resolution by Angelo et al.

Angelo et al. [49] also propose a resolution to the apparent paradox that they have raised in their paper. Let us recapitulate their resolution in our terminology and compare the two approaches. First, they introduce an operator $T := e^{-2i(a+b)\widehat{p}_{r_2}}$ which in our notation simply implements a translation of the two particles, given by $T|g_1, g_2\rangle = |g_1 - 2a, g_2 + 2b\rangle$. Computing the expectation value in the two-particle state of Eq. (11) yields $\langle\psi|T|\psi\rangle = \frac{1}{2}e^{i\theta}$. Thus, this operator (or rather its real and imaginary parts) admit the measurement of the phase θ.

Since we are in the framework of Example 8, we can use the explicit form of the characters to see that

$$T|h;\chi_k\rangle = \chi_k(-2a)|h + 2a + 2b;\chi_k\rangle.$$

In particular, the invariant part of this translation can be expressed in the form

$$T|h;\mathbf{1}\rangle = |h + 2a + 2b;\mathbf{1}\rangle. \tag{30}$$

That is, T increases the relative distance of the particles by $2a + 2b$. Indeed, if we define

$$T_{\text{inv}} := \Pi_{\text{inv}}(T) = \Pi_{\text{alg}}(T) = \widehat{\Pi}_{\text{phys}}(T),$$

then $T_{\text{inv}} \in \mathcal{A}_{\text{phys}}$ satisfies Eq. (30). Note, however, that T contains strictly more information than T_{inv}: not only does it tell us that the relative distance of the particles increases by $2a + 2b$, but it also tells us what happens to their *absolute* positions.

Angelo et al. write: *"The crucial (and surprising) observation is that T actually shifts the relative coordinate of particle 3 as well as that of particle 2."* Strictly speaking, this is not a claim about T, but about the embedding $T^{(3)}$ of T into the three-particle observables. Using the obvious embedding $T^{(3)} = T \otimes \mathbf{1}$, we obtain

$T^{(3)}|g_1, g_2, g_3\rangle = |g_1 - 2a, g_2 + 2b, g_3\rangle$, and thus

$$T^{(3)}|h_1, h_2; \mathbf{1}\rangle = |h_1 + 2a + 2b, h_2 + 2a; \mathbf{1}\rangle.$$

That is, the relative coordinate of particle 3 is also shifted by $2a$. This reproduces Angelo et al.'s claim, but it is important to understand where the shift of $h_2 \mapsto h_2 + 2a$ comes from. It is certainly not possible to deduce this shift from T_{inv} alone. Instead, it comes form *the specific choice* of implementing T_{inv} (a relative shift of $2a + 2b$) via T (absolutely shifting particle 1 by $-2a$ and particle 2 by $2b$). And the latter choice comes from Angelo et al.'s decision of *preserving the center of mass*, which fixes the non-invariant action of T.

In summary: Angelo et al.'s proposed resolution of the paradox comes from *deciding to embed the two-particle observables via the center-of-mass embedding* that we have described in Subsection IV B. As shown there, this leads to a preservation of the phase θ. However, as also discussed in Subsection IV B, there exist other equally well-motivated, but physically inequivalent choices of embedding which may or may not preserve θ.

In contrast, our resolution amounts to the construction of a *relational embedding* for which no such choice has to be made in the first place. Nonetheless, Angelo et al's insight is still important: embedding fewer into more particles will in general "do something" to the additional particles, and care has to be taken of how the embedding is accomplished.

V. CONCLUSIONS

The aim of this article is to elucidate the operational essence and interpretation of the recent structural approach to QRFs [38–40], and to also clarify the meaning of its QRF transformations as symmetry transformations. These insights have then been exploited to illuminate the physics behind the apparent 'paradox of the third particle' of Ref. [49] and to resolve it at a formal level through relational observables.

We began by providing a careful conceptual comparison of the quantum information [1–8] and structural approaches to QRFs, illustrating the difference in their operational essence in terms of two communication scenarios. While both approaches focus on an external relatum independent description of physical observables and quantum states, they do so in different manners and with different goals. As we have seen, technically a distinction can be drawn between the two in terms of how they describe external relatum independent states fundamentally: they are the *incoherently* and *coherently* group-averaged states in the QI and structural approach, respectively.

A key ambition of the QI approach is to elucidate how to perform communication protocols between different parties in the absence of a shared external laboratory frame. To this end, it suffices to focus on physical properties of the communicated quantum system that are meaningful relative to an arbitrary choice of external frame. For example, this can be achieved by restricting to speakable information that is encoded in decoherence-free subspaces or by communicating an *additional* reference quantum system that serves as a token for the sender's reference frame. Either way, the QI approach maintains the reference frame external to the system of interest. For successfully carrying out such operational protocols it is also not necessary to take an extra step and choose an internal reference frame within the system of interest, and to ask how the quantum system is described relative to one of its subsystems.

However, this additional step is precisely what the structural approach aims for. Its primary goal is not the implementation of protocols for communicating physical information. It has rather a more fundamental ambition: to dissolve the distinction between quantum systems and reference frames and thereby to extend the set of available reference frame choices to include subsystems of the physical system of interest. Its focal point are thus not only external relatum independent state descriptions, but *internal* state descriptions. The operational essence of the structural approach can be illustrated in a scenario in which different agents agree on a (redundancy-free) *description* of physical quantum states *without* adhering to an external relatum. They can always achieve this task by invoking certain "canonical choices" in the representation of quantum states that exploit the *internal* structure of the quantum system to be described. In particular, these canonical choices of representation are related by transformations that coincide with the QRF transformations in the structural approach.

To show this explicitly, we have then formalized these conceptual observations in the context of an N-particle quantum system ("$\mathcal{G}$-system") where the configuration space of each particle is a finite Abelian group $\mathcal{G}$. We chose this simple setting in order to avoid technicalities and to render all appearing structures completely transparent. But we emphasize that our observations are of more general validity. They apply directly to laboratory situations in which agents simply disregard, or do not have access to, a relatum external to the system of interest, but in principle also to the case that no external physical relatum exists in the first place as, e.g., in quantum cosmology (see Refs. [43, 68] for a related discussion).

We determined the symmetry group $\mathcal{G}_{\text{sym}}$ associated with a $\mathcal{G}$-system, which preserves all its external-relatum-independent structure. We showed that states from the corresponding symmetry equivalence class are alignable to a choice of reference system through a $\mathcal{G}_{\text{sym}}$ transformation: each equivalence class contains "canonical choices" of state representations, and these correspond to selecting any one of the N particles as a reference system to define the origin and to describing the remaining $N - 1$ particles relative to it. These canonical choices are the 'internal QRF perspectives' on the

N-particle system of Refs. [38–40].

The symmetry group $\mathcal{G}_{\rm sym}$ contains the 'classical translation group' $\mathcal{G}$ as a strict *subgroup*, but it also contains 'relation-conditional translations' which turn out to include the QRF transformations of Ref. [40], which are equivalent to those of Refs. [38, 39]. While it is evident from these works that the QRF transformations are conditional translations, the present article clarifies that they are *symmetry* transformations with a precise and transparent physical interpretation.

Being translations conditional on the particle relations, the QRF transformations make sense in a classical context when dealing with, for example, statistical mixtures of particle positions rather than superpositions, and indeed have classical analogs that have been exhibited in Refs. [39–43]. Nevertheless, just like the CNOT gate has a classical meaning, but can generate entanglement, the QRF transformations similarly lead to interesting quantum effects such as a QRF dependence of, e.g. entanglement and superpositions [38–40], classicality [39, 71, 72], spin [46, 47], certain quantum resources [73], temporal locality [44, 48], and of comparing quantum clock readings [42, 45].[11]

Given the two groups $\mathcal{G}_{\rm sym}$ and $\mathcal{G}$ in the setup, one has *a priori* two distinct ways to construct invariant states and observables. Interestingly, as we have shown, the invariant (pure) states and thus the subspace $\mathcal{H}_{\rm phys}$ of relational states do *not* in fact depend on whether one requires invariance under the action of $\mathcal{G}_{\rm sym}$ or its subgroup $\mathcal{G}$. By contrast, the set of invariant observables *does* depend on which group one works with: the operator algebra invariant under $\mathcal{G}_{\rm sym}$ is a strict subset of the operator algebra invariant under its subgroup $\mathcal{G}$. However, the two invariant operator algebras coincide again in their restriction to the space of relational states $\mathcal{H}_{\rm phys}$, which is the algebra $\mathcal{A}_{\rm phys}$ generated by so-called relational observables [18–26, 42–45]. The space of relational states $\mathcal{H}_{\rm phys}$ and the relational operator algebra $\mathcal{A}_{\rm phys}$ are key structures in constraint quantization [19, 66, 67] and the platform of the perspective-neutral approach to QRFs [39, 41–45] (part of the structural approach), which are thus independent of the distinction between $\mathcal{G}_{\rm sym}$ and $\mathcal{G}$. The difference between $\mathcal{G}_{\rm sym}$ and its subgroup $\mathcal{G}$ is, however, crucial when aligning the *non-invariant* description of quantum states to a particle at the level of the full N-particle Hilbert space.

These observations also permitted us to first clarify the physics behind the 'paradox of the third particle' discussed in Ref. [49], and subsequently to resolve it at a formal level. First, we have illuminated why the usual partial trace is *not* suitable in the context of QRFs, because it ignores the equivalence classes of states that are operationally indistinguishable in the absence of an external relatum. Next, we have explained that, in order to take the observational equivalence classes into account, one has to construct a partial trace in terms of the invariant observables. However, even when attempting to do so, we have seen that there does not exist a physically distinguished choice for such an invariant partial trace *outside the space of relational states*. The reason is that an invariant partial trace demands a suitable embedding of the two-particle invariant observables into the three-particle invariant observables. Yet such an embedding (while invariant under symmetry transformations) depends on how one defines the relation of the third to the first two particles, and there are multiple physically inequivalent ways (e.g., distance to the first particle, center of mass, etc.). The two-particle reduced state then depends on one's convention of how to define the relation between the third and the first two particles, despite restricting attention to invariant observables.

However, when restricting attention further to the algebra $\mathcal{A}_{\rm phys}$ generated by relational observables and the space of relational states $\mathcal{H}_{\rm phys}$, we showed that there *does* exist a physically distinguished embedding of the relational two-particle observables and states into the relational three-particle observables and states. Physically, this embedding corresponds to *coherently averaging* over all possible relations between the third and the first two particles, and thereby defines an entirely invariant embedding. This permitted us to define an unambiguous relational partial trace that determines the expectation values of relational observables on subsets of particles. In particular, this trace achieves for relational states what a consistent partial trace should do: if a third particle is independently prepared, then the two-particle reduced state, obtained from the relational three-particle state, coincides with the relational two-particle state prior to taking the third particle into account. At the level of relational observables and relational states, the paradox of the third particle of Ref. [49] is thus resolved; in this sense, the perspective-neutral approach does not feature any paradox of additional particles. However, we have not discussed what it would mean for an agent to operationally implement this resolution in the lab and, specifically, how they may operationally restrict to relational states and observables (although we believe this to be possible). In this light, our resolution of the paradox is formal.

The paradox of the third particle and our resolution can be viewed as a finite-dimensional analog of the problem of boundaries and edge modes in gauge theory and gravity [51–56]. Boundaries in space or spacetime usually break gauge-invariance and constitute challenges for gauge-invariant observables. The latter are typically non-local (such as Wilson loops) and can thus have support in two neighbouring regions separated by a boundary. Those gauge-invariant observables with support in both regions determine the physical relation between the two and are accounted for in terms

[11] It would be interesting to study the recent proposals [77–79] for quantum time dilation effects in terms of the temporal QRF transformations as in Refs. [42–45, 48].

of so-called edge modes when one of the regions is ignored. This is analogous to the joining of two groups of N and M particles and asking for the invariant relations between the two groups. The relative distances between the two groups of particles are the finite-dimensional analog of the gauge-invariant observables in gauge theories and gravity that have support in two neighbouring regions. As we have seen, ignoring one group of particles by simply taking the standard partial trace may indeed lead to an invariance breaking in analogy to the field theory case, i.e. N-particle states that are *not* relational. This is because the set of relational observables for the joint $(N + M)$-particle system is *not* only the union of the sets of relational N- and M-particle observables, again in analogy to two neighbouring subregions in spacetime. Our relational trace defines a purely relational, i.e. invariant way of 'ignoring' a group of particles, and it would be interesting to extend this tool to the study of edge modes in gauge theories and gravity.

Lastly, we emphasize that our novel interpretation of the structural approach applies in particular also to temporal quantum reference frames, i.e. quantum clocks. For instance, the example of the cyclic group could model a set of quantum clocks each with a finite set of readings. The external frame would then be some laboratory clock that one external observer has access to, but another may not. Nevertheless, the two observers can agree on the description of the flow of time by focusing on a purely *internal* choice of clock that leads to a *relational notion of time* entirely independent of any external clock. It is in this sense that one can interpret the relational quantum dynamics defined by temporal relational observables [18–25, 42–45, 80] or the Page-Wootters formalism [44, 45, 48, 78, 81–84] (which recently have been shown to be equivalent [44, 45]) in the context of laboratory situations.[12] Indeed, this is precisely how the experimental illustration of the Page-Wootters dynamics reported in [85] is to be understood.

In this manuscript, we focused purely on kinematical aspects of quantum reference frame physics. In forthcoming work [62], we will study in detail how the insights gained here are affected when we take the dynamics of the N-particle system into account. This question will link also with the perspective-neutral approach to QRFs [39, 41–45], and we will establish in detail the equivalence of its "quantum coordinate changes" with the QRF transformations exhibited here.

ACKNOWLEDGMENTS

We thank Thomas Galley, Leon Loveridge and Alexander R. H. Smith for many useful comments on an earlier draft version. MK is grateful to Časlav Brukner, Esteban Castro-Ruiz, and David Trillo Fernandez for helpful discussions. MPM would like to thank Anne-Catherine de la Hamette for inspiring discussions in the initial phase of this project. MK acknowledges the support of the Vienna Doctoral School in Physics (VDSP) and the support of the Austrian Science Fund (FWF) through the Doctoral Programme CoQuS. PAH is grateful for support from the Foundational Questions Institute under grant number FQXi-RFP-1801A. This work was supported in part by funding from Okinawa Institute of Science and Technology Graduate University. MK and MPM thank the Foundational Questions Institute and Fetzer Franklin Fund, a donor advised fund of Silicon Valley Community Foundation, for support via grant number FQXi-RFP-1815. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

[1] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, *Reference frames, superselection rules, and quantum information*, Rev. Mod. Phys. **79**, 555 (2007).

[2] A. R. H. Smith, *Communicating without shared reference frames*, Phys. Rev. **A 99**, 052315 (2019).

[3] I. Marvian, *Symmetry, Asymmetry and Quantum Information*, PhD thesis, University of Waterloo, 2012.

[4] G. Gour and R. W. Spekkens, *The resource theory of quantum reference frames: manipulations and monotones*, New J. Phys. **10**, 033023 (2008).

[5] G. Gour, I. Marvian, and R. W. Spekkens, *Measuring the quality of a quantum reference frame: The relative entropy of frameness*, Phys. Rev. A **80**, 012307 (2009).

[6] I. Marvian and R. W. Spekkens, *Modes of asymmetry: The application of harmonic analysis to symmetric quantum dynamics and quantum reference frames*, Phys. Rev. A **90**, 062110 (2014).

[7] M. C. Palmer, F. Girelli, and S. D. Bartlett, *Changing quantum reference frames*, Phys. Rev. A **89**, 052121 (2014).

[8] A. R. H. Smith, M. Piani, and R. B. Mann, *Quantum reference frames associated with noncompact groups: the case of translations and boosts, and the role of mass*, Phys. Rev. **A 94**, 012333 (2016)

[9] J. Åberg, *Catalytic Coherence*, Phys. Rev. Lett. **113**, 150402 (2014).

[10] M. Lostaglio, D. Jennings, and T. Rudolph, *Description of quantum coherence in thermodynamic processes requires constraints beyond free energy*, Nat. Comm. **6**, 6383 (2015).

[11] M. Lostalgio, K. Korzekwa, D. Jennings, and T. Rudolph, *Quantum Coherence, Time-Translation Symmetry, and Thermodynamics*, Phys. Rev. X **5**, 021001 (2015).

[12] These frameworks, however, also apply in the absence of external agents such as in quantum cosmology or gravity.

[12] M. Lostaglio and M. P. Müller, *Coherence and asymmetry cannot be broadcast*, Phys. Rev. Lett. **123**, 020403 (2019).

[13] I. Marvian and R. W. Spekkens, *No-Broadcasting Theorem for Quantum Asymmetry and Coherence and a Trade-off Relation for Approximate Broadcasting*, Phys. Rev. Lett. **123**, 020404 (2019).

[14] P. Erker, M. T. Mitchison, R. Silva, M. P. Woods, N. Brunner, and M. Huber, *Autonomous quantum clocks: does thermodynamics limit our ability to measure time?*, Phys. Rev. X **7**, 031022 (2017).

[15] P. Ćwikliński, M. Studziński, M. Horodecki, and J. Oppenheim, *Limitations on the Evolution of Quantum Coherences: Towards Fully Quantum Second Laws of Thermodynamics*, Phys. Rev. Lett. **115**, 210403 (2015).

[16] M P. Woods, R. Silva, and J. Oppenheim, *Autonomous quantum machines and finite-sized clocks* Ann. Henri Poincaré **20**, 125 (2019).

[17] M. P. Woods, and M Horodecki, *The resource theoretic paradigm of quantum thermodynamics with control*, arXiv:1912.05562 [quant-ph].

[18] C. Rovelli, *Quantum gravity*, Cambridge University Press, 2004.

[19] T. Thiemann, *Modern canonical quantum general relativity*, Cambridge University Press, 2007.

[20] J. Tambornino, *Relational Observables in Gravity: a Review,* SIGMA **8**, 017 (2012)

[21] C. Rovelli, *What Is Observable in Classical and Quantum Gravity?*, Class. Quant. Grav. **8**, 297 (1991)

[22] C. Rovelli, *Quantum reference systems*, Class. Quant. Grav. **8**, 317 (1991)

[23] C. Rovelli, *Time in Quantum Gravity: Physics Beyond the Schrödinger Regime*, Phys. Rev. D **43**, 442-456 (1991)

[24] B. Dittrich, *Partial and complete observables for Hamiltonian constrained systems*, Gen. Rel. Grav. **39**, 1891 (2007)

[25] B. Dittrich, *Partial and complete observables for canonical general relativity*, Class. Quant. Grav. **23**, 6155 (2006)

[26] L. Chataignier, *Construction of quantum Dirac observables and the emergence of WKB time*, Phys. Rev. D **101**, no.8, 086001 (2020).

[27] Y. Aharonov and L. Susskind, *Charge Superselection Rule*, Phys. Rev. **155**, 1428 (1967).

[28] Y. Aharonov and L. Susskind, *Observability of the Sign Change of Spinors under 2π Rotations*, Phys. Rev. **158**, 1237 (1967).

[29] Y. Aharonov and T. Kaufherr, *Quantum frames of reference*, Phys. Rev. D **30**, 368 (1984).

[30] E. Wigner, *Die Messung quantenmechanischer Operatoren*, Z. Phys. **133**, 101 (1952).

[31] H. Araki and M. M. Yanase, *Measurement of Quantum Mechanical Operators*, Phys. Rev. **120**, 622 (1960).

[32] M. M. Yanase, *Optimal Measuring Apparatus*, Phys. Rev. **123**, 666 (1961).

[33] L. Loveridge, B. Busch, and T. Miyadera, *Relativity of quantum states and observables*, EPL **117**, 40004 (2017).

[34] L. Loveridge, T. Miyadera, and P. Busch, *Symmetry, Reference Frames, and Relational Quantities in Quantum Mechanics*, Found. Phys. **48**, 135–198 (2018).

[35] T. Miyadera, L. Loveridge, and P. Busch, *Approximating relational observables by absolute quantities: a quantum accuracy-size trade-off*, J. Phys. A: Mathematical and Theoretical, **49**(18), 185301 (2016).

[36] L. Loveridge, *A relational perspective on the Wigner-Araki-Yanase theorem*, arXiv:2006.07047 [quant-ph].

[37] P. A. Höhn and M. P. Müller, *An operational approach to spacetime symmetries: Lorentz transformations from quantum communication*, New J. Phys. **18**, 063026 (2016).

[38] F. Giacomini, E. Castro-Ruiz, and Č. Brukner, *Quantum mechanics and the covariance of physical laws in quantum reference frames*, Nat. Comm. **10**, 494 (2019).

[39] A. Vanrietvelde, P. A. Höhn, F. Giacomini, and E. Castro-Ruiz, *A change of perspective: switching quantum reference frames via a perspective-neutral framework*, Quantum **4**, 225 (2020).

[40] A. de la Hamette and T. Galley, *Quantum reference frames for general symmetry groups*, arXiv:2004.14292.

[41] A. Vanrietvelde, P. A. Höhn, and F. Giacomini, *Switching quantum reference frames in the N-body problem and the absence of global relational perspectives*, arXiv:1809.05093 [quant-ph].

[42] P. A. Höhn and A. Vanrietvelde, *How to switch between relational quantum clocks*, arXiv:1810.04153 [gr-qc].

[43] P. A. Höhn, *Switching Internal Times and a New Perspective on the 'Wave Function of the Universe'*, Universe **5**, no.5, 116 (2019).

[44] P. A. Höhn, A. R. H. Smith and M. P. E. Lock, *The Trinity of Relational Quantum Dynamics*, arXiv:1912.00033 [quant-ph].

[45] P. A. Höhn, A. R. H. Smith and M. P. E. Lock, *Equivalence of approaches to relational quantum dynamics in relativistic settings*, arXiv:2007.00580 [gr-qc].

[46] F. Giacomini, E. Castro-Ruiz and Č. Brukner, *Relativistic Quantum Reference Frames: The Operational Meaning of Spin*, Phys. Rev. Lett. **123**, 090404 (2019)

[47] L. F. Streiter, F. Giacomini, and Č. Brukner, *A Relativistic Bell Test within Quantum Reference Frames*, arXiv:2008.03317 [quant-ph].

[48] E. Castro-Ruiz, F. Giacomini, A. Belenchia, and Č. Brukner, *Quantum clocks and the temporal localisability of events in the presence of gravitating quantum systems*, Nat. Comm. **11**, 2672 (2020).

[49] R. M. Angelo, N. Brunner, S. Popescu, A. J. Short, and P. Skrzypczyk, *Physics within a quantum reference frame*, J. Phys. A: Math. Theor. **44**, 145304 (2011).

[50] C. Rovelli, *Why Gauge?*, Found. Phys. **44**, 91–104 (2014).

[51] W. Donnelly and A. C. Wall, *Entanglement entropy of electromagnetic edge modes*, Phys. Rev. Lett. **114**(11), 111603 (2015)

[52] W. Donnelly and L. Freidel, *Local subsystems in gauge theory and gravity*, JHEP **09**, 102 (2016)

[53] M. Geiller and P. Jai-akson, *Extended actions, dynamics of edge modes, and entanglement entropy*, JHEP **20**, 134 (2020)

[54] L. Freidel, M. Geiller and D. Pranzetti, *Edge modes of gravity – I: Corner potentials and charges*, [arXiv:2006.12527 [hep-th]].

[55] H. Gomes and A. Riello, *Unified geometric framework for boundary charges and particle dressings*, Phys. Rev. D **98**, no.2, 025013 (2018)

[56] A. Riello, *Symplectic reduction of Yang-Mills theory with boundaries: from superselection sectors to edge modes, and back*, arXiv:2010.15894 [hep-th].

[57] S. Gielen and L. Menéndez-Pidal, *Singularity resolution depends on the clock*, Class. Quant. Grav. **37**, no.20, 205018 (2020).

[58] K. Giesel, L. Herold, B. F. Li and P. Singh, *Mukhanov-Sasaki equation in a manifestly gauge-invariant linearized cosmological perturbation theory with dust reference fields*, Phys. Rev. D **102**, no.2, 023524 (2020).

[59] K. Giesel, B. F. Li and P. Singh, *Towards a reduced phase space quantization in loop quantum cosmology with an inflationary potential*, arXiv:2007.06597 [gr-qc].
[60] P. A. Dirac, *Lectures on Quantum Mechanics*, Yeshiva University Press, 1964.
[61] M. Henneaux and C. Teitelboim, *Quantization of Gauge Systems*, Princeton University Press, 1992.
[62] M. Krumm, P. Höhn, and M. Müller, *Perspective-neutral approach to quantum reference frames for finite groups*, to appear.
[63] B. Simon, *Representations of Finite and Compact Groups*, American Mathematical Society, 1996.
[64] K. R. Davidson, *C^*-Algebras by Example*, American Mathematical Society, 1996.
[65] A. Savage, *Modern Group Theory*, lecture notes, University of Ottawa, 2017. Available at https://alistairsavage.ca/mat5145/notes/MAT5145-Modern_group_theory.pdf
[66] D. Giulini and D. Marolf, *A Uniqueness theorem for constraint quantization*, Class. Quant. Grav. **16**, 2489 (1999)
[67] D. Marolf, *Group averaging and refined algebraic quantization: Where are we now?*,, arXiv:gr-qc/0011112 [gr-qc].
[68] P. A. Höhn, *Reflections on the information paradigm in quantum and gravitational physics*, J. Phys. Conf. Ser. **880** no.1, 012014 (2017).
[69] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, New York, 2010.
[70] M. Tinkham, *Group Theory and Quantum Mechanics*, Dover Publications, 1992.
[71] T. P. Le, P. Mironowicz, and P. Horodecki, *Blurred quantum Darwinism across quantum reference frames*, arXiv:2006.06364 [quant-ph].
[72] J. Tuziemski, *Decoherence and information encoding in quantum reference frames*, arXiv:2006.07298 [quant-ph].
[73] M. F. Savi, and R. M. Angelo, *Quantum Resources Covariance*, arXiv:2005.09612 [quant-ph].
[74] J. M. Yang, *Switching Quantum Reference Frames for Quantum Measurement*, Quantum **4**, 283 (2020).
[75] L. Hardy, *Implementation of the Quantum Equivalence Principle*, arXiv:1903.01289.
[76] P. A. Guérin and Č. Brukner, *Observer-dependent locality of quantum events*, New J. Phys. **20**, 103031 (2018).
[77] S. Khandelwal, M. P. E. Lock and M. P. Woods, *Universal quantum modifications to general relativistic time dilation in delocalised clocks*, Quantum **4** 309 (2020).
[78] A. R. H. Smith and M. Ahmadi, *Quantum clocks observe classical and quantum time dilation*, Nat. Comm. **11**, no.1, 5360 (2020)
[79] P. T. Grochowski, A. R. H. Smith, A. Dragan and K. Debski, *Quantum time dilation in atomic spectra*, arXiv:2006.10084 [quant-ph].
[80] R. Gambini and J. Pullin, *The Montevideo Interpretation: How the inclusion of a Quantum Gravitational Notion of Time Solves the Measurement Problem*, arXiv:2010.14519.
[81] D. N. Page, and W. K. Wootters, *Evolution without evolution: Dynamics described by stationary observables*, Phys. Rev. **D 27**, 2885 (1983).
[82] W. K. Wootters, *"Time" replaced by quantum correlations*, Int. J. Theor. Phys. **23**, 701 (1984).
[83] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum time*, Phys. Rev. D **79**, 945933 (2015).
[84] A. R. H. Smith and M. Ahmadi, *Quantizing time: interacting clocks and systems*, Quantum **3**, 160 (2019).
[85] E. Moreva, G. Brida, M. Gramegna, V. Giovannetti, L. Maccone, and M. Genovese, *Time from quantum entanglement: An experimental illustration*, Phys. Rev. A **89**, 052122 (2014).

Chapter 5

Semi-device-independent information processing with spatiotemporal degrees of freedom

Authors: Andrew J. P. Garner, Marius Krumm, Markus P. Müller
Published in: Phys. Rev. Research **2**, 013112 (2020)
DOI: https://doi.org/10.1103/PhysRevResearch.2.013112
Preprint available as: arXiv:1907.09274
https://arxiv.org/abs/1907.09274

Contribution: The authors are in alphabetical order. I made essential contributions to all parts of the research. I provided the leading input for the local hidden variable models.

Abstract: Nonlocality, as demonstrated by the violation of Bell inequalities, enables device-independent cryptographic tasks that do not require users to trust their apparatus. In this article, we consider devices whose inputs are spatiotemporal degrees of freedom, e.g. orientations or time durations. Without assuming the validity of quantum theory, we prove that the devices' statistical response must respect their input's symmetries, with profound foundational and technological implications. We exactly characterize the bipartite binary quantum correlations in terms of local symmetries, indicating a fundamental relation between spacetime and quantum theory. For Bell experiments characterized by two input angles, we show that the correlations are accounted for by a local hidden variable model if they contain enough noise, but conversely must be nonlocal if they are pure enough. This allows us to construct a "Bell witness" that certifies nonlocality with fewer measurements than possible without such spatiotemporal symmetries, suggesting a new class of semi-device-independent protocols for quantum technologies.

I. INTRODUCTION

Quantum theory radically challenges our classical intuitions. A famous example is provided by the violation of Bell inequalities [1–6], demonstrating that local hidden variable models are inadequate to account for all observable correlations in quantum theory. While this so-called *nonlocality* was initially of foundational concern, it transpires to have a very powerful practical use: it enables *device-independent* protocols in quantum information theory (e.g. [7–10]). In this paradigm, one can perform certain tasks (e.g. cryptography) without trusting one's apparatus, or even necessarily assuming the full formalism of quantum mechanics. These protocols rely on the readily believable *no-signalling* constraint, which forbids the instantaneous transmission of information between sufficiently distant laboratories. Since this constraint originates in special relativity, it may be thought of as a property of spacetime itself.

A pillar of the device-independent formalism is its abstract *black box* description: experimental devices are fully characterized by probability tables of outputs given a supplied input (figure 1a). In this article, we supplement these inputs with physical structure, and adopt a semi-device-independent approach that makes no assumptions about the inner workings of the devices, or the physical theories governing them (i.e. quantum or otherwise), but assumes that their ensemble statistics can be characterized by a finite number of parameters. Specifically, we consider when inputs are *spatiotemporal* degrees of freedom, e.g. some orientation in space or duration of time. This includes, for example, the bias of a magnetic field, duration of a Rabi pulse, or angle of a polarizer (figure 1b). Spatiotemporal degrees of freedom bring with them a symmetry structure, which can be mathematically described using Lie group theory.

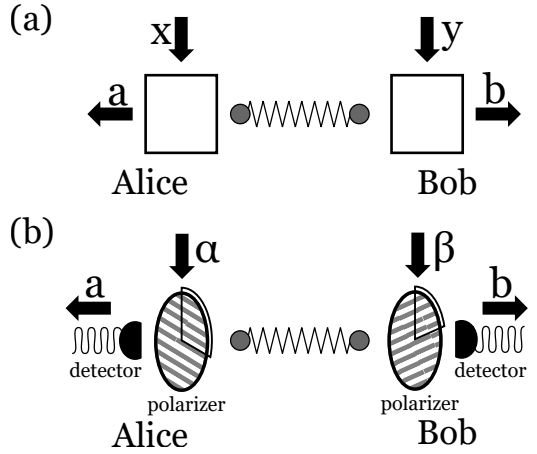


FIG. 1: **Bell scenario: abstract vs. spatiotemporal inputs.** Spatially-separate Alice and Bob independently choose measurement settings x, y and receive some outputs a, b, yielding the joint conditional probability distribution $\mathrm{P}(a,b\,|\,x,y)$. (a) In the usual black box formalism, the inputs x and y are abstract labels. (b) Here, we consider the physical situation where the inputs are spatiotemporal degrees of freedom (e.g. angles $x=\alpha$, $y=\beta$ of polarizers).

In this article, we introduce a general framework for spatiotemporal black boxes. We prove that the probability tables associated with spatiotemporal inputs must encode a linear representation of the corresponding symmetry groups (section II A). We demonstrate the power of this approach with two examples in Bell test scenarios: First, if each laboratory controls a single angle (section II B), we find—independently of the theory—that the response to rotations can in some cases certify the existence of a local hidden-variable model, or the violation of a Bell inequality. Consequently, we present a novel protocol for witnessing nonlocality, similar in spirit to [11, 12], but without prerequiring the validity of quantum theory. Secondly, we consider when both inputs are chosen via rotations in d-dimensional space. We show that natural assumptions on the local response to those rotations recovers the set of bipartite binary quantum correlations exactly (section II C), indicating a fundamental relation between the structures of spacetime and of quantum mechanics. Finally, we discuss the implications of these results (section III), particularly for the construction of novel experimental tests of quantum mechanics and of new semi-device-independent protocols for quantum technologies.

II. RESULTS

A. Representation theorem for spatiotemporal degrees of freedom

The device-independent formalism abstracts experiments into a table of output statistics conditional on some choice of input. This is imbued with *causal structure* [13] by separating the inputs and outputs into local choices and responses made and observed by different *local agents*, acting in potentially different locations and times. The simplest structure is one agent at a single point in time. More commonly considered is the *Bell scenario* [6], where two spatially separated agents each independently select an input (measurement choice) and record the resulting local output. Theorem 1 of this paper applies to *any* casual structure, but looking towards application the later examples will use the Bell scenario.

Here, we shall consider experiments where the local inputs correspond to *spatiotemporal* degrees of freedom: for example, the direction of inhomogeneity of the magnetic field in a Stern–Gerlach experiment, or the angle of a polarization filter (figure 1b). Crucially, we will describe such experiments without assuming the validity of quantum mechanics.

Let us first consider a single laboratory, say, Alice's. For concreteness, assume for the moment that Alice's input is given by the direction $\vec{x}$ of a magnetic field. She chooses her input by applying a rotation $R\in\mathrm{SO}(3)$ to some initial magnetic field direction $\vec{x}_0$, i.e. $\vec{x}=R\vec{x}_0$. Her statistics of obtaining any outcome a will now depend on this direction, giving her a *black box* $\mathrm{P}(a\,|\,\vec{x})$.

In general, Alice will have a set of inputs $\mathcal{X}$ and a symmetry group $\mathcal{G}$ that acts on $\mathcal{X}$. Given

some arbitrary $x_0 \in \mathcal{X}$, we assume that Alice can generate every possible input $x \in \mathcal{X}$ by applying a suitable transformation $R \in \mathcal{G}$, such that $x = Rx_0$. Mathematically, $\mathcal{X}$ is then a *homogeneous space* [14], which can be written $\mathcal{X} = \mathcal{G}/\mathcal{H}$, where $\mathcal{H} \subseteq \mathcal{G}$ is the subgroup of transformations R' with $R'x_0 = x_0$. In the example above, $\mathcal{G} = \mathrm{SO}(3)$ describes the full set of rotations that Alice can apply to $\vec{x}_0$, while $\mathcal{H} = \mathrm{SO}(2)$ describes the subset of rotations that leave $\vec{x}_0$ invariant (i.e. the axial symmetry of the magnetic field vector). Then, $\mathcal{X} = \mathrm{SO}(3)/\mathrm{SO}(2) = S^2$ is the 2-sphere of unit vectors (i.e. directions) in 3-dimensional space. Similarly, the polarizer (figure 1b) corresponds to $\mathcal{G} = \mathrm{SO}(2)$, $\mathcal{H} = \{\mathbb{1}\}$, and $\mathcal{X} = S^1$, which we identify with the unit circle.

Temporal symmetries also fit into this formalism. Suppose Alice's input corresponds to letting her system evolve for some time, then $\mathcal{G} = (\mathbb{R}, +)$ is the group of *time translations*. If we know that the system evolves periodically over intervals $\tau \in \mathbb{R}^+$, which we model as a symmetry subgroup $\mathcal{H} = (\tau \cdot \mathbb{Z}, +)$, then the input domain $\mathcal{X} = \mathcal{G}/\mathcal{H} \simeq S^1$. Physically, this could correspond to applying a controlled-duration Rabi pulse to an atomic system of trusted periodicity before recording an outcome.

Now suppose Alice has a black box P, where on spatiotemporal input $x \in \mathcal{X}$, the outcome a is observed with probability $\mathrm{P}(a\,|\,x)$. Then, Alice can "rotate" her apparatus by $R \in \mathcal{G}$, and induce a new black box P′ with outcome probabilities $\mathrm{P}'(a|x) = \mathrm{P}(a\,|Rx)$. Physically, R could be an active rotation within Alice's laboratory (e.g. spinning a polarizer), of the incident system (e.g. adding a phase plate), or could be a passive change of coordinates.

Thus, a given black box and a spatiotemporal degree of freedom defines a family of black boxes, and transformations $R \in \mathcal{G}$ map a given black box to another one in this family. Suppose we denote the action of R on the black boxes by $T_R : \mathrm{P} \mapsto \mathrm{P}'$. If rotating the input first by R then by R' is equivalent to a single rotation $R'' = R' \circ R$, it follows the black box formed by applying T_R and then $T_{R'}$ is equivalent to applying the single transformation $T_{R''} = T_{R'} \circ T_R$ on P. We can say more about this action if we consider ensembles of black boxes. For any family of black boxes $\{\mathrm{P}_i\}_{i=1}^n$ and probabilities $\{\lambda_i\}_{i=1}^n$, $\sum_i \lambda_i = 1$, $\lambda_i \geq 0$, the experiment of first drawing i with probability λ_i and then applying black box P_i defines a new, effective black box P, with statistics $\mathrm{P}(a\,|\,x) = \sum_i \lambda_i \mathrm{P}_i(a\,|\,x)$. All these black boxes are in principle operationally accessible to Alice. However, a priori, we cannot say much about the resulting set of boxes – it could be a complicated uncountably-infinite-dimensional set defying simple analysis. Thus, we make a minimal assumption that this set is not "too large":

Assumption (i). *Ensembles of black boxes can be characterized by a finite number of parameters.*

The mathematical consequence is that the space of possible boxes for Alice is finite-dimensional. This is a weaker abstraction of a stronger assumption typically made in the *semi-device-independent* framework of quantum information: that the systems involved in the protocols are described by Hilbert spaces of bounded (usually small) dimension [15, 16]. For example, BB84 [17] quantum cryptography assumes that the information carriers are two-dimensional, excluding additional degrees of freedom that could serve as a side channel for eavesdroppers [18]. Assumption (i) is much weaker; it does not presume that we have Hilbert spaces in the first place. It is for this assumption (and not the spatiotemporal structure of the input space) that the results presented in this article lie in the *semi*-device-independent regime.

We thus arrive at our first theorem. Recall that Alice chooses her input $x_R \in \mathcal{X}$ by selecting some $R \in \mathcal{G}$ and applying it to a default input x_0, i.e. $x = Rx_0$. Then:

Theorem 1. *There is a representation of the symmetry group $\mathcal{G}$ in terms of real orthogonal matrices $R \mapsto T_R$, such that for each outcome a, the outcome probabilities $\mathrm{P}(a\,|\,x_R)$ are a fixed (over R) linear combination of matrix entries of T_R.*

The proof is given in appendix A, and is based on the observation that T_R becomes a linear group representation on the space of ensembles. Motivated by this characteristic response, we refer to black boxes whose inputs are selected through the action of $\mathcal{G}$ as *$\mathcal{G}$–boxes.*

A few comments are in order. First, this theorem applies to any causal structure, including the case of two parties performing a Bell experiment. If Alice and Bob have inputs and transformations $\mathcal{X}_A$, $\mathcal{G}_A$ and $\mathcal{X}_B$, $\mathcal{G}_B$ respectively, then the full setup can be seen as an experiment with $\mathcal{X} = \mathcal{X}_A \times \mathcal{X}_B$ and $\mathcal{G} = \mathcal{G}_A \times \mathcal{G}_B$, to which Theorem 1 applies directly.

Secondly, there may be more than one transformation that generates the desired input x, i.e. both $x = Rx_0$ and $x = R'x_0$ for $R \neq R'$; this is precisely the case if $R^{-1}R' \in \mathcal{H}$. For example, a magnetic field can be rotated from the y- to z-direction in many different ways. In this case, Theorem 1 applies to both R and R', which yields additional constraints.

Finally, quantum theory is contained as a special case. Typically, one argues that due to preservation of probability, transformations R must be represented in quantum mechanics via unitary matrices U_R acting on density matrices via $\rho \mapsto U_R \rho U_R^\dagger$. This projective action can be written as an orthogonal matrix on the real space of Hermitian operators, in concordance with Theorem 1.

As a specific example, consider a quantum harmonic oscillator with frequency ω, initially in state ρ_0, left to evolve for a variable time t before it is measured by a fixed POVM [19] $\{M_a\}_{a \in \mathcal{A}}$. The free dynamics are given by the Hamiltonian H, whose discrete set of eigenvalues $\{E_n = \hbar\omega\left(\frac{1}{2}+n\right)\}$ correspond to allowed "energy levels". The evolution is periodic, so (recalling earlier) $\mathcal{G} = (\mathbb{R}, +)$, $\mathcal{H} = \left(\frac{2\pi}{\omega} \cdot \mathbb{Z}, +\right)$ and $\mathcal{X} \simeq S^1$. The associated black box is thus $\mathrm{P}(a\,|\,t) = \mathrm{Tr}\left[M_a \exp\left(-\frac{iHt}{\hbar}\right) \rho_0 \exp\left(\frac{iHt}{\hbar}\right)\right]$. For any given ρ_0 and M_a, this evaluates to an affine-linear

combination of terms of the form $\cos[(n-m)\hbar\omega t]$ and $\sin[(n-m)\hbar\omega t]$, involving all pairs of energy levels that have non-zero occupation probability in ρ_0 (and non-zero support in M_a). This is a linear combination of entries of the matrix representation

$$T_t = \bigoplus_{\alpha=E_n-E_m} \begin{pmatrix} \cos(\alpha t) & \sin(\alpha t) \\ -\sin(\alpha t) & \cos(\alpha t) \end{pmatrix}, \tag{1}$$

in accordance with Theorem 1. For T_t to be a finite matrix, there must only be a finite number of occupied energy differences $E_m - E_n$.

Here, Assumption (i) is equivalent to an upper (and lower) bound on the system's energy. In the general framework that does not assume the validity of quantum mechanics (or presuppose trust in our devices, or our assignment of Hamiltonians), we can view Assumption (i) as a natural generalization of this to other symmetry groups and beyond quantum theory. By assuming a concrete upper bound on the representation label (such as α in eq. (1)), we can establish powerful theory- and device-independent consequences for the resulting correlations, as we will now demonstrate by means of several examples.

B. Example: Two angles and Bell witnesses

Let us consider the simplest non-trivial spatiotemporal freedom, where Alice and Bob each have the choice of a single continuous angle: respectively $\alpha, \beta \in [0, 2\pi)$, and each obtain a binary output $a, b \in \{+1, -1\}$. Physically, this would arise, say, in experiments where a pair of photons is distributed to the two laboratories, each of which contains an rotatable polarizer followed by a photodetector (figure 1b).

Due to Theorem 1, the probabilities $\mathrm{P}(a, b \,|\, \alpha, \beta)$ are linear combinations of matrix entries of an orthogonal representation of $\mathrm{SO}(2)\times\mathrm{SO}(2)$. From the classification of these representations (see appendix B 1), it follows that all $\mathrm{SO}(2)\times\mathrm{SO}(2)$-boxes are of the form

$$\mathrm{P}(a, b \,|\, \alpha, \beta) := \sum_{m=0}^{2J} \sum_{n=-2J}^{2J} c_{mn}^{ab} \cos(m\alpha - n\beta) + s_{mn}^{ab} \sin(m\alpha - n\beta), \tag{2}$$

resulting in a correlation function

$$\begin{aligned} C(\alpha, \beta) &:= \mathrm{P}(+1, +1 \,|\, \alpha, \beta) + \mathrm{P}(-1, -1 \,|\, \alpha, \beta) \\ &\quad - \mathrm{P}(+1, -1 \,|\, \alpha, \beta) - \mathrm{P}(-1, +1 \,|\, \alpha, \beta) \qquad (3) \\ &= \sum_{m=0}^{2J} \sum_{n=-2J}^{2J} C_{mn} \cos(m\alpha - n\beta) + S_{mn} \sin(m\alpha - n\beta), \qquad (4) \end{aligned}$$

where $J \in \{0, \frac{1}{2}, 1, \frac{3}{2}, \ldots\}$ is some finite maximum "spin".

If Alice and Bob's laboratories are spatially separated, the laws of relativity forbid Alice from sending signals to Bob instantaneously. This "no-signalling" principle constrains the set of valid joint probability distributions: namely Bob's marginal statistics cannot depend on Alice's choice of measurement, and vice versa. However, for *any* given correlation function of the form eq. (4), there is always at least one set of valid no-signalling probabilities (see appendix B 2) – for example, those where the marginal distributions are "maximally mixed" such that independent of α, a is $+1$ or -1 with equal probability (likewise for β and b), consistent with an observation of Popescu and Rohrlich [20].

Consider a quantum example: two photons in a Werner state [21, 22] $\rho_W := p|\psi^-\rangle\langle\psi^-| + \frac{1}{4}(1-p)\mathbb{1}_4$ where $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle - |1\rangle|0\rangle)$ and $p \in [0, 1]$. Alice and Bob's polarizer/detector setups are described by the observables $M_\theta := \left(\begin{smallmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{smallmatrix}\right)$ for orientations $\theta = \alpha, \beta$ respectively. Then, $C(\alpha, \beta) = \mathrm{Tr}(\rho_W M_\alpha \otimes M_\beta) = -p\cos[2(\alpha - \beta)]$. This fits the form of eq. (4) for $J = 1$, with $C_{22} = -p$ and all other coefficients as zero.

A paradigmatic question in this setup is whether the statistics can be explained by a *local hidden variable* (LHV) model. Namely, is there a single random variable λ over some space Λ such that $\mathrm{P}(a, b \,|\, \alpha, \beta) = \int_\Lambda \mathrm{d}\lambda\, \mathrm{P}_\Lambda(\lambda)\, \mathrm{P}_A(a \,|\, \alpha, \lambda)\, \mathrm{P}_B(b \,|\, \beta, \lambda)$, where $\mathrm{P}_\Lambda(\lambda)$ is a classical probability distribution, and $\mathrm{P}_A(a \,|\, \alpha, \lambda)$ and $\mathrm{P}_B(b \,|\, \beta, \lambda)$ are respectively Alice and Bob's local response functions (conditioned on their input choices α and β and the particular realization of the hidden variable λ)? If no LHV model exists, then the scenario is said to be *nonlocal*. Famously, Bell's theorem shows that quantum theory admits correlations that are nonlocal in this sense [1, 2]. This follows from the violation of *Bell inequalities* that are satisfied by all distributions with LHV models, the archetypical example being the Clauser–Horne–Shimony–Holt (CHSH) inequality [3]:

$$\big|C(\alpha_1, \beta_2) + C(\alpha_3, \beta_2) + C(\alpha_3, \beta_4) - C(\alpha_1, \beta_4)\big| \leq 2, \tag{5}$$

where α_1, α_3 are two choices of Alice's angle, and β_2, β_4 of Bob's. Classical systems always satisfy this bound, but quantum theory admits states and measurements that violate it. When working with a continuous parameter, Bell inequalities need not be limited to a subset of angles, but can also be formulated as a *functional* of the entire correlation function [23, 24].

Not all correlations of the form in eq. (4) are allowed by quantum theory. For example, "science fiction" polarizers with the correlation function $C(\alpha, \beta) = \frac{2}{7}\cos[3(\alpha - \beta)] - \cos[\alpha - \beta]$ would yield a CHSH value of 3.63, under choices of angles $\alpha_1 = 1.5$, $\alpha_3 = 0$, $\beta_2 = 3.9$ and $\beta_4 = 2.3$, violating quantum theory's maximum achievable value of $2\sqrt{2}$ [25].

With this general form, we can make broad statements about whether correlations are local or nonlocal. First, if the correlations are sufficiently "noisy", we can systematically construct a LHV model by generalizing a procedure by Werner [21] (see appendix B 3). If the only constraint on the correlations is that is has some maximum J, then the existence of a LHV is guaranteed if the magnitude of angle-dependent changes in C is less than γ_J where

$$\gamma_J := \sqrt{2} e^{-1} [4J(2J+1)]^{-\frac{3}{2}}. \tag{6}$$

Subject to extra restrictions that keep the form of C simple, more permissive bounds are also derived. For instance, if there is only one non-zero coefficient in eq. (4), then $\gamma = \sqrt{2}/\pi \approx 0.4502$. Recall the correlation function for projective measurements on a Werner state, $-p\cos[2(\alpha-\beta)]$, and identify γ with p. In this case, our bound is comparable with that in Hirsch *et al.* [26] of $p \leq 0.6829$.

Conversely, we can give a simple sufficient criterion for nonlocality if we separate the terms in eqs. (2) and (4) into *relational* and *non-relational* components. The relational components where $m = n$ account for behaviour that depends only on the difference between the two angles. Purely relational correlations, i.e. ones with $C(\alpha,\beta) \equiv C(\alpha-\beta)$, can be motivated by symmetry (i.e. that in the absence of external references, *only* the relative angle should have operational meaning). Here, the $J = \frac{1}{2}$ case contains the bipartite *rotational invariant* correlations discussed in Nagata *et al.* [27]. Conversely, the correlations resulting from any experiment can be actively made relational as we will describe in more detail below.

If the relational part of a correlation function C_{rel} has an angle difference Θ_+ which results in near perfect (anti-)correlations, and another angle difference Θ_- that does not, then one can systematically construct a (Braunstein–Caves [28]) Bell inequality that will be violated (see appendix B 4). Specifically, "near perfect" means that for a given J, $C_{\text{rel}}(\Theta_+) \geq 1 - \varepsilon_J$ with

$$\varepsilon_J := -K_J + \sqrt{K_J^2 + \frac{\Delta^2}{4}} = \frac{\Delta^2}{8K_J} + \mathcal{O}\big(K_J^{-2}\big)\,, \qquad (7)$$

where $K_J := \sqrt{2}\pi^2 J(2J+1)(4J+1)/3$, and $C_{\text{rel}}(\Theta_-) \leq 1 - \Delta$ bounds the "other" value measured at Θ_-. (See appendix B 4 for proof).

We summarize these results (see also figure 2):

Theorem 2. *Consider a two-angle Bell experiment with correlations C in the form of eq. (4), with an upper bound J on the representation labels.*

A. If C is sufficiently "noisy", in the sense that

$$\max_{\alpha,\beta} |C(\alpha,\beta) - C_{00}| \leq \gamma_J(1 - |C_{00}|) \qquad (8)$$

with γ_J as in eq. (6), then the correlations can always be exactly accounted for by a LHV model.

B. If the relational part of C is sufficiently "pure" for some angle Θ_+ (above $1-\varepsilon_J$, as defined in eq. (7)), but also sufficiently different (below $1-\Delta$) for some other angle Θ_-, then the correlations violate a Bell inequality.

This is a powerful result: with a choice between two experimental settings for Alice, and no choice made by Bob, we can *witness* nonlocality. This can be done by the following protocol:

- Alice and Bob share some random angle λ, uniformly distributed in the interval $[0, 2\pi)$.
- Alice chooses locally freely between the two possible angles $\alpha \in \{\Theta_+, \Theta_-\}$.
- Alice now inputs $\alpha + \lambda$ into her half of the box, while Bob inputs λ.
- By repeating the protocol, they determine the correlations $C_{\text{rel}}(\Theta_+)$ and $C_{\text{rel}}(\Theta_-)$, and verify that they violate the inequality above.

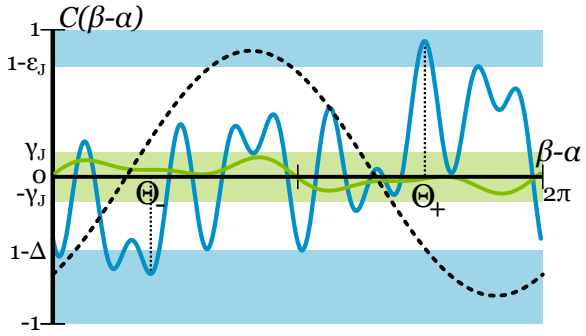


FIG. 2: **Two-angle relational correlation functions.** A "sufficiently noisy" correlation function can always be reproduced exactly by a LHV model (Theorem 2a). This is represented by the green curve completed contained within the central green-shaded region (drawn for $C_{00} = 0$). Conversely, if the function is "pure enough", then it must be nonlocal (Theorem 2b). This is represented by the blue curve with values in both extremal blue-shaded regions. Not all curves can be realized within quantum theory, but simple sinusoidal curves certainly can (such as the dashed black curve), following from Theorem 3 in two dimensions.

Randomization over λ effectively projects C onto its relational part $C_{\text{rel}}(\alpha,\beta) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\lambda\, C(\alpha+\lambda, \beta+\lambda)$, which only depends on $\alpha - \beta$. The protocol above fixes β to zero, while $\alpha \in \{\Theta_+, \Theta_-\}$. This is sufficient to determine the two correlation values.

The protocol assumes that Alice and Bob have some physically motivated promise on the maximum representation label J (e.g. by assuming an upper bound on the total energy of the system, or the number of elementary particles transmitted), and that they know the angles Θ_+ and Θ_- beforehand. The latter assumption is analogous to standard Bell experiments, where the relevant measurement settings are assumed to be known.

Witnessing Bell nonlocality is not the same as directly demonstrating nonlocality (i.e. collecting all the statistics for a Bell test, which is only possible if Bob has some free choice too) but rather, subject to Assumption (i), implies the existence of an experiment that *would* demonstrate nonlocality. In contrast to a full Bell experiment, a Bell witness has the advantage of being experimentally easier to implement: the protocol above allows one to witness nonlocality with only *two* measurement settings instead of four. Note that *only* making the correlation function relational (i.e. going from C to C_{rel} as above) *without* any additional assumption on J is *not* sufficient to obtain this reduction, as we show in appendix B 5.

Our protocol hence demonstrates that natural assumptions on the response of devices to spatiotemporal transformations can give additional constraints that allow for the construction of new Bell witnesses. This opens up the possibility of new methods of experimentally certifying nonlocal behaviour, similar to [11, 12, 29], but without the need to presume the validity of quantum theory or to trust all involved measurement devices.

Theorem 2 shows us that smaller values of J (and hence "simpler" responses to changes in angles) result in more permissive bounds for finding LHV models, or witnessing non-locality. In our next example, we shall move from angles (SO(2)) to directions (SO(d)), but consider arguably the simplest non-trivial response.

C. Example: Characterizing quantum correlations

For our last example, we shall apply our framework to characterize the set of correlations that can be realized by two parties sharing a quantum state, each locally choosing one of two binary-outcome measurements – the thus called quantum "(2,2,2)"-behaviours. The set of quantum $(2,2,2)$-behaviours $\mathcal{Q}$ is a strict superset of the classical $(2,2,2)$-behaviours $\mathcal{C}$ (i.e. those admitting a LHV model). However, the set of all no-signalling behaviours $\mathcal{NS}$ is strictly larger: $\mathcal{C} \subsetneq \mathcal{Q} \subsetneq \mathcal{NS}$ [20, 30]. This has led to the search for simple physical or information-theoretic principles that would explain "why" nature admits no more correlations than in $\mathcal{Q}$. Several candidates have been suggested over the years, including *information causality* [31], *macroscopic locality* [32], or *non-trivial communication complexity* [33], but none of these have been able to single out $\mathcal{Q}$ uniquely [34].

Here, we will provide such a characterization by considering black boxes that transform in arguably the simplest manner. Over a spherical input domain $\mathcal{X} = S^{d-1}$ an SO(d)–box $\mathrm{P}(a\,|\,\vec{x})$ is said to *transform fundamentally* if the representation matrix T_R in Theorem 1 can be chosen as the block matrix $\mathbb{1}_1 \oplus R$, where $\mathbb{1}_1 := (1)$ and R is the fundamental representation of SO(d) (e.g. for $d = 3$, $\{R\}$ are the familiar rotation matrices). Consequently, a black box that transforms fundamentally has an *affine* representation, $\mathrm{P}(a\,|\,\vec{x}) = c_0^a + \vec{c}^{\,a} \cdot \vec{x}$ where $\vec{x} \in S^{d-1}$ is the input, and $c_0^a \in \mathbb{R}^+$, $\vec{c}^{\,a} \in \mathbb{R}^d$ (proof in appendix C).

Motivated by symmetry, we consider a class of *unbiased* black boxes that do not prefer any particular output when averaged over all possible inputs. This implies that $c_0^a = 1/|\mathcal{A}|$ for every a. For example, this symmetry holds for measurements on quantum spin-$\frac{1}{2}$ particles: spin $+\frac{1}{2}$ in one direction is the same as spin $-\frac{1}{2}$ in the opposite, and hence neither outcome is preferred on average.

Imagine Alice and Bob residing in d-dimensional space ($d \geq 2$), sharing a non-signalling box $\mathrm{P}(a, b\,|\,\vec{x}, \vec{y})$, where both inputs $\vec{x}, \vec{y} \in S^{d-1}$ are spatial directions, and a, b each can take two values. Suppose that their *conditional boxes* transform fundamentally and are unbiased. A conditional box $\mathrm{P}_A^{b,\vec{y}}(a|\vec{x}) := \mathrm{P}(a, b\,|\,\vec{x}, \vec{y})\,/\mathrm{P}_B(b\,|\,\vec{y})$ describes the local black box Alice would have if she was told Bob's measurement choice $\vec{y}$ and outcome b. If all conditional boxes for Alice and Bob transform fundamentally, then the bipartite box is said to *transform fundamentally locally*. Similarly, if all conditional boxes are unbiased, $\mathrm{P}(a, b\,|\,\vec{x}, \vec{y})$ is said to be *locally unbiased*.

Surprisingly, these local symmetries severely constrain the *global* correlations: they allow for only and exactly those correlations that can be realized by two parties who share a *quantum* state and choose between two possible two-outcome quantum measurements each—the *quantum* $(2,2,2)$*-behaviours*:

Theorem 3. *The quantum* $(2,2,2)$*-behaviours are exactly those that can be realised by binary-outcome bipartite* SO(d)×SO(d)*-boxes that* transform fundamentally locally *and are* locally unbiased*, restricted to two choices of input direction per party per box, and statistically mixed via shared randomness.*

The proof is given in appendix C.

A few remarks are in place. First, the *unbiasedness* refers to the total set S^{d-1} of possible inputs per party, not to the two inputs to which the box is restricted. Even if the unrestricted behaviour is unbiased in the sense described above, the resulting $(2,2,2)$-behaviour can be biased. Secondly, this unbiasedness of the underlying $\mathrm{SO}(d) \times \mathrm{SO}(d)$-box is necessary to recover the quantum correlations – without it, one can realize arbitrary nonsignalling correlations, including PR–box behaviour, in a way that still transforms locally fundamentally (we give an example in appendix C). Finally, shared randomness is necessary to realize explicitly *non-extremal* quantum correlations by such boxes, following on the observation that the set of $(2,2,2)$–behaviours realizable by POVMs on two qubits is not convex [35, 36]. Namely, if both parties share the $(2,2,2)$-behaviours P_0 and P_1 and a random bit $c \in \{0,1\}$ that equals 0 with probability λ, they can statistically implement the mixed behaviour $\lambda P_0 + (1-\lambda)P_1$ by feeding their inputs into box P_c.

For $n=2$ parties with $m=2$ measurements and $k=2$ outcomes each, our result provides a characterization of the quantum set. Although Theorem 3 cannot be extended to general (m,n,k)-behaviours [37] without modification, this result shows that our framework of $\mathcal{G}$-boxes offers a very natural perspective on physical correlations, and reinforces earlier observations that hint at a deep fundamental link between the structures of spacetime and quantum theory [38–41].

III. DISCUSSION AND OUTLOOK

We have introduced a general framework for semi-device-independent information processing, without assuming quantum mechanics, for black boxes whose inputs are degrees of freedom that break spatiotemporal symmetries. Such black boxes have characteristic probabilistic responses to symmetry transformations,

and natural assumptions about this behaviour can certify technologically important properties like the presence or absence of Bell correlations.

Specifically, we have shown that the quantum $(2, 2, 2)$-behaviours can be exactly classified as those of bipartite boxes that transform *locally* in the simplest possible way – by the fundamental representation of $\mathrm{SO}(d)$ rotations, respecting the unbiasedness of outcomes. For Bell experiments with $\mathrm{SO}(2) \times \mathrm{SO}(2)$-boxes, we have shown that correlations that are quantifiably "noisy enough" always admit a local hidden variable model, whereas relational correlations for which there are settings with differing "purity" must violate a Bell inequality. Since the underlying technical tools (e.g. Schur orthogonality [42]) hold in greater generality, many of our results could be applied to other groups.

Furthermore, these results have allowed us to construct a protocol to *witness* the violation of a Bell inequality within a causal structure that is otherwise too simple to admit the direct detection of nonlocality. We believe that our approach can be applied to experimental settings, such as the recent demonstration of Bell correlations in a Bose–Einstein condensate [12], and potentially eliminate the necessity to trust all detectors or to assume the exact validity of quantum mechanics. Many of these experiments *do* work with spatiotemporal inputs like Rabi pulses, which makes our approach particularly natural for analyzing them.

We have predominantly worked under the assumption that ensembles of black boxes are characterized by a *finite* number of parameters, and – more specifically – that an upper bound on the representation label (say, the "spin" J) of the boxes is known. On one hand, this assumption can likely be weakened, by employing group-theoretic results such as the Peter–Weyl theorem [42]. On the other hand, we have argued that this assumption is natural: it is weaker than assuming a Hilbert space with bounded dimension (standard in the semi-device-independent framework [16]) and constitutes a generalization of an "energy bound" beyond quantum theory (cf. [43]). Moreover, it incorporates an intuition conceptually closer to particle physics: to quantify the potential eavesdropping side channels, one might not count Hilbert space dimensions, but rather representation labels, since these are intuitively (and sometimes rigorously) related to the total number of particles.

Our framework opens up several potential avenues for future work. First, as the witness example demonstrates, our formalism hints at novel semi-device-independent protocols based on assumptions with firmer physical motivation than the usual dimension bounds. In contrast to recent proposals for using energy bounds [44–46], our assumption on the devices' symmetry behaviour does not presume the validity of quantum mechanics, but rather embodies a natural upper bound to the "fine structure" of the devices' response. Meanwhile, one might apply the functional approach [23, 24] to our framework by taking Haar integrals over spatiotemporal input spaces to derive a device–independent family of generalized Bell–Żukowski inequalities for various limits of fine structure.

Secondly, our framework informs novel experimental searches for conceivable physics beyond quantum theory. Previous proposals (e.g. superstrong nonlocality [20] or higher-order interference [47, 48]) have simply described the probabilistic effects without predicting how they could actually occur within spacetime as we know it. This has made the search for such effects seem like the search for a needle in a haystack [49]. Our formalism promises a more direct spatiotemporal description of such effects – hopefully leading to predictions that are more tied to experiments and in greater compatibility with spacetime physics.

Combining the principles of quantum theory with special relativity has historically been an extremely fruitful strategy. Here, we propose to extend this strategy to device-independent quantum information and even beyond quantum physics. In principle, suitable extensions of our framework would allow us to address questions such as: *which probability rules are compatible with Lorentz invariance*? Any progress on these kind of questions has the potential to give us fascinating insights into the logical architecture of our physical world.

ACKNOWLEDGMENTS

We are grateful to Miguel Navascués, Matt Pusey, and Valerio Scarani for discussions. This project was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. We acknowledge the support of the Austrian Science Fund (FWF) through the Doctoral Programme CoQuS. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

[1] A. Einstein, B. Podolsky, and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" Physical Review **47**, 777–780 (1935).

[2] J. S. Bell, "On the Einstein Podolsky Rosen paradox," Physics Physique Fizika **1**, 195–200 (1964).

[3] J. Clauser, M. Horne, A. Shimony, and R. Holt, "Proposed Experiment to Test Local Hidden-Variable Theories," Physical Review Letters **23**, 880–884 (1969).

[4] A. Aspect, P. Grangier, and G. Roger, "Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's In-

equalities," Physical Review Letters **49**, 91–94 (1982).

[5] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, "Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres," Nature **526**, 682–686 (2015).

[6] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, "Bell nonlocality," Reviews of Modern Physics **86**, 419–478 (2014).

[7] D. Mayers and A. Yao, "Quantum Cryptography with Imperfect Apparatus," in *Proceedings of the 39th Annual Symposium on Foundations of Computer Science* (IEEE, Palo Alto, 1998).

[8] J. Barrett, L. Hardy, and A. Kent, "No Signaling and Quantum Key Distribution," Physical Review Letters **95**, 010503 (2005).

[9] R. Colbeck and R. Renner, "Free randomness can be amplified," Nature Physics **8**, 450–453 (2012).

[10] U. Vazirani and T. Vidick, "Fully Device-Independent Quantum Key Distribution," Physical Review Letters **113**, 140501 (2014).

[11] J. Tura, R. Augusiak, A. B. Sainz, T. Vértesi, M. Lewenstein, and A. Acín, "Quantum nonlocality. Detecting nonlocality in many-body quantum states." Science (New York, N.Y.) **344**, 1256–8 (2014).

[12] R. Schmied, J.-D. Bancal, B. Allard, M. Fadel, V. Scarani, P. Treutlein, and N. Sangouard, "Bell correlations in a Bose-Einstein condensate." Science (New York, N.Y.) **352**, 441–4 (2016).

[13] J. Pearl, *Causality: Models, reasoning, and inference* (Cambridge University Press, 2000).

[14] M. Nakahara, *Geometry, Topology and Physics* (IOP Publishing, 2003).

[15] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. A. Méthot, and V. Scarani, "Testing the Dimension of Hilbert Spaces," Physical Review Letters **100**, 210503 (2008).

[16] M. Pawłowski and N. Brunner, "Semi-device-independent security of one-way quantum key distribution," Physical Review A **84**, 010302 (2011).

[17] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, Vol. 3 (1984) pp. 175–179.

[18] A. Acín, N. Gisin, and Ll. Masanes, "From Bell's Theorem to Secure Quantum Key Distribution," Physical Review Letters **97**, 120405 (2006).

[19] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[20] S. Popescu and D. Rohrlich, "Quantum nonlocality as an axiom," Found. Phys. **24**, 379–385 (1994).

[21] R. F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model," Physical Review A **40**, 4277–4281 (1989).

[22] R. Augusiak, M. Demianowicz, and A. Acín, "Local hidden–variable models for entangled quantum states," Journal of Physics A: Mathematical and Theoretical **47**, 424002 (2014).

[23] M. Żukowski, "Bell theorem involving all settings of measuring apparatus," Physics Letters A **177**, 290–296 (1993).

[24] A. Sen(De), U. Sen, and M. Żukowski, "Functional Bell inequalities can serve as a stronger entanglement witness than conventional Bell inequalities," Physical Review A - Atomic, Molecular, and Optical Physics **66**, 4 (2002).

[25] B. S. Cirel'son, "Quantum generalizations of Bell's inequality," Letters in Mathematical Physics **4**, 93–100 (1980).

[26] F. Hirsch, M. T. Quintino, T. Vértesi, M. Navascués, and N. Brunner, "Better local hidden variable models for two-bit Werner states and an upper bound on the Grothendieck constant $K_G(3)$," Quantum **1**, 3 (2017).

[27] K. Nagata, W. Laskowski, M. Wieśniak, and M. Żukowski, "Rotational invariance as an additional constraint on local realism," Physical Review Letters **93** (2004), 10.1103/PhysRevLett.93.230403.

[28] S. L. Braunstein and C. M. Caves, "Wringing out better Bell inequalities," Annals of Physics **202**, 22–56 (1990).

[29] Z. Wang, S. Singh, and M. Navascués, "Entanglement and Nonlocality in Infinite 1D Systems," Physical Review Letters **118**, 230401 (2017).

[30] L. A. Khalfin and B. S. Tsirelson, "Quantum and Quasi-classical Analogs Of Bell Inequalities," in *Symposium on the foundations of modern physics*, edited by P. Lahti and P. Mittelstaedt (World Scientific Publishing Co., 1985) pp. 441–460.

[31] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski, "Information causality as a physical principle," Nature **461**, 1101–1104 (2009).

[32] T. H. Yang, M. Navascués, L. Sheridan, and V. Scarani, "Quantum Bell inequalities from macroscopic locality," Physical Review A **83**, 022105 (2011).

[33] G. Brassard, H. Buhrman, N. Linden, A. Méthot, A. Tapp, and F. Unger, "Limit on Nonlocality in Any World in Which Communication Complexity Is Not Trivial," Physical Review Letters **96**, 250401 (2006).

[34] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín, "Almost quantum correlations," Nature Communications **6**, 6288 (2015).

[35] K. F. Pál and T. Vértesi, "Concavity of the set of quantum probabilities for any given dimension," Physical Review A **80**, 042114 (2009).

[36] J. M. Donohue and E. Wolfe, "Identifying nonconvexity in the sets of limited-dimension quantum correlations," Physical Review A **92**, 062120 (2015).

[37] A. Acín, R. Augusiak, D. Cavalcanti, C. Hadley, J. K. Korbicz, M. Lewenstein, Ll. Masanes, and M. Piani, "Unified Framework for Correlations in Terms of Local Quantum Observables," Physical Review Letters **104**, 140404 (2010).

[38] W. K. Wootters, *The acquisition of information from quantum measurements*, Ph.D. thesis, University of Texas at Austin (1980).

[39] M. P. Müller and Ll. Masanes, "Three-dimensionality of space and the quantum bit: an information-theoretic approach," New Journal of Physics **15**, 053040 (2013).

[40] P. A. Höhn and M. P. Müller, "An operational approach to spacetime symmetries: Lorentz transformations from quantum communication," New Journal of Physics **18**, 063026 (2016).

[41] A. J. P. Garner, M. P. Müller, and O. C. O. Dahlsten, "The complex and quaternionic quantum bit from relativity of simultaneity on an interferometer," Proceedings of the Royal Society A **473**, 20170596 (2017).

[42] M. R. Sepanski, *Complex Lie groups* (Springer, New York, 2007).

[43] D. Branford, O. C. O. Dahlsten, and A. J. P. Garner, "On Defining the Hamiltonian Beyond Quantum

Theory," Foundations of Physics **48**, 982–1006 (2018).
[44] T. van Himbeeck, E. Woodhead, N. J. Cerf, R. García-Patrón, and S. Pironio, "Semi-device-independent framework based on natural physical assumptions," Quantum **1**, 33 (2017).
[45] T. Van Himbeeck and S. Pironio, "Correlations and randomness generation based on energy constraints," arXiv:1905.09117 (2019).
[46] D. Rusca, T van Himbeeck, A. Martin, J. B. Brask, W. Shi, S. Pironio, N. Brunner, and H. Zbinden, "Practical self-testing quantum random number generator based on an energy bound," arXiv:1904.04819 (2019).
[47] R. D. Sorkin, "Quantum Mechanics as Quantum Measure Theory," Modern Physics Letters A **09**, 3119–3127 (1994).
[48] C. Ududec, H. Barnum, and J. Emerson, "Three Slit Experiments and the Structure of Quantum Theory," Foundations of Physics **41**, 396–405 (2010).
[49] U. Sinha, C. Couteau, T. Jennewein, R. Laflamme, and G. Weihs, "Ruling out multi-order interference in quantum mechanics." Science (New York, N.Y.) **329**, 418–21 (2010).
[50] "Finite-dimensional representation." Encyclopedia of Mathematics. (2016).
[51] V. F. Molchanov, A. U. Klimyk, and N. Ya. Vilenkin, *Representation Theory and Noncommutative Harmonic Analysis II*, edited by A. A. Kirillov (Springer Verlag, 1995).
[52] M. Kleinmann, T. J. Osborne, V. B. Scholz, and A. H. Werner, "Typical Local Measurements in Generalized Probabilistic Theories: Emergence of Quantum Bipartite Correlations," Physical Review Letters **110**, 040403 (2013).
[53] H. Barnum, S. Beigi, S. Boixo, M. B. Elliott, and S. Wehner, "Local Quantum Measurement and No-Signaling Imply Quantum Correlations," Phys. Rev. Lett. **104**, 140401 (2010).
[54] Ll. Masanes, "Extremal quantum correlations for N parties with two dichotomic observables per site," arXiv:quant-ph/0512100 (2005).
[55] B. Toner and F. Verstraete, "Monogamy of Bell correlations and Tsirelson's bound," arXiv:quant-ph/0611001 (2006).

Appendix A: The representation of spatiotemporal degrees of freedom in black box statistics

Let us first furnish a mathematical description of a black box as an input–output process. We begin with the single party case (say, Alice). Suppose the domain of Alice's inputs is the set $\mathcal{X}$, and of her outputs is the finite set $\mathcal{A}$. As motivated in the main text, we are interested in the case where $\mathcal{X}$ is a homogeneous space. That is, we have a group $\mathcal{G}$ that acts transitively on the set of inputs $\mathcal{X}$, such that $\mathcal{X} = \mathcal{G}/\mathcal{H}$, and $\mathcal{H} \subseteq \mathcal{G}$ is the corresponding stabilizer subgroup. The paradigmatic example is given by $\mathcal{X} = S^{d-1}$, $\mathcal{G} = \mathrm{SO}(d)$ and $\mathcal{H} = \mathrm{SO}(d-1) \subset \mathcal{G}$, such that the inputs $\vec{x} \in \mathcal{X}$ are unit vectors. Even though the inputs need not be vectors in general, we will use the vector notation in the following for convenience. We will assume that $\mathcal{G}$ is a locally compact group, such that all bounded finite-dimensional representations are unitary [50].

For such an input domain, we can assign an arbitrary "default input" $\vec{x}_0 \in \mathcal{X}$, such that every other input $\vec{x} \in \mathcal{X}$ can be written as $\vec{x} = R_{\vec{x}}\vec{x}_0$ for some suitable transformation $R_{\vec{x}} \in \mathcal{G}$. Physically, we can imagine that Alice chooses her input by "rotating" the default input $\vec{x}_0$ into her desired direction $\vec{x}$, and she can do so by applying a suitable rotation $R_{\vec{x}}$. In general, $R_{\vec{x}}$ is not unique, and Alice's freedom of choice of transformation is given by $\mathcal{H}$.

A **black box** P is then a map $\mathrm{P} : \mathcal{X} \to \mathbb{R}^{|\mathcal{A}|}$ such that for $\vec{x} \in \mathcal{X}$, $\mathrm{P}^a : \vec{x} \mapsto \mathrm{P}(A = a \,|\, X = \vec{x})$, where P^a is the a^{th} element of the vector map. Since for probabilities $0 \leq \mathrm{P}(A = a \,|\, X = \vec{x}) \leq 1$, each P^a is a non-negative real bounded function on $\mathcal{X}$. For probabilities, we also have the constraint that for all $\vec{x}$, $\sum_a \mathrm{P}^a\vec{x} = 1$; so the range of the vector function P is actually that of $(|\mathcal{A}| - 1)$–dimensional simplices (a compact convex subspace of $\mathbb{R}^{|\mathcal{A}|}$). As such, $\mathrm{P} \in \mathcal{B}(\mathcal{X})^{|\mathcal{A}|}$ where $\mathcal{B}(\mathcal{X})$ is the set of bounded functions on $\mathcal{X}$.

Definition 1 ($\mathcal{G}$-box)**.** *A black box (formalized above) whose input domain $\mathcal{X}$ is a homogeneous space acted transitively upon by the group $\mathcal{G}$ is known as a $\mathcal{G}$-box.*

Proof of Theorem 1. *Consider a $\mathcal{G}$-box whose ensemble behaviour can be characterized by a finite number of parameters (Assumption (i)). There is a representation of the symmetry group $\mathcal{G}$ in terms of real orthogonal matrices $R \mapsto T_R$, such that for each outcome a, the outcome probabilities $\mathrm{P}(a \,|\, x_R)$ are a fixed (over R) linear combination of matrix entries of T_R.*

Proof. Suppose Alice has a black box P, and access to a geometric freedom $\mathcal{G}$ acting on $\mathcal{X}$. For each $R \in \mathcal{G}$, Alice can induce a new black box P′ by first applying R to her input $\vec{x}$ and then supplying the input $R\vec{x}$ to P, which acts as $\mathrm{P}'^a : \vec{x} \mapsto \mathrm{P}(a \,|\, R\vec{x})$, i.e. $\mathrm{P}'(a|\vec{x}) = \mathrm{P}(a \,|R\vec{x})$.

For each R, we can define a map $T_R : \mathrm{P} \mapsto \mathrm{P}'$, acting on each component of P via $T_R\mathrm{P}^a = \mathrm{P}'^a$. Obviously, $T_R \circ T_S = T_{RS}$, so if we denote the "space of black boxes" accessible to Alice by $\Omega_{\mathcal{G}} := \{T_R\mathrm{P} \,|\, R \in \mathcal{G}\} \subseteq \mathcal{B}(\mathcal{X})^{|\mathcal{A}|}$, then T_R defines a group action on $\Omega_{\mathcal{G}}$.

Consider the linear extension $\Omega_{\mathcal{G}}^{\mathbb{R}} := \mathrm{span}\,(\Omega_{\mathcal{G}})$, a linear subspace of $\mathcal{B}(\mathcal{X})^{|\mathcal{A}|}$, with elements $Q = \sum_{i=1}^{n} \lambda_i \mathrm{P}_i$, where $n \in \mathbb{N}$ is arbitrary but finite, all $\lambda_i \in \mathbb{R}$, and $\mathrm{P}_i \in \Omega_{\mathcal{G}}$. Note $Q : \mathcal{X} \to \mathbb{R}^{|\mathcal{A}|}$, but without further restriction on $\{\lambda_i\}$ this may map to outside of the simplex of normalized probabilities.

Now, consider the effect of $R \in \mathcal{G}$ on some object Q. Since $Q : \vec{x} \mapsto \sum_i \lambda_i \mathrm{P}(a \,|\, \vec{x})$, applying R first to take $R : \vec{x} \mapsto \vec{x}'$ gives us $Q \circ R : \vec{x} \mapsto \sum_i \lambda_i \mathrm{P}(a \,|\, R\vec{x})$, and hence $Q \circ R = \sum_i \lambda_i T_R P_i$. Since $T_R P = P \circ R$ for $P \in \Omega_{\mathcal{G}}$, we can define the map $\tilde{T}_R : \Omega_{\mathcal{G}}^{\mathbb{R}} \to \Omega_{\mathcal{G}}^{\mathbb{R}}$ via $\tilde{T}_R Q := Q \circ R$ as an extension of the map T_R. By construction, every $\tilde{T}_R$ is a linear map, and

$$\tilde{T}_R \tilde{T}_S(Q) = Q \circ R \circ S = Q \circ (R \circ S) = \tilde{T}_{RS}(Q), \quad \text{(A1)}$$

hence $R \mapsto \tilde{T}_R$ is a real linear representation of $\mathcal{G}$. Since $\tilde{T}_R$ is an extension of T_R, we drop the tilde from our notation. As we have assumed that ensembles of black boxes can be characterized by a finite number of parameters, the linear space $\Omega_{\mathcal{G}}^{\mathbb{R}}$ is finite-dimensional. Then T_R, as linear maps acting on a finite-dimensional real vector space, may be expressed as real matrices.

Next, we need to show that the representation $R \mapsto T_R$ is *bounded*, i.e. that $\sup_{R \in \mathcal{G}} \|T_R\| < \infty$. This will exclude, for example, cases like $\mathcal{G} = (\mathbb{R}, +)$ and $T_t := \left(\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right)$. To this end, let $P_1, \ldots, P_D \in \Omega_{\mathcal{G}}$ be a linearly independent set of boxes that spans $\Omega_{\mathcal{G}}^{\mathbb{R}}$ (that is, a basis of boxes, hence $D = \dim \Omega_{\mathcal{G}}^{\mathbb{R}}$). Then, every $\mathrm{P} \in \Omega_{\mathcal{G}}^{\mathbb{R}}$ has a unique representation $\mathrm{P} = \sum_{i=1}^{D} \alpha_i P_i$, and $\|\mathrm{P}\|_1 := \sum_{i=1}^{D} |\alpha_i|$ defines a norm on $\Omega_{\mathcal{G}}^{\mathbb{R}}$. We can define another norm on this space via

$$\|P\| := \sup_{\vec{x} \in \mathcal{X}} \sum_{a \in \mathcal{A}} |\mathrm{P}(a \,|\, \vec{x})| \,. \quad \text{(A2)}$$

This is finite since $\mathrm{P} \in \mathcal{B}(\mathcal{X})^{|\mathcal{A}|}$, and it is easy to check that it satisfies the properties of a norm. Since all norms on a finite-dimensional vector space are equivalent, there is some $c > 0$ such that $\| \bullet \|_1 \leq c \| \bullet \|$. Furthermore, all $\mathrm{P}' \in \Omega_{\mathcal{G}}$ satisfy $\|\mathrm{P}'\| = 1$. Thus, noting that $T_R P_i \in \Omega_{\mathcal{G}}$ for all $i = 1, \ldots, D$, we get

$$\begin{aligned} \|T_R\mathrm{P}\| &= \left\| \sum_{i=1}^{D} \alpha_i T_R P_i \right\| \leq \sum_{i=1}^{D} |\alpha_i| \cdot \|T_R P_i\| \\ &= \|\mathrm{P}\|_1 \leq c \cdot \|\mathrm{P}\|. \end{aligned} \quad \text{(A3)}$$

This establishes that the operator norm of all T_R with respect to $\| \bullet \|$ (and hence with respect to all other norms) is uniformly bounded. Since we have assumed that $\mathcal{G}$ is locally compact, this implies that there is a basis of $\Omega_{\mathcal{G}}^{\mathbb{R}}$ in which the T_R are orthogonal matrices.

Consider now the *evaluation functional* $\delta_{\vec{x}}^a : \Omega_{\mathcal{G}}^{\mathbb{R}} \to \mathbb{R}$; namely, the map from the space of black boxes to the particular probability of outcome a given input $\vec{x}$. It follows that the statistics $\mathrm{P}(a \,|\, \vec{x}) = \mathrm{P}(a \,|\, R\vec{x}_0) = T_R P^a(\vec{x}_0) = \delta_{\vec{x}_0}^a (T_R P)$. Since the evaluation functional is a linear map, we then find that the probabilities are given by a linear combination of elements from T_R. For all $\vec{x} \in \mathcal{X}$, we use the same P and the same $\delta_{\vec{x}_0}^a$ such

that the only element that changes is the representation matrix T_R. □

Arguing via harmonic analysis on homogeneous spaces [51], we expect that Theorem 1 can be extended: it is not only entries of T_R that appear in the probability table $P(a|x_R)$, but, more specifically, generalized spherical harmonics. A taste of this appears in Lemma VIII, but since the formulation of Theorem 1 is sufficient for the purpose of this article, we defer this extension to future work.

Appendix B: $\mathrm{SO}(2) \times \mathrm{SO}(2)$ Bell experiment setting

1. General form of correlations

Lemma I. *Consider a bipartite* $\mathrm{SO}(2) \times \mathrm{SO}(2)$*-box – i.e. Alice and Bob can each choose their inputs as angles* $\alpha, \beta \in [0, 2\pi)$ *– with local binary outcomes* $a \in \{+1, -1\}$ *and* $b \in \{+1, -1\}$*. Then, the most general joint probability distribution consistent with Theorem 1 is*

$$\mathrm{P}(a, b \,|\, \alpha, \beta) := \sum_{m=0}^{2J} \sum_{n=-2J}^{2J} c_{mn}^{ab} \cos(m\alpha - n\beta) + s_{mn}^{ab} \sin(m\alpha - n\beta), \tag{B1}$$

where J is some non-negative integer or half-integer. (Note that this does not yet assume the no-signalling principle.)

Proof. While the representation $T_R = T_{\alpha,\beta}$ from Theorem 1 acts on a real vector space V of finite dimension D, we can also regard it as a representation on the complexification $W = V \oplus iV$. Since $\mathrm{SO}(2) \times \mathrm{SO}(2)$ is an Abelian group, all its irreducible representations are one-dimensional [42]. Thus, we can decompose W as $W = \bigoplus_{j=1}^{D} W_j$, where each W_j is a one-dimensional invariant subspace on which $T_{\alpha,\beta}$ acts as a complex phase. It follows that $T_{\alpha,\beta} = \bigoplus_{j=1}^{D} \exp(i(m_j\alpha - n_j\beta))$ with suitable integers $m_j, n_j \in \mathbb{Z}$ (to see this, write $T_{\alpha,\beta}$ as a composition of the $\mathrm{SO}(2)$-representations $T_{\alpha,0}$ and $T_{0,\beta}$). Then, due to Theorem 1, $\mathrm{P}(a, b \,|\, \alpha, \beta)$ must be a linear combination of real and imaginary parts of $T_{\alpha,\beta}$, which proves that it is of the form (B1). □

2. Generic no-signalling correlations

It is well-known (e.g. [20]) that the no-signalling principle does not impose any constraints on the form of the correlation function if we have a bipartite box with two outcomes $a, b \in \{+1, -1\}$ each. Namely, if $\mathcal{X}, \mathcal{Y}$ denote two arbitrary sets of inputs, given an arbitrary function $C : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ with $-1 \le C(x, y) \le 1$ for all $x, y \in \mathcal{X}, \mathcal{Y}$, the simple prescription

$$\mathrm{P}(a, b \,|\, x, y) := \frac{1}{4} + \frac{1}{4} ab\, C(x, y) \tag{B2}$$

generates a valid no-signalling distribution that has $C(x, y)$ as its correlation function. It is a simple exercise to check that C is non-negative, normalized and no-signalling, and that C is the correlation function for P.

3. Local hidden variable models for $\mathrm{SO}(2) \times \mathrm{SO}(2)$ settings

Generalizing ideas of Werner [21], we can show that for noisy enough correlation functions of $\mathrm{SO}(2) \times \mathrm{SO}(2)$ settings, we can always construct a LHV model that achieves these correlations.

Lemma II. *Consider any two-angle function*

$$f(\alpha, \beta) = \sum_{j=1}^{N} \left[c_j \cos(m_j\alpha - n_j\beta) + s_j \sin(m_j\alpha - n_j\beta) \right], \tag{B3}$$

for $(m_j, n_j) \in \mathbb{Z} \times \mathbb{Z} \setminus (0, 0)$ (i.e. disallowing constant terms). Without loss of generality[1] *in $f(\alpha, \beta)$, we disallow $(m_i, n_i) = (m_j, n_j)$ when $i \neq j$, choose $m_j \ge 0$, and if $m_j = 0$ then we choose $n_j > 0$. Suppose $-1 \le f(\alpha, \beta) \le 1$ for all α, β. Then $C(\alpha, \beta) := \gamma\, f(\alpha, \beta)$ is a correlation function that has a LHV model whenever $0 \le \gamma \le \gamma_N$, where*

$$\gamma_N := \sqrt{\frac{2}{N}} \max_{0 \le x \le \pi} \left(\frac{x}{\pi}\right)^{N-1} \frac{\sin x}{\pi}. \tag{B4}$$

Proof. In this proof, we will express all angles as numbers in the interval $[-\pi, \pi)$. Under the inner product $\langle f, g \rangle := \frac{1}{2\pi^2} \int_{-\pi}^{\pi} \mathrm{d}\alpha \int_{-\pi}^{\pi} \mathrm{d}\beta f(\alpha, \beta) g(\alpha, \beta)$, the set of functions

$$\cos(m_j\alpha - n_j\beta), \quad \sin(m_j\alpha - n_j\beta) \tag{B5}$$

(with m_j and n_j defined as above) is an orthonormal system (this follows from Schur orthonormality for $\mathrm{SO}(2) \times \mathrm{SO}(2)$, and can be verified by direct integration). Hence the L^2-norm $\|f\|^2 := \langle f, f \rangle$ satisfies

$$\|f\|^2 = \int_{-\pi}^{\pi} \mathrm{d}\alpha \int_{-\pi}^{\pi} \mathrm{d}\beta \frac{f(\alpha, \beta)^2}{2\pi^2} = \sum_{j=1}^{N} (c_j^2 + s_j^2) \le 2, \tag{B6}$$

since $|f(\alpha, \beta)| \le 1$ everywhere.

Our goal is to construct a LHV model of the form

$$\mathrm{P}(a, b \,|\, \alpha, \beta) = \int_{\Lambda} \mathrm{d}\mu(\lambda)\, \mathrm{P}_A(a \,|\, \alpha, \lambda)\, \mathrm{P}_B(b \,|\, \beta, \lambda). \tag{B7}$$

We will have a hidden variable $\lambda = (\vec{\lambda}_1, \vec{\lambda}_2, \ldots, \vec{\lambda}_N)$, where each $\vec{\lambda}_j = (\cos\phi_j, \sin\phi_j)^{\mathrm{T}}$ is independently and uniformly distributed on the unit circle, hence $\mathrm{d}\mu(\lambda) = (2\pi)^{-N} \mathrm{d}\phi_1 \ldots \mathrm{d}\phi_N$. This measure is invariant under $\mathrm{SO}(2)$ rotations of the individual $\vec{\lambda}_j$.

[1] These restrictions ensure that the coefficients c_j and s_j are associated with unique trigonometric functions.

We will construct local probabilities that implement the dependence on α, β, λ in the following form:

$$\mathrm{P}_A(\pm\,|\,\alpha,\lambda) = q_A\left(\pm\middle|R_{m_1\alpha}\vec{\lambda}_1,\ldots,R_{m_N\alpha}\vec{\lambda}_N\right), \quad \text{(B8)}$$

$$\mathrm{P}_B(\pm\,|\,\beta,\lambda) = q_B\left(\pm\middle|R_{n_1\beta}\vec{\lambda}_1,\ldots,R_{n_N\beta}\vec{\lambda}_N\right), \quad \text{(B9)}$$

where q_A and q_B are response functions defined in the following way: $q_A(-|\vec{\lambda}'_1,\ldots,\vec{\lambda}'_N) = 1 - q_A(+|\vec{\lambda}'_1,\ldots,\vec{\lambda}'_N)$,

$$q_A(+|\vec{\lambda}'_1,\ldots,\vec{\lambda}'_N) := \begin{cases} 1 & \text{if } \phi'_j \in [-\xi,\xi] \text{ for all } j \\ 0 & \text{otherwise} \end{cases} \quad \text{(B10)}$$

where $\xi \in (0,\pi)$ is some (small) constant and ϕ'_j is the angle such that $\vec{\lambda}'_j = (\cos\phi'_j, \sin\phi'_j)^{\mathrm{T}}$. Furthermore,

$$q_B\left(\pm|\vec{\lambda}'_1,\ldots,\vec{\lambda}'_N\right) := \frac{1}{2}\left(1 \pm \frac{1}{\sqrt{2N}}\sum_{j=1}^{N}\vec{b}_j\cdot\vec{\lambda}'_j\right), \quad \text{(B11)}$$

where $\vec{b}_j := (c_j, -s_j)^{\mathrm{T}}$. Note that $\sum_j \vec{b}_j\cdot\vec{\lambda}'_j = \vec{b}\cdot\vec{\lambda}'$, where $\vec{b} := \oplus_j\vec{b}_j$ and $\vec{\lambda}' := \oplus_j\vec{\lambda}'_j$. But since $|\vec{b}|^2 = \sum_j|\vec{b}_j|^2 = \sum_j(c_j^2+s_j^2) \le 2$ and $|\lambda'|^2 = \sum_j|\vec{\lambda}'_j|^2 = N$, the sum hence is upper-bounded by $\sqrt{2N}$ due to the Cauchy-Schwarz inequality. This shows that q_B yields valid probabilities.

We calculate the joint probability distribution obtained in the Bell test scenario:

$$\begin{aligned}\mathrm{P}(+,\pm\,|\,\alpha,\beta) &= \int_\Lambda \mathrm{d}\mu(\lambda)\ \mathrm{P}_A(+\,|\,\alpha,\lambda)\,\mathrm{P}_B(\pm\,|\,\beta,\lambda) \\ &= \int_\Lambda \mathrm{d}\mu(\lambda)\ q_A(+|R_{m_1\alpha}\vec{\lambda}_1,\ldots,R_{m_N\alpha}\vec{\lambda}_N) \\ &\qquad \cdot q_B(\pm|R_{n_1\beta}\vec{\lambda}_1,\ldots,R_{n_N\beta}\vec{\lambda}_N). \quad \text{(B12)}\end{aligned}$$

We apply the substitution $\vec{\lambda}'_j := R_{m_j\alpha}\vec{\lambda}_j$ and $\lambda' := (\vec{\lambda}'_1,\ldots,\vec{\lambda}'_N)$, noting that this does not change the integral due to our choice of measure:

$$\begin{aligned}\mathrm{P}(+,\pm\,|\,\alpha,\beta) = \int_\Lambda \mathrm{d}\mu(\lambda')\ q_A(+|\vec{\lambda}'_1,\ldots,\vec{\lambda}'_N) \\ \cdot q_B(\pm|R_{n_1\beta-m_1\alpha}\vec{\lambda}'_1,\ldots,R_{n_N\beta-m_N\alpha}\vec{\lambda}'_N). \quad \text{(B13)}\end{aligned}$$

Due to the definition of q_A and q_B, this equals

$$\int_{-\xi}^{\xi}\frac{\mathrm{d}\phi'_1}{2\pi}\cdots\int_{-\xi}^{\xi}\frac{\mathrm{d}\phi'_N}{2\pi}\frac{1}{2}\left(1 \pm \frac{1}{\sqrt{2N}}\sum_j f_j(\phi'_j)\right), \quad \text{(B14)}$$

where

$$f_j(\phi'_j) = c_j\cos(\phi'_j+n_j\beta-m_j\alpha) - s_j\sin(\phi'_j+n_j\beta-m_j\alpha). \quad \text{(B15)}$$

Noting that

$$\begin{aligned}\int_{-\xi}^{\xi} f_j(\phi'_j)\mathrm{d}\phi'_j = 2\sin\xi\,\big[c_j\cos(m_j\alpha-n_j\beta) \\ +s_j\sin(m_j\alpha-n_j\beta)\big], \quad \text{(B16)}\end{aligned}$$

we can evaluate the integral explicitly, obtaining

$$\begin{aligned}&\mathrm{P}(+\pm\,|\,\alpha,\beta) = \\ &\quad \frac{1}{2}\left(\frac{\xi}{\pi}\right)^N \pm \frac{1}{2\sqrt{2N}}\left(\frac{\xi}{\pi}\right)^{N-1}\frac{\sin\xi}{\pi}\,f(\alpha,\beta). \quad \text{(B17)}\end{aligned}$$

Next, let us look at the other probabilities:

$$\begin{aligned}\mathrm{P}(-,\pm\,|\,\alpha,\beta) &= \int_\Lambda \mathrm{d}\mu(\lambda)\mathrm{P}_A(-\,|\,\alpha,\lambda)\,\mathrm{P}_B(\pm\,|\,\beta,\lambda) \\ &= \int_\Lambda \mathrm{d}\mu(\lambda)(1-\mathrm{P}_A(+\,|\,\alpha,\lambda))\mathrm{P}_B(\pm\,|\,\beta,\lambda) \\ &= -\mathrm{P}(+,\pm|\alpha,\beta) + \int_\Lambda \mathrm{d}\mu(\lambda)q_B(\pm|R_{n_1\beta}\vec{\lambda}_1,\ldots,R_{n_N\beta}\vec{\lambda}_N), \quad \text{(B18)}\end{aligned}$$

and the final integral vanishes on all $\vec{b}_j\cdot\vec{\lambda}'_j$-terms of q_B, leaving only the constant term $1/2$. That is,

$$\mathrm{P}(-,\pm\,|\,\alpha,\beta) = \frac{1}{2} - \mathrm{P}(+,\pm|\alpha,\beta). \quad \text{(B19)}$$

These give the correlation function

$$C(\alpha,\beta) = \sqrt{\frac{2}{N}}\left(\frac{\xi}{\pi}\right)^{N-1}\frac{\sin\xi}{\pi}f(\alpha,\beta). \quad \text{(B20)}$$

Finally, we define γ_N as the largest admissible prefactor among all possible choices of ξ. □

Let us now introduce a constant term:

Lemma III. *Consider any two-angle correlation function*

$$C(\alpha,\beta) = c_0 + \sum_{j=1}^{N}\left[c_j\cos(m_j\alpha-n_j\beta) + s_j\sin(m_j\alpha-n_j\beta)\right], \quad \text{(B21)}$$

where $(m_j,n_j) \in \mathbb{Z}\times\mathbb{Z}\setminus(0,0)$, and (as above) without loss of generality we choose $m_j \ge 0$ and $n_j > 0$ if $m_j = 0$, and disallow $(m_i,n_i) = (m_j,n_j)$ if $i \ne j$. If

$$\max_{\alpha,\beta}|C(\alpha,\beta) - c_0| \le \gamma_N(1-|c_0|) \quad \text{(B22)}$$

with constant γ_N given by

$$\gamma_N = \begin{cases} \sqrt{2}/\pi & \text{if } N=1, \\ 0.184375\ldots & \text{if } N=2, \\ 0.103893\ldots & \text{if } N=3, \\ \sqrt{2}e^{-1}\,N^{-3/2} & \text{if } N\ge 4, \end{cases} \quad \text{(B23)}$$

then this correlation function has a local hidden variable model.

Proof. First, we obtain the form of γ_N by solving the optimization problem of Lemma II exactly for $N = 1,2,3$, and by substituting $x = \pi\left(1-\frac{1}{N}\right)$ and using

$$\left(1-\frac{1}{N}\right)^{N-1}\sin\left[\pi\left(1-\frac{1}{N}\right)\right] \ge \frac{\pi}{Ne} \text{ for } N \ge 4. \quad \text{(B24)}$$

We can add the constant function $1/\sqrt{2}$ to the orthonormal system in (eq. (B5)); similar reasoning as in

the proof of Lemma II shows that $(\sqrt{2}c_0)^2 \leq 2$, i.e. that $-1 \leq c_0 \leq 1$, and $|c_0| = 1$ is only possible if $C(\alpha, \beta) = c_0$ (i.e. with no angle-dependent terms). Now consider the case $0 \leq c_0 < 1$. We can write

$$C(\alpha, \beta) = c_0 \mathbf{1} + (1 - c_0) f(\alpha, \beta), \tag{B25}$$

where $\mathbf{1}$ is the constant function that takes the value 1 on all angles, and $f(\alpha, \beta) = (C(\alpha, \beta) - c_0)/(1 - c_0)$ is of the form of the function in Lemma II. If inequality (B22) holds, then

$$\max_{\alpha,\beta} |f(\alpha, \beta)| = \frac{1}{1 - c_0} \max_{\alpha,\beta} |C(\alpha, \beta) - c_0| \leq \gamma_N, \tag{B26}$$

and so Lemma II proves that $f(\alpha, \beta)$ is a classical correlation function. Moreover, $\mathbf{1}$ is trivially a classical correlation function, and thus so must be $C(\alpha, \beta)$, which is a convex combination of the two. Then case $-1 < c_0 < 0$ can be treated analogously, using that $-\mathbf{1}$ is a classical correlation function too. □

Proof of Theorem 2A. Consider an SO(2) × SO(2) box, with a correlation function in the form of eq. (4) with maximum (half-)integer $J \neq 0$. If

$$\max_{\alpha,\beta} |C(\alpha, \beta) - C_{00}| \leq \gamma_J (1 - |C_{00}|), \tag{B27}$$

where C_{00} is the angle-independent contribution to the correlation function (as in eq. (4)), and γ_J is a given

$$\gamma_J = \sqrt{2} e^{-1} \left[4J \left(2J + 1\right)\right]^{-\frac{3}{2}}, \tag{B28}$$

then there is a LHV model that accounts for these correlations.

Proof. This follows as a corollary of Lemma III. We convert between the form of correlations in eq. (4) and eq. (B21) by counting the maximum number N of unique terms that could appear for a given positive (half-)integer J. The double sum contributes $(2J + 1)(4J + 1)$ terms, from which we remove $2J$ cases corresponding to negative n where $m = 0$, and the one completely constant case $m = n = 0$. This gives a maximum of $N = 4J(2J + 1)$. Since the lowest value ($J = \frac{1}{2}$) already yields $N = 4$ unique terms, we only need the final case of eq. (B23), and hence the constant $\gamma_J = \sqrt{2} e^{-1} \left[4J \left(2J + 1\right)\right]^{-\frac{3}{2}}$. □

4. Witnessing nonlocality

Bell inequalities can be *chained* by direct addition. For instance, suppose one takes a CHSH inequality (eq. (5)) with measurements $\{x_1, y_2, x_3, y_4\}$ and a second with measurements $\{x_1, y_4, x_5, y_6\}$. Adding these together yields $\big|C(x_1, y_2) + C(x_3, y_2) + C(x_3, y_4) + C(x_5, y_4) + C(x_5, y_6) - C(x_1, y_6)\big| \leq 4$. This can inductively be done for a set of N measurements ($\frac{N}{2}$ each for Alice and Bob), leading to a chained Bell inequality, known as the *Braunstein–Caves inequality* (BCI) [28]:

$$\Big|C(x_1, y_2) + C(x_3, y_2) + C(x_3, y_4) + \cdots + C(x_{N-1}, y_N) - C(x_1, y_N)\Big| \leq N - 2. \tag{B29}$$

If such an equation is violated, then no LHV can account for these statistics[2].

Recall, eq. (4) gives the generic SO(2) × SO(2) correlation function. If we restrict ourselves to *relational* correlations, this amounts to setting $S_{mn} = C_{mn} = 0$ when $m \neq n$, such that the correlation function has a single parameter form

$$C(\beta - \alpha) = \sum_{m=0}^{2J} C_m \cos\left[m \left(\beta - \alpha\right)\right] + S_m \sin\left[m \left(\beta - \alpha\right)\right] \tag{B30}$$

where J is some positive (half-)integer, and $C_m := C_{mm}$, $S_m := S_{mm}$.

Lemma IV. *Consider relational* SO(2) × SO(2) *correlations (of the form of eq.* (B30)*) for finite positive (half-)integer J. If there is some $\Theta_+ \in [0, 2\pi)$ such that $C(\Theta_+) = +1$, and $\Theta_- \in [0, 2\pi)$ where $C(\Theta_-) = -1$, then there exists a BCI that demonstrates a Bell violation.*

Proof. We show this by construction. For even N, define

$$\delta_N := \frac{(\Theta_- - \Theta_+) \mod 2\pi}{N - 1}. \tag{B31}$$

We use the notation "$x \mod 2\pi$" to indicate $x - 2\pi n$ where $n \in \mathbb{Z}$ is chosen such that $x - 2\pi n \in [0, 2\pi)$, mapping the angle to the principal range.

We construct a N-measurement BCI, as defined in eq. (B29). Since the correlation function is relational, we write $C(\alpha, \beta)$ as the single parameter function $C(\beta - \alpha)$, and assign the measurement settings:

$$\begin{aligned} a_i &= (i - 1)\delta_N && \text{for odd } i, \\ b_i &= \Theta_+ + (i - 1)\delta_N && \text{for even } i. \end{aligned} \tag{B32}$$

(Illustrated in figure 3.) This amounts to setting the arguments of the correlation functions featured in the BCI to

$$\begin{aligned} b_2 - a_1 &= \ldots = b_{2m} - a_{2m-1} = \Theta_+ + \delta_N, \\ b_2 - a_3 &= \ldots = b_{2m} - a_{2m+1} = \Theta_+ - \delta_N, \\ b_N - a_1 &= \Theta_-, \end{aligned} \tag{B33}$$

where equality is taken modulo 2π.

[2] The BCI can also be directly justified, just as the CHSH inequality. One writes $\sigma_{x_1}\left(\sigma_{y_2} - \sigma_{y_N}\right) + \sigma_{x_3}\left(\sigma_{y_2} + \sigma y_4\right) + \ldots + \sigma_{x_{N-1}}\left(\sigma_{y_{N-2}} + \sigma_{y_N}\right)$ for spins $\{\sigma_i \in \{+1, -1\}\}$, and notes that if $\sigma_{y_2} = \sigma_{y_4} = \ldots = \sigma_{y_N}$, then $\sigma_{y_2} - \sigma_{y_N} = 0$. This bounds the expression to $N - 2$. Convex combinations, such as eq. (B29), cannot exceed this value.

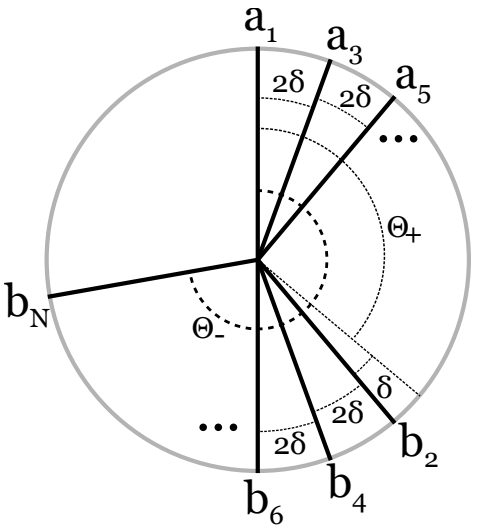


FIG. 3: **Measurement angles for generic correlation function.** The first choice of measurement angles are chosen such that $b_2 - a_1 = \Theta_+ + \delta$. Subsequent choices then precess by 2δ, such that ultimately $b_N - a_1 = \Theta_-$.

With such assignments, the BCI is then written:

$$\frac{N}{2}C(\Theta_+ + \delta_N) + \left(\frac{N}{2} - 1\right) C(\Theta_+ - \delta_N) - C(\Theta_-) \leq N - 2. \tag{B34}$$

Recall that $C(\Theta_-) = -1$. $C(\Theta_+) = +1$ must be a local maximum, and a finite J allows us to assume the function C is smooth at this point. Thus, in the limit of small δ_N, $C(\Theta_+ \pm \delta_N) = 1 - k{\delta_N}^2 + \mathcal{O}({\delta_N}^3)$ where $k \geq 0$ is some constant. We then rewrite eq. (B34) as

$$\begin{aligned}(N-1)\left(1 - k{\delta_N}^2 + \mathcal{O}({\delta_N}^3)\right) + 1 &\leq N - 2, \\ N + \mathcal{O}\left((N-1)^{-1}\right) &\leq N - 2,\end{aligned} \tag{B35}$$

which for large enough N will eventually be violated. □

The above construction can be shown to have some robustness to noise (tolerating a smaller maximum value than 1). To show this, we first prove an auxilary lemma:

Lemma V. *Consider relational* $\mathrm{SO}(2) \times \mathrm{SO}(2)$ *correlations (of the form of eq.* (B30)*) for finite (half-)integer* J*. Expressed as a single parameter function* $C(\beta - \alpha)$*, the second derivative is everywhere bounded:*

$$\left|\frac{\mathrm{d}^2\, C(\beta - \alpha)}{\mathrm{d}(\beta - \alpha)^2}\right| \leq \frac{\sqrt{2}J\,(2J+1)\,(4J+1)}{3}. \tag{B36}$$

Proof. In polar form, the correlation function is:

$$C(\beta - \alpha) = A_0 + \sum_{m=1}^{2J} A_m \cos\left(m\,(\beta - \alpha) - \phi_m\right), \tag{B37}$$

where $A_0 = C_0$, $A_m = \sqrt{C_m^2 + S_m^2}$, and

$$\phi_m = \begin{cases} \arctan(S_m/C_m) & \text{if } C_m \neq 0 \\ \pi/2 & \text{if } C_m = 0, S_m \geq 0 \\ -\pi/2 & \text{if } C_m = 0, S_m < 0. \end{cases} \tag{B38}$$

The second derivative with respect to $\beta - \alpha$ is:

$$\frac{\mathrm{d}^2}{\mathrm{d}(\beta - \alpha)^2} C(\beta - \alpha) = \sum_{m=1}^{2J} -m^2 A_m \cos\left[m\,(\beta - \alpha) - \phi_m\right]. \tag{B39}$$

Because $C(\beta - \alpha)$ is bounded everywhere to $[-1, 1]$, its L^2 norm $\|C\|^2 := \frac{1}{2\pi}\int_0^{2\pi} d\theta\, |C(\theta)|^2$ is bounded within $[0, 1]$. Thus we may determine a maximum value over all m for the amplitude A_m. Using the orthonormality of the functions 1, $\sqrt{2}\cos(m\theta + a)$, and $\sqrt{2}\cos(n\theta + b)$ under the corresponding inner product $\langle f, g\rangle := \frac{1}{2\pi}\int_0^{2\pi} d\theta f(\theta) g(\theta)$, we get $\|C\|^2 = A_0^2 + \frac{1}{2}\sum_{m=1}^{2J} A_m^2$. Since this is upper-bounded by 1, we get $A_m^2 \leq 2$ for all $m \geq 1$, and so

$$\left|\frac{\mathrm{d}^2 C(\beta - \alpha)}{\mathrm{d}(\beta - \alpha)^2}\right| \leq \sum_{m=1}^{2J} \sqrt{2} m^2 = \frac{\sqrt{2}J\,(2J+1)\,(4J+1)}{3}. \tag{B40}$$

□

Lemma VI. *Consider relational* $\mathrm{SO}(2) \times \mathrm{SO}(2)$ *correlations (of the form of eq.* (B30)*) for finite positive (half-)integer* J*. Let there be some angle difference* Θ_+*, where* $C(\Theta_+) \geq 1 - \varepsilon$*, and some other angle difference* Θ_- *where* $C(\Theta_-) \leq 1 - \Delta$*, with* $\varepsilon \geq 0$ *and* $\Delta > 0$*. If*

$$\varepsilon < -K_J + \sqrt{K_J^2 + \frac{\Delta^2}{4}} \quad = \frac{\Delta^2}{8K_J} + \mathcal{O}(K_J^{-2}), \tag{B41}$$

where $K_J = \sqrt{2}\pi^2 J(2J+1)(4J+1)/3$*, then there will be a BCI that is violated.*

Proof. First, since the correlation function C is continuous, it attains its global maximum at some Θ'_+. Since $C(\Theta'_+) \geq C(\Theta_+) \geq 1 - \varepsilon$, the premises of this lemma are also satisfied if Θ_+ is replaced by Θ'_+ – i.e., we can assume without loss of generality that C attains its global maximum at Θ_+.

With these Θ_+ and Θ_-, we use the prescription in Lemma IV, with the angle choices in eq. (B32) to generate the following BCI, which must hold for all even integers $N \geq 2$ if there exists a LHV model:

$$\frac{N}{2}C(\Theta_+ + \delta_N) + \left(\frac{N}{2} - 1\right) C(\Theta_+ - \delta_N) - C(\Theta_-) \leq N - 2. \tag{B42}$$

Let us write $\delta\Theta := \Theta_- - \Theta_+ \mod 2\pi$, such that $\delta_N = \frac{\delta\Theta}{N-1}$. Let $K \in \mathbb{R}$ be any constant such that $C''(x) \geq -K$ for all x; it follows from Lemma V that such K exists, and we will fix K later in accordance with that lemma. Since Θ_+ is a local maximum, we know that $0 \geq C''(\Theta_+) \geq -K$, i.e. $K \geq 0$. The global bound on the second derivative of C gives us

$$C(\Theta_+ \pm x) \geq C(\Theta_+) - K\frac{x^2}{2} \text{ for all } x. \tag{B43}$$

Thus, eq. (B42) implies

$$(N-1)\left(1 - \varepsilon - K\frac{\delta_N^2}{2}\right) - (1 - \Delta) \leq N - 2. \tag{B44}$$

Under what conditions does there exist an even integer $N \geq 2$ such that this inequality is violated, i.e. the existence of a LHV model is ruled out? The negation of this inequality can be rearranged into a quadratic equation in $(N-1)$:

$$\varepsilon(N-1)^2 - \Delta(N-1) + \frac{K}{2}(\delta\Theta)^2 < 0. \tag{B45}$$

If this equation has a solution for some even integer N, then the non-existence of a LHV model follows. If $\varepsilon = 0$, then there will always be a solution for large enough N, recovering Lemma IV. Thus, we here only give further consideration to the case where $\varepsilon > 0$.

Since this quadratic function in $(N-1)$ is positive for large values of $\pm(N-1)$, it is necessary for the existence of a negative value that this function has zeroes over the real numbers. The zeroes are

$$N_\pm - 1 = \frac{\Delta \pm \sqrt{\Delta^2 - 2\varepsilon K(\delta\Theta)^2}}{2\varepsilon}, \tag{B46}$$

and so the following inequality is necessary for the existence of a solution of eq. (B45):

$$\Delta^2 > 2\varepsilon K(\delta\Theta)^2. \tag{B47}$$

If it is satisfied, then the values of $(N-1)_\pm$ are well-defined, and we can continue to argue as follows. The quadratic function in eq. (B45) is negative for all real numbers $N \in (N_-, N_+)$, where $0 < N_- < N_+$. Now, this interval definitely contains an even integer N if $N_+ - N_- > 2$. Since $N_+ - N_- = \sqrt{\Delta^2 - 2\varepsilon K(\delta\Theta)^2}/\varepsilon$, this difference is larger than two if and only if

$$4\varepsilon^2 + 2\varepsilon K(\delta\Theta)^2 - \Delta^2 < 0. \tag{B48}$$

The two solutions of the corresponding quadratic equation are

$$\varepsilon_\pm = \frac{-K(\delta\Theta)^2 \pm \sqrt{K^2(\delta\Theta)^4 + 4\Delta^2}}{4}. \tag{B49}$$

They are both real, and $\varepsilon_- < 0 < \varepsilon_+$. Thus, $\varepsilon < \varepsilon_+$ implies a suitable solution of eq. (B45), i.e. rules out the existence of a LHV model.

In fact, if $\varepsilon < \varepsilon_+$, then we automatically get

$$4\varepsilon < -K(\delta\Theta)^2 + K(\delta\Theta)^2\sqrt{1 + \frac{4\Delta^2}{K^2(\delta\Theta)^4}} \leq \frac{2\Delta^2}{K(\delta\Theta)^2}, \tag{B50}$$

i.e. eq. (B47) is automatically satisfied. Now, considering ε_+ as a function in $\delta\Theta$, this function is decreasing for $\delta\Theta > 0$. Since $\delta\Theta \leq 2\pi$, $\varepsilon < \varepsilon_+(2\pi)$ implies $\varepsilon < \varepsilon_+(\delta\Theta) \equiv \varepsilon_+$. Thus, the inequality

$$\varepsilon < \frac{-K(2\pi)^2 + \sqrt{K^2(2\pi)^4 + 4\Delta^2}}{4} \tag{B51}$$

implies a violation of a BCI. The statement of the lemma now follows from taking the value of K from Lemma V, and by substituting $K_J := \pi^2 K$. □

This has consequence for generic (possibly non-relational) $\mathrm{SO}(2) \times \mathrm{SO}(2)$ settings.

Proof of Theorem 2B. *Consider* $\mathrm{SO}(2) \times \mathrm{SO}(2)$ *correlations* C *for finite maximum (half-)integer* J*. Let* C_{rel} *the relational core of* C *(that is, the function of the form eq.* (B30)*) formed by only including terms of eq.* (4) *where* $m = n$*). Let there be some angle difference* Θ_+*, where* $C_{\mathrm{rel}}(\Theta_+) \geq 1-\varepsilon$*, and some other angle difference* Θ_- *where* $C_{\mathrm{rel}}(\Theta_-) \leq 1 - \Delta$*, with* $\varepsilon \geq 0$ *and* $\Delta > 0$*. If*

$$\varepsilon < -K_J + \sqrt{K_J^2 + \frac{\Delta^2}{4}} \;=\; \frac{\Delta^2}{8K_J} + \mathcal{O}(K_J^{-2}), \tag{B52}$$

where $K_J = \sqrt{2}\pi^2 J(2J+1)(4J+1)/3$*, then there will be a BCI for the (possibly non-relational) correlation function* C *that is violated.*

Proof. Subtracting the "non-relational" parts of $C(\alpha, \beta)$ is equivalent to performing the following integration:

$$C_{\mathrm{rel}}(\alpha, \beta) = \frac{1}{2\pi}\int_0^{2\pi} d\phi\, C(\alpha + \phi, \beta + \phi) \tag{B53}$$

This may be directly verified by noting that terms of the form $\cos(m\alpha - n\beta + (m-n)\,\phi)$ and $\sin(m\alpha - n\beta + (m-n)\,\phi)$ individually integrate to 0 over ϕ except when $m = n$. This allows us to interpret taking the relational core of a correlation function as mixing C over many settings offset by a shared uniform random angle.

It then follows from the convexity of Bell inequalities that if the BCI implied by Lemma VI for the relational core is violated "on average" for this mixture of settings, there must be at least one single set of input settings that also results in that BCI being violated. □

5. Necessity of a bound on J

We will now show that our protocol for witnessing nonlocality does not work if we simply drop the assumption that J is finite.

A correlation function $C(\alpha, \beta)$ has a LHV model if and only if there exists a variable $\lambda \in \Lambda$, distributed via some $P_\Lambda(\lambda)$, and a family of local response functions $C_A(\alpha, \lambda) := \sum_{a\in\{-1,+1\}} a\, P_A(a|\alpha, \lambda)$ and $C_B(\beta, \lambda) := \sum_{b\in\{-1,+1\}} b\, P_B(b|\beta, \lambda)$ such that

$$C(\alpha, \beta) = \int_\Lambda d\lambda P_\Lambda(\lambda) C_A(\alpha, \lambda) C_B(\beta, \lambda). \tag{B54}$$

Suppose that there are two angles θ'_+, θ'_- such that our protocol gives correlation values $C_{\mathrm{rel}}(\theta'_\pm)$ very close to ± 1. If this is the *only* experimental syndrome, without further assumptions on the form of the correlation function (in particular, without any assumption on J as explained in the main text), then this experimental behavior can always be reproduced to arbitrary accuracy by local hidden variables. Namely, $\theta'_\pm$ can be arbitrarily well approximated by angles $\theta_\pm$ that satisfy the premises of the following lemma:

Lemma VII. *Suppose that $\theta_+, \theta_- \in [0, 2\pi)$ are such that $\theta_- - \theta_+ = \frac{m}{n}\pi$, where $n \in \mathbb{N}$ and $m \leq n$ is an odd integer. Then there exists a local relational correlation function $C(\alpha, \beta) \equiv C(\alpha - \beta)$ such that*

$$C(\theta_+) = +1 \text{ and } C(\theta_-) = -1. \tag{B55}$$

Proof. We set $\Lambda = [0, 2\pi)$ and $P_\Lambda(\lambda) = \frac{1}{2\pi}$ – the uniform measure on this interval. Without loss of generality, assume that $\Theta_- > \Theta_+$ and $\Theta_+ = 0$ (we can choose our local coordinates α, β to make this the case). Define $C_A(x, \lambda) = C_B(x, \lambda) := f(x + \lambda)$, where $f : \mathbb{R} \to \{-1, +1\}$ is the 2π-periodic extension of

$$f(x) := \begin{cases} +1 & \text{if } \frac{x}{\pi} \in \left[0, \frac{1}{n}\right) \cup \left[\frac{2}{n}, \frac{3}{n}\right) \cup \ldots \cup \left[\frac{2n-2}{n}, \frac{2n-1}{n}\right) \\ -1 & \text{for all other } x \in [0, 2\pi). \end{cases} \tag{B56}$$

That is, f is a square-wave function of period $\frac{2\pi}{n}$. This f is piecewise continuous and satisfies $f\left(x + \frac{\pi}{n}\right) = -f(x)$ for all x. Thus, $f\left(x + \frac{3}{n}\pi\right) = -f\left(x + \frac{2}{n}\pi\right) = f\left(x + \frac{1}{n}\pi\right) = -f(x)$ for all x, and in particular, by induction, $f\left(x + \frac{m}{n}\pi\right) = -f(x)$ for all x since m is odd by assumption. Now, defining $C(\alpha, \beta)$ as in (B54), this correlation function is relational, since

$$\begin{aligned} C(\alpha + x, \beta + x) &= \frac{1}{2\pi}\int_0^{2\pi} f(\alpha + x + \lambda) f(\beta + x + \lambda)\, d\lambda \\ &= \frac{1}{2\pi}\int_0^{2\pi} f(\alpha + \lambda') f(\beta + \lambda')\, d\lambda' \\ &= C(\alpha, \beta) \end{aligned} \tag{B57}$$

by substitution and due to the (2π)-periodicity of f. Furthermore,

$$\begin{aligned} C(\theta_+) &= C(0, 0) = \frac{1}{2\pi}\int_0^{2\pi} (f(\lambda))^2\, d\lambda = 1, \\ C(\theta_-) &= C\left(\frac{m}{n}\pi, 0\right) = \frac{1}{2\pi}\int_0^{2\pi} f\left(\frac{m}{n}\pi + \lambda\right) f(\lambda)\, d\lambda \\ &= -\frac{1}{2\pi}\int_0^{2\pi} (f(\lambda))^2\, d\lambda = -1. \end{aligned} \tag{B58}$$

□

Therefore, simply following our protocol but relaxing our assumption on J (while not imposing any other assumptions) cannot be sufficient to certify nonlocality.

Appendix C: Characterizing quantum correlations

Let us consider black boxes that have a particularly simple transformation behaviour under rotations:

Definition 2 (Transforming fundamentally)**.** *Consider an $\mathrm{SO}(d)$-box $\mathrm{P}(a \,|\, x)$, where $d \geq 2$. Let $x_0 \in \mathcal{X}$. We say that this box* transforms fundamentally *under rotations if for all $x \in \mathcal{X}$ and all $R_x \in \mathcal{G}$ with $R_x x_0 = x$ one finds*

$$\mathrm{P}(a \,|\, x) \equiv \mathrm{P}(a \,|\, R_x x_0) = c_0^a + \sum_{i,j=1}^{d} (R_x)_{i,j} c_{i,j}^a, \tag{C1}$$

where $(R_x)_{i,j}$ is the fundamental matrix representation of $R_x \in \mathcal{G}$, and $c_0^a, c_{i,j}^a$ are constants independent of x and R_x.

Equivalently, a black box transforms fundamentally if the corresponding representation $R \mapsto T_R$ from Theorem 1 can be chosen as a direct sum of copies of the trivial and the fundamental representations of $\mathcal{G} = \mathrm{SO}(d)$. Since $\mathcal{G}$ is transitive on $\mathcal{X}$, the existence of the above representation is independent of the particular x_0: any alternative $x_0' \in \mathcal{X}$ satisfies $x_0' = S x_0$ for some $S \in \mathcal{G}$, satisfying the above with $R' = RS^{-1}$.

Lemma VIII. *Suppose that $\mathcal{X} = S^{d-1}$ (the unit sphere), and $\mathrm{P}(a \,|\, \vec{x})$ transforms fundamentally under rotations in $\mathcal{G} = \mathrm{SO}(d)$. Then,*

$$\mathrm{P}(a \,|\, \vec{x}) = c_0^a + \vec{c}^{\,a} \cdot \vec{x}, \tag{C2}$$

where constants $c_0^a \in \mathbb{R}$ and $\vec{c}^{\,a} \in \mathbb{R}^d$ satisfy

$$\sum_{a \in \mathcal{A}} c_0^a = 1, \quad \sum_{a \in \mathcal{A}} \vec{c}^{\,a} = \vec{0} \tag{C3}$$

such that for all a, $c_0^a \geq 0$ and $|\vec{c}^{\,a}| \leq \min(c_0^a, 1 - c_0^a)$.

Conversely, if a black box has this form, then it transforms fundamentally under rotations.

In other words, an $\mathrm{SO}(d)$-box transforms fundamentally if and only if $\mathrm{P}(a \,|\, \vec{x})$ is affine-linear in $\vec{x}$ (and non-negativity and normalization of probabilities holds).

Proof. Set $\vec{x}_0 := \vec{e}_1 = (1, 0, \ldots, 0)^{\mathrm{T}}$. Fix some choice of rotations $\vec{x} \mapsto R_{\vec{x}}$ with $R_{\vec{x}}\vec{x}_0 = \vec{x}$. Consider the stabilizer subgroup $\mathcal{G}_{\vec{e}_1} := \{R \in \mathrm{SO}(d) \mid R\vec{e}_1 = \vec{e}_1\}$:

$$\mathcal{G}_{\vec{e}_1} = \left\{ \begin{pmatrix} 1 & \vec{0}^{\mathrm{T}} \\ \vec{0} & T \end{pmatrix} \mid T \in \mathrm{SO}(d-1) \right\}. \tag{C4}$$

The fact that this group is isomorphic to $\mathrm{SO}(d-1)$ is precisely due to the fact that our set of inputs is the homogeneous space $\mathcal{X} = \mathrm{SO}(d)/\mathrm{SO}(d-1)$, i.e. the $(d-1)$-sphere. For $d \geq 3$, we have $\int_{\mathrm{SO}(d-1)} T \,\mathrm{d}T = 0$, and thus

$$\int_{\mathcal{G}_{\vec{e}_1}} S \,\mathrm{d}S = Q := \begin{pmatrix} 1 & \vec{0}^{\mathrm{T}} \\ \vec{0} & \mathbf{0} \end{pmatrix}. \tag{C5}$$

Since $R_{\vec{x}} S \vec{e}_1 = \vec{x}$ for every $S \in \mathcal{G}_{\vec{e}_1}$, every rotation matrix $R_{\vec{x}}S$ can be substituted for R_x into definition 2. Thus $\mathrm{P}(a \,|\, \vec{x}) = c_0^a + \sum_{i,j=1}^{d} (R_{\vec{x}}S)_{i,j} c_{i,j}^a$. By taking the average over all $S \in \mathcal{G}_{\vec{e}_1}$ according to the Haar measure, we get

$$P(a|\vec{x}) = c_0^a + \sum_{i,j=1}^{d} \left(R_{\vec{x}} \int_{\mathcal{G}_{\vec{e}_1}} S \,\mathrm{d}S \right)_{i,j} c_{i,j}^a. \tag{C6}$$

But $R_{\vec{x}}Q = (\vec{x}, \vec{0}, \ldots, \vec{0})$, i.e. a matrix with first column equal to $\vec{x}$ and all further columns equal to zero. This proves that $P(a|\vec{x})$ is affine-linear in $\vec{x}$ as claimed, in the case $d \geq 3$.

Now consider the case $d = 2$. Here, $\vec{x} = (x_1, x_2)^{\mathrm{T}}$, and there is a unique choice of $R_{\vec{x}}$, namely $R_{\vec{x}} =$

$\begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix}$. Then, $P(a|\vec{x})$ being affine-linear in $R_{\vec{x}}$ is equivalent to being affine-linear in $\vec{x}$.

From normalization, $\sum_a \mathrm{P}(a\,|\,\vec{x}) = 1\ \forall \vec{x} \in \mathcal{X}$, which by transitivity of $\mathcal{G}$ on $\mathcal{X}$ can be re-written $\sum_a \mathrm{P}(a\,|\,R\vec{x}_0) = 1\ \forall R \in \mathcal{G}$. Suppose we take the Haar average of $\mathcal{G}$ over both sides of this constraint:

$$\int_{\mathrm{SO}(d)} dR \sum_a \mathrm{P}(a\,|\,R\vec{x}_0) = \sum_a \left(c_0^a + \vec{c}^{\,a} \cdot \int_{\mathrm{SO}(d)} dR\, R\vec{x} \right) = \sum_a c_0^a = 1, \tag{C7}$$

where we have used $\int_{\mathrm{SO}(d)} RdR = 0$. Since $\sum_a \mathrm{P}(a\,|\,R\vec{x}_0) = 1$ for each individual $R \in \mathcal{G}$, we have $(\sum_a \vec{c}^{\,a}) \cdot (R\vec{x}_0) = 0$. Since by transitivity $\dim[\,\mathrm{span}(\{R\vec{x}_0\}_{R\in\mathcal{G}})] = d$, it follows that $\sum_a \vec{c}^{\,a} = \vec{0}$.

For any $\vec{x}$, one may find some $R \in \mathrm{SO}(d)$ such that $R\vec{x} = -\vec{x}$ (since both $\vec{x}$ and $-\vec{x} \in S^{d-1}$ and $\mathrm{SO}(d)$ is transitive on S^{d-1}). Hence, for the black box $\mathrm{P}(a\,|\,\vec{x})$, there is always another black box $\mathrm{P}(a\,|-\vec{x})$ such that the average statistics of these two boxes is given by $\frac{1}{2}\left[\mathrm{P}(a\,|\,\vec{x}) + \mathrm{P}(a\,|-\vec{x})\right] = c_0^a$. Clearly, then $c_0^a \geq 0$. Finally, from the definition of the dot product, $\min_{\vec{x}\in S^{d-1}} (\vec{c}\cdot\vec{x}) = -|\vec{c}|$. Thus, if $|\vec{c}^{\,a}| > c_0^a$, there will be some $\vec{x}$ such that $\mathrm{P}(a\,|\,\vec{x}) = c_0^a - |\vec{c}^{\,a}| < 0$, which is not a valid probability. Similarly $\max_{\vec{x}\in S^{d-1}} (\vec{c}\cdot\vec{x}) = |\vec{c}|$, so if $|\vec{c}^{\,a}| > 1 - c_0^a$, there will be some $\vec{x}$ such that $\mathrm{P}(a\,|\,\vec{x}) = c_0^a + |\vec{c}^{\,a}| \geq 1$. Hence $|\vec{c}^{\,a}| \leq \min(c_0^a, 1 - c_0^a)$.

The converse follows from the transitivity of $\mathrm{SO}(d)$ on S^{d-1}: any $\vec{x}$ can be expressed as $R\vec{x}_0$ for some fixed $\vec{x}_0$ and $R \in \mathrm{SO}(d)$. Thus eq. (C2) can be written $\mathrm{P}(a\,|\,\vec{x}) = c_0^a + \vec{c}^{\,a} \cdot R\vec{x}_0$ which has the form of eq. (C1). □

We can formally define the concept of an "unbiased" black box where if the input orientation is randomized, no particular outcome is preferred:

Definition 3 (Unbiased)**.** *Consider a $\mathcal{G}$-box $\mathrm{P}(a\,|\,x)$ for some compact group $\mathcal{G}$. We say that this box is* unbiased *if the Haar average of $\mathrm{P}(a\,|\,Rx)$ over $R \in \mathcal{G}$ is the same for all $a \in \mathcal{A}$.*

It follows from normalization that if a black box transforms fundamentally under rotations and is unbiased, it may be written in the form $\mathrm{P}(a\,|\,\vec{x}) = \frac{1}{|\mathcal{A}|} + \vec{c}^{\,a} \cdot \vec{x}$.

We extend both these definitions to the local parts of a bipartite system by considering the *conditional boxes*

$$P_A^{b,\vec{y}}(a|\vec{x}) := \mathrm{P}(a,b\,|\,\vec{x},\vec{y})\,/\mathrm{P}_B(b\,|\,\vec{y}) \quad (b,\vec{y}\text{ fixed}), \tag{C8}$$

$$P_B^{a,\vec{x}}(b|\vec{y}) := \mathrm{P}(a,b\,|\,\vec{x},\vec{y})\,/\mathrm{P}_A(a\,|\,\vec{x}) \quad (a,\vec{x}\text{ fixed}), \tag{C9}$$

defined whenever $\mathrm{P}_B(b\,|\,\vec{y}) > 0$ and $\mathrm{P}_A(a\,|\,\vec{x}) > 0$ respectively.

A *conditional box* can be thought of as the black box Alice has if she is told Bob's measurement and outcome. (This is in contrast to a *marginal* black box, which quantifies Alice's statistics when she knows nothing of Bob's measurement or outcome.) No-signalling implies the existence of well-defined marginal boxes $\mathrm{P}_B(b\,|\,\vec{y})$ and $\mathrm{P}_A(a\,|\,\vec{x})$.

Definition 4. *We say that a no-signalling bipartite box $\mathrm{P}(a,b\,|\,\vec{x},\vec{y})$* transforms fundamentally locally *(is* locally unbiased*) if all conditional boxes transform fundamentally (are unbiased).*

The next two lemmas show that these properties are preserved by convex combinations of boxes.

Lemma IX. *Let $\{\mathrm{P}_i(a,b\,|\,\vec{x},\vec{y})\}_{i=1,\ldots N}$ be a set of no-signalling bipartite black boxes that transform fundamentally locally. Any convex combination $\mathrm{P}(a,b\,|\,\vec{x},\vec{y}) := \sum_i \lambda_i \mathrm{P}_i(a,b\,|\,\vec{x},\vec{y})$ where all $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$ also transforms fundamentally locally.*

Proof. First, we calculate the marginal distribution: $\mathrm{P}_A(a\,|\,\vec{x}) := \sum_b \mathrm{P}(a,b\,|\,\vec{x},\vec{y}) = \sum_b \sum_i \lambda_i \mathrm{P}_i(a,b\,|\,\vec{x},\vec{y}) = \sum_i \lambda_i \sum_b \mathrm{P}_i(a,b\,|\,\vec{x},\vec{y}) = \sum_i \lambda_i \mathrm{P}_{A,i}(a\,|\,\vec{x})$. Similarly, $\mathrm{P}_B(b\,|\,\vec{y}) = \sum_i \lambda_i \mathrm{P}_{B,i}(b\,|\,\vec{y})$. First, we note that $\mathrm{P}_B(b\,|\,\vec{y}) = 0$ only if $\mathrm{P}_{B,i}(b\,|\,\vec{y}) = 0$ for all i. In this case, the combined conditional box is undefined, and there is nothing to prove. Thus, we may proceed with the case that $\mathrm{P}_B(b\,|\,\vec{y}) > 0$.

With $\mathrm{P}(a,b\,|\,\vec{x},\vec{y}) = \sum_i \lambda_i \mathrm{P}_i(a,b\,|\,\vec{x},\vec{y}) = {\sum_i}' \lambda_i \mathrm{P}_{B,i}(b\,|\,\vec{y})\, \mathrm{P}_{A,i}^{b,\vec{y}}(a|\vec{x})$ we obtain

$$\mathrm{P}_A^{b,\vec{y}}(a|\vec{x}) = {\sum_i}' \frac{\lambda_i \mathrm{P}_{B,i}(b\,|\,\vec{y})}{\mathrm{P}_B(b\,|\,\vec{y})} \mathrm{P}_{A,i}^{b,\vec{y}}(a|\vec{x}) \tag{C10}$$

Here, ${\sum_i}'$ denotes a sum over all those i for which $P_{B,i}(b|\vec{y}) > 0$. These are exactly the i for which $\mathrm{P}_{A,i}^{b,\vec{y}}(a|\vec{x})$ is well-defined. Meanwhile, ${\sum_i}' \frac{\lambda_i \mathrm{P}_{B,i}(b\,|\,\vec{y})}{\mathrm{P}_B(b\,|\,\vec{y})} = 1$, and hence we may define $\mu_i := \frac{\lambda_i \mathrm{P}_{B,i}(b\,|\,\vec{y})}{\mathrm{P}_B(b\,|\,\vec{y})}$ for those i with $P_{B,i}(b|\vec{y}) > 0$, and $\mu_i = 0$ for all other i, such that $\mu_i \geq 0$ and moreover $\sum_i \mu_i = {\sum_i}' \mu_i = 1$. Thus, the new conditional box is itself a convex combination of the constituent conditional boxes. A similar convex combination can be found for $\mathrm{P}_B^{a,\vec{x}}(b|\vec{y})$.

Since the constituent conditional boxes transform fundamentally, from Lemma VIII we write $\mathrm{P}_{A,i}^{b,\vec{y}}(a|\vec{x}) = c_{i,0}^{(b,\vec{y}),a} + \vec{c}_i^{\,(b,\vec{y}),a} \cdot \vec{x}$. Then, $\mathrm{P}_A^{b,\vec{y}}(a|\vec{x}) = \sum_i \mu_i \left(c_{i,0}^{(b,\vec{y}),a} + \vec{c}_i^{\,(b,\vec{y}),a} \cdot \vec{x} \right) = c_0^{(b,\vec{y}),a} + \vec{c}^{\,(b,\vec{y}),a} \cdot \vec{x}$ where $c_0^{(b,\vec{y}),a} := \sum \mu_i c_{i,0}^{(b,\vec{y}),a}$ and $\vec{c}^{\,(b,\vec{y}),a} := \sum \mu_i \vec{c}_i^{\,(b,\vec{y}),a}$. By the converse part of Lemma VIII, $\mathrm{P}_A^{b,\vec{y}}(a|\vec{x})$ is thus a valid black box that transforms fundamentally. The same argument holds for Bob's conditional boxes. □

Lemma X. *Let $\{\mathrm{P}_i(a,b\,|\,\vec{x},\vec{y})\}_{i=1,\ldots N}$ be a set of non-signalling black boxes that are locally unbiased with respect to $\mathcal{G}$. Any convex combination $\mathrm{P}(a,b\,|\,\vec{x},\vec{y}) := \sum_i \lambda_i \mathrm{P}_i(a,b\,|\,\vec{x},\vec{y})$ where all $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$ is also locally unbiased with respect to $\mathcal{G}$.*

Proof. In the proof of Lemma IX we have seen that there exists a probability distribution $\{\mu_i\}_i$ such that $\mathrm{P}_A^{b,\vec{y}}(a|\vec{x}) = \sum_i \mu_i \mathrm{P}_{A,i}^{b,\vec{y}}(a|\vec{x})$. Thus we find

$$\int \mathrm{P}_A^{b,\vec{y}}(a|R\vec{x})\,\mathrm{d}R = \sum_i \mu_i \int \mathrm{P}_{A,i}^{b,\vec{y}}(a|R\vec{x})\,\mathrm{d}R = \frac{\sum_i \mu_i}{|\mathcal{A}|} = \frac{1}{|\mathcal{A}|}. \tag{C11}$$

Likewise holds for Bob's conditional boxes, and hence, $\mathrm{P}(a,b\,|\,\vec{x},\vec{y})$ is also locally unbiased. □

Lemma XI. *Consider a bipartite black box with inputs $\mathcal{X}=\mathcal{Y}=S^{d-1}$ and binary outcomes $\mathcal{A}=\mathcal{B}=\{+1,-1\}$. If this box transforms fundamentally locally and is locally unbiased, then it describes quantum correlations.*

Proof. From Lemma VIII, binary-outcome conditional boxes that transform fundamentally and are unbiased can be written:

$$\mathrm{P}_A^{b,\vec{y}}(a|\vec{x}) := \begin{pmatrix} \frac{1}{2} \\ a\vec{c}^{(b,\vec{y})} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \vec{x} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \vec{c}^{(b,\vec{y})} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ a\vec{x} \end{pmatrix}, \tag{C12}$$

$$\mathrm{P}_B^{a,\vec{x}}(b|\vec{y}) := \begin{pmatrix} \frac{1}{2} \\ b\vec{c}^{(a,\vec{x})} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \vec{y} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \vec{c}^{(a,\vec{x})} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ b\vec{y} \end{pmatrix}, \tag{C13}$$

such that the joint probability distribution is given by:

$$\mathrm{P}(a,b\,|\,\vec{x},\vec{y}) = \mathrm{P}_B(b|\vec{y}) \begin{pmatrix} \frac{1}{2} \\ \vec{c}^{(b,\vec{y})} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ a\vec{x} \end{pmatrix}, \tag{C14}$$

$$\mathrm{P}(a,b\,|\,\vec{x},\vec{y}) = \mathrm{P}_A(a|\vec{x}) \begin{pmatrix} \frac{1}{2} \\ \vec{c}^{(a,\vec{x})} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ b\vec{y} \end{pmatrix}. \tag{C15}$$

This defines a map $\tilde{\omega}_{AB}$ acting on $\vec{e}_{a,\vec{x}} := (1, a\vec{x})^{\mathrm{T}}$ and $\vec{e}_{b,\vec{y}} := (1, b\vec{y})^{\mathrm{T}}$, such that $\tilde{\omega}_{AB}(\vec{e}_{a,\vec{x}}, \vec{e}_{b,\vec{y}}) = \mathrm{P}(a,b\,|\,\vec{x},\vec{y})$. Moreover, $\mathrm{span}(\vec{e}_{a,\vec{x}}) = \mathrm{span}(\vec{e}_{b,\vec{y}}) = \mathbb{R}^{d+1}$, and so this function has a unique bilinear extension $\omega_{AB} : \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \to \mathbb{R}$.

The set of non-negative linear combinations of $\vec{e}_{a,\vec{x}}$ define a *positive Euclidean cone* $A^+ \subset \mathbb{R}^{d+1}$, whose extremal rays are the non-negative multiples of $(1,\vec{z})^{\mathrm{T}}$ for $\vec{z} \in S^{d-1}$. We may then define an *Archimedean order unit* (AOU) [52], $\vec{u} := (2,0,\ldots,0)^{\mathrm{T}} \in \mathbb{R}^{d+1}$ and define an AOU-space $(\mathbb{R}^{d+1}, A^+, \vec{u})$. An identical AOU-space $(\mathbb{R}^{d+1}, B^+, \vec{u})$ can be defined using the non-negative linear combinations of $\vec{e}_{b,\vec{y}}$.

Now, we shall employ a result from Kleinmann *et al.* [52] (generalizing a result by Barnum *et al.* [53]) that pertains to bilinear maps on positive Euclidean cones. If a bilinear map ω_{AB} on such cones is both *unital* and *positive*, then there exists a bipartite quantum system ρ_{AB} and a map from each point $\vec{a} \in A^+$, $\vec{b} \in B^+$ onto local quantum POVM elements M_a, M_b such that $\omega_{AB}(\vec{a},\vec{b}) = \mathrm{tr}\,(\rho_{AB} M_a \otimes M_b)$.

We show that ω_{AB} satisfies these conditions. First, for any given a, $\vec{x}$ (likewise b, $\vec{y}$), it can be seen that

$$\vec{e}_{+a,\vec{x}} + \vec{e}_{-a,\vec{x}} = \vec{u} = \vec{e}_{+b,\vec{y}} + \vec{e}_{-b,\vec{y}}, \tag{C16}$$

and hence $\omega_{AB}(\vec{u},\vec{u}) = \sum_{a,b} \mathrm{P}(a,b\,|\,\vec{x},\vec{y}) = 1$, which means that ω_{AB} is *unital*. Next, since every $\vec{p} \in A^+$ can be written as a non-negative linear combination of finitely many $e_{a,\vec{x}}$ (likewise for $\vec{q} \in B^+$), then $\omega_{AB}(\vec{p},\vec{q}) \geq 0$ for all $\vec{p},\vec{q} \in A^+, B^+$, showing that ω_{AB} is *positive*. Hence, ω_{AB} can be realised by a quantum system, and $\mathrm{P}(a,b\,|\,\vec{x},\vec{y})$ is a quantum behaviour. □

The premise of local unbiasedness cannot be removed: if we only demand that a box transforms fundamentally locally, then it can generate correlations that are disallowed by quantum theory. To see this, let P_0 be any non-signalling $(2,2,2)$-behaviour (for example, a PR-box), and define

$$\begin{aligned}\mathrm{P}(a,b\,|\,\vec{x},\vec{y}) := {} & \lambda_A\lambda_B P_0(a,b|0,0) + \lambda_A\bar{\lambda}_B P_0(a,b|0,1) \\ & + \bar{\lambda}_A\lambda_B P_0(a,b|1,0) + \bar{\lambda}_A\bar{\lambda}_B P_0(a,b|1,1),\end{aligned} \tag{C17}$$

where $a,b \in \{-1,+1\}$, $\vec{x},\vec{y} \in S^{d-1}$, $\lambda_A := \frac{1}{2}(1+x_1)$, $\bar{\lambda}_A := 1-\lambda_A$, $\lambda_B := \frac{1}{2}(1+y_1)$, and $\bar{\lambda}_B := 1-\lambda_B$. If $d=3$, for example, this describes a situation in which two possible local inputs $x,y \in \{0,1\}$ are encoded into the first Bloch vector component of a qubit, the qubits are locally measured, and the measurement results $x,y \in \{0,1\}$ are input into the original box P_0. This defines a valid non-signalling box, and the linear dependence of the outcome probabilities on $\vec{x}$ (resp. $\vec{y}$) demonstrate, via Lemma VIII, that P transforms fundamentally locally. However, it reproduces P_0 via $P_0(a,b|r,t) = \mathrm{P}(a,b\,|\,\vec{x}_r,\vec{y}_t)$, where $\vec{x}_0 = \vec{y}_0 = (1,0,\ldots,0)$ and $\vec{x}_1 = \vec{y}_1 = (-1,0,\ldots,0)$.

Now we show a converse statement, so that we get an exact classification of the quantum $(2,2,2)$-behaviours: i.e. the family of probabilities obtained in quantum theory during a two party Bell test, where each agent has two choices of input and obtains one of two outcomes.

Lemma XII. *Let $d \geq 2$. Then all extremal quantum (2,2,2)-behaviours can be realized by locally unbiased $\mathrm{SO}(d)$-boxes that transform fundamentally locally with $\mathcal{X}_A = \mathcal{X}_B = S^{d-1}$; the two settings (inputs) correspond to two choices of directions.*

Proof. It has been shown [25, 34, 54, 55] that the extremal quantum $(2,2,2)$-behaviours can be realised by rank-1 projective measurements on two-qubit pure states. Any extremal quantum $(2,2,2)$-behaviour $\mathrm{P}(a,b\,|\,r,t)$ can then be written in the form $\mathrm{P}(a,b\,|\,r,t) = \mathrm{Tr}[\rho_{AB}(E_r^{(a)} \otimes F_t^{(b)})]$ where ρ is a pure state of two qubits and $E_r^{(a)}$ and $F_t^{(b)}$ are qubit rank-1 projectors, $a,b \in \{-1,+1\}$ and $r,t \in \{1,2\}$. We shall show that there exists a non-signalling $\mathrm{SO}(d) \times \mathrm{SO}(d)$-box $\mathrm{P}(a,b\,|\,\vec{x},\vec{y})$ that transforms fundamentally locally, is locally unbiased, and has choices $\vec{x}_r,\vec{y}_t$ such that $\mathrm{P}(a,b\,|\,\vec{x}_r,\vec{y}_t) = \mathrm{P}(a,b\,|\,r,t)$.

Write $E := E_1^{(+1)}$ and $\tilde{E} := E_2^{(+1)}$. As $\mathrm{Tr}\,E = 1$, its expansion in the Pauli operator basis is of the form $E = \frac{1}{2}(\mathbb{1}_2 + E_x\sigma_x + E_y\sigma_y + E_z\sigma_z)$, and the associated Bloch vector $\vec{x}_1' := (E_x,E_y,E_z)^{\mathrm{T}}$ has Euclidean norm 1. Let $\vec{x}_2'$ be the Bloch vector similarly associated with $\tilde{E}$. By changing the local bases unitarily, we can ensure that $\vec{x}_1' = (1,0,0)^{\mathrm{T}}$ and $\vec{x}_2' = (\cos\theta, \sin\theta, 0)^{\mathrm{T}}$, where $0 \leq \theta < 2\pi$. Similarly, we can define the Bloch vectors $\vec{y}_t'$ for $t=1,2$ via the rank-1 projections $F_t^{(b)}$.

Define a linear map $\Pi : \mathbb{R}^d \to \mathbb{R}^3$ in the following way. If $d=2$, set $\Pi(v_1,v_2)^{\mathrm{T}} := (v_1,v_2,0)^{\mathrm{T}}$; if $d \geq 3$, set $\Pi(v_1,\ldots,v_d)^{\mathrm{T}} := (v_1,v_2,v_3)^{\mathrm{T}}$, which is an orthogonal

projection (and the identity if $d = 3$). Furthermore, for $\vec{v} \in \mathbb{R}^3$, define $E_{\vec{v}} := \frac{1}{2}(\mathbb{1}_2 + \vec{v} \cdot \vec{\sigma})$, which is positive-semidefinite whenever $|\vec{v}| \leq 1$. Consider the non-signalling $\mathrm{SO}(d) \times \mathrm{SO}(d)$-box

$$\mathrm{P}(a, b \,|\, \vec{x}, \vec{y}) := \mathrm{Tr}\left[\rho_{AB} E_{a\Pi\vec{x}} \otimes E_{b\Pi\vec{y}}\right]. \tag{C18}$$

Since $|a\Pi\vec{x}| \leq |\vec{x}| = 1$ for $\vec{x} \in S^{d-1}$ (and similarly for $b\Pi\vec{y}$), this defines a valid (quantum) behaviour. The conditional boxes are

$$P_A^{b,\vec{y}}(a|\vec{x}) = \frac{\frac{1}{2}\,\mathrm{Tr}\left[\rho_{AB}(\mathbb{1} + a\vec{x} \cdot (\Pi^{\mathrm{T}}\vec{\sigma})) \otimes E_{b\Pi\vec{y}}\right]}{\mathrm{Tr}(\rho_B E_{b\Pi\vec{y}})}. \tag{C19}$$

This expression yields well-defined probabilities by construction, and it is affine-linear in $\vec{x}$. Analogous statements hold for the other conditional boxes. Thus, according to Lemma VIII, $\mathrm{P}(a, b \,|\, \vec{x}, \vec{y})$ transforms fundamentally locally. Furthermore, averaging the above conditional box uniformly over $\vec{x}$ replaces $\vec{x}$ by zero and annihilates all dependence on a; hence this box is locally unbiased.

Let $\vec{x}_r \in \mathbb{R}^d$ be the vector whose first two components are the first two components of $\vec{x}'_r$, and all other $(d-2)$ components are zero; define $\vec{y}_t$ analogously. Then $\mathrm{P}(a, b \,|\, r, t) = \mathrm{P}(a, b \,|\, \vec{x}_r, \vec{y}_t)$. □

Proof of Theorem 3. *Let $d \geq 2$. The quantum $(2, 2, 2)$-behaviours are exactly those that can be realised by binary-outcome bipartite $\mathrm{SO}(d) \times \mathrm{SO}(d)$-boxes that transform fundamentally locally and are locally unbiased, restricted to two choices of input direction per party per box, and statistically mixed via shared randomness.*

Proof. Lemma XI tells us that "$(2, S^{d-1}, 2)$–behaviours" that transform fundamentally locally, and are locally unbiased, can be realised by local measurements on a bipartite quantum system. If we restrict our choice of inputs from the full S^{d-1} freedom to just two choices of orientation per party, then these will be $(2, 2, 2)$–behaviours, and since they can be realised by a quantum system, they are quantum $(2, 2, 2)$–behaviours.

The other direction follows from Lemma XII: all *extremal* quantum $(2, 2, 2)$-behaviours can be realised by restricting binary-outcome bipartite $\mathrm{SO}(d) \times \mathrm{SO}(d)$-boxes, transforming fundamentally locally and being locally unbiased, to two possible input directions per party. Additional shared randomness allows the two parties to generate all statistical mixtures of these behaviours, yielding all further quantum $(2, 2, 2)$-behaviours. □

Theorem 3 cannot hold for all $d \geq 2$ without allowing shared randomness. For example, suppose that $d = 3$, then the proof of Lemma XI shows that all correlations realizable with binary-outcome bipartite $\mathrm{SO}(3) \times \mathrm{SO}(3)$-boxes that transform fundamentally locally and are locally unbiased can be realized via unital positive bilinear forms on the positive semidefinite qubit cone. Consequently, the result by Barnum *et al.* [53] implies that all these correlations can also be realized via POVMs on ordinary two-qubit quantum state space. However, Donohue and Wolfe [36] (extending results by Pál and Vértesi [35]) have shown that the set of $(2, 2, 2)$-behaviours realizable on two qubits via POVMs is not convex, and thus not equal to the convex set of quantum $(2, 2, 2)$-behaviours.

Remarks about the Bell witness protocol:
In our Bell witness protocol, one agent (Bob) does not need a random number generator. In that regard, our protocol is simpler than a usual Bell experiment. Instead, there is a random number generator in the beginning of the protocol which generates the shared random angle λ. In the main text, our explanation of the protocol says that Bob inputs λ. This could be interpreted as Bob having an input after all, although he cannot freely choose it while separated from the other agent (in particular, the other agent Alice knows this input of Bob, i.e. λ). However, one could also consider a modified scenario in which the shared random angle is a property of the boxes themselves, or the preparation procedure. In this modification, the device of Bob is hard-wired to λ during preparation, i.e. his device is already set to λ in the preparation phase and Bob does not input anything at all. Similarly, Alice's device is preset to λ and she just modifies this angle by adding Θ_+ or Θ_- to it. In this modified setting, Bob indeed does not input anything at all, while Alice only chooses one of two angles.

The most important simplifications arise if one knows in advance that the full correlation function already is relational, i.e. $C(\alpha, \beta) = C_{\text{rel}}(\alpha - \beta)$. For example, there might be physical arguments that imply that a considered preparation procedure has no reference for an absolute angle. In such cases, the randomization with the shared random angle λ is unnecessary. Then we can completely remove the initial random number generator, i.e. λ, as a common cause.

Once more, we point out that the Bell witness protocol is **NOT** a full Bell test, because the input of Bob's device is known to Alice or her device in advance. Instead, the protocol just collects statistics that implies that there exists a Bell test one could perform instead to find Bell non-locality.

Chapter 6

Quantum computation is the unique reversible circuit model for which bits are balls

Authors: Marius Krumm, Markus P. Müller
Published in: npj Quantum Information **5**, 7 (2019)
DOI: https://doi.org/10.1038/s41534-018-0123-x
Preprint available as: arXiv:1804.05736
https://arxiv.org/abs/1804.05736

Contribution: Markus suggested the original idea for a project to construct interactions of higher dimensional Bloch balls. Instead, I discovered the no-go theorem and developed the first full proof.

Abstract: The computational efficiency of quantum mechanics can be defined in terms of the qubit circuit model, which is characterized by a few simple properties: each computational gate is a reversible transformation in a connected matrix group; single wires carry quantum bits, i.e. states of a three-dimensional Bloch ball; states on two or more wires are uniquely determined by local measurement statistics and their correlations. In this paper, we ask whether other types of computation are possible if we relax one of those characteristics (and keep all others), namely, if we allow wires to be described by d-dimensional Bloch balls, where d is different from three. Theories of this kind have previously been proposed as possible generalizations of quantum physics, and it has been conjectured that some of them allow for interesting multipartite reversible transformations that cannot be realized within quantum theory. However, here we show that all such potential beyond-quantum models of computation are trivial: if d is not three, then the set of reversible transformations consists entirely of single-bit gates, and not even classical computation is possible. In this sense, qubit quantum computation is an island in theoryspace.

I. INTRODUCTION

Since the discovery of quantum algorithms that outperform all known classical ones in certain tasks [1], improving our understanding of the possibilities and limitations of quantum computation has become one of the central goals of quantum information theory. While it is notoriously difficult to prove unconditional separation of polynomial-time classical and quantum computation [2], an approach that is often regarded more tractable is to analyze how certain modifications of quantum computing affect its computational power. For instance, one may consider restrictions on the set of allowed quantum resources, and ask under which condition the possibility of universal quantum computation is preserved despite the restriction. Notable results along these lines, among many others, include the Gottesman-Knill theorem [3–5], insights on the necessity of contextuality as a resource for magic state distillation [6], or bounds on the noise threshold of quantum computers [7].

In a complementary and in some sense more radical approach, going back to Abrams and Lloyd [8], one considers modifications of the quantum formalism itself and studies the impact of those modifications on the computational efficiency, resembling strategies of classical computer science such as the introduction of oracles [9]. For example, it has been shown that availability of closed timelike curves leads to implausible computational power [10], that stronger-than-quantum nonlocality reduces the set of available transformations [11–14], that tomographic locality forces computations to be contained in a class called AWPP [15, 16], and that in some theories (satisfying additional axioms) higher-order interference does not lead to a speed-up in Grover's algorithm [17]. Further examples can be found e.g. in [18–21].

In this paper, we consider a specific modification of the quantum formalism that is arguably among the simplest and most conservative possibilities. This modification dates back to ideas by Jordan, von Neumann, and Wigner [22], and it has several independent motivations as we will explain further below. This generalization keeps all characteristic properties of quantum computation unchanged, but modifies a single aspect: namely, it allows the quantum bit to have any number of $d \geq 2$ degrees of freedom, instead of standard quantum theory's $d = 3$ (or the classical bit's $d = 1$). It has been conjectured [23] that the resulting theories allow for interesting "beyond quantum" reversible multipartite dynamics, which would make the corresponding models of computation highly relevant objects of study within the research program mentioned above. However, here we show that, quite on the contrary, these models are so constrained that they do not even allow for classical computation; hence, in Aaronson's terminology, the $d = 3$ case of the standard qubit circuit model can be seen as an "island in theoryspace" [24].

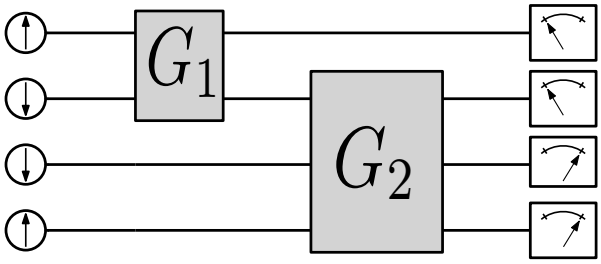


FIG. 1. The circuit model that we consider in this paper. We have an arbitrary finite number n of wires (here $n = 4$), and each wire carries a "gbit" which is a state in a d-dimensional Bloch ball state space. Initially, a product state is prepared (encoding, for example, the classical input to the algorithm), then a finite number of gates G_i is applied, each acting on an arbitrary number of gbits, and finally local measurements are performed. We assume that the G_i are elements of an (arbitrary unspecified) closed connected matrix group, and that the global state of n wires is uniquely determined by the statistics and correlations of single-wire measurements ("tomographic locality"). If $d = 3$, i.e. if the gbits are qubits, it has been shown in [50] that these assumptions uniquely characterize unitary quantum computation as the only computationally non-trivial theory. Here we analyze the case $d \neq 3$, and prove that — despite conjectures to the opposite [23] — the corresponding models do not allow for any non-trivial computation at all. We do *not* assume that wires can be swapped, or that all transformations can be composed out of two-gbit transformations. See the main text for details.

Our paper is organized as follows. Section II gives the mathematical framework. We define single bits that generalize the qubit ("gbits"), and then give three postulates that allow us to reason about circuits that are constructed out of n of these gbits. We formulate the problem that is addressed in this work and describe how it relates to earlier results in the literature. In Section III, we state and prove our main result: namely, while our principles uniquely determine quantum computation in the case that the single gbits have $d = 3$ degree of freedom, any other value of d does not even allow for classical computation. We give the full proof for the case $d \geq 4$ (the $d = 2$ case is deferred to the appendix), and illustrate the main idea of some of the proof steps by a circuit diagram, before concluding in Section IV.

II. GENERALIZED BITS AND GBIT CIRCUITS

In both classical and quantum computation, we can restrict our attention to the circuit model (as in Figure 1) where each of the wires (the single systems that enter and exit logical gates) corresponds to a two-level system. Quantum two-level systems (qubits) are different from classical ones (bits): they allow for a more complex behavior which encompasses phenomena like coherent superposition, interference, or uncertainty relations. Yet, both classical and quantum bits can be for-

malized in a unified way that we now describe (for both single and multiple bits, i.e. circuits, we follow the constructions and notation from [50]).

A. Single gbits

To any $d \in \mathbb{N}$, we associate a "generalized bit" (gbit) that has the d-dimensional Bloch ball, $B^d = \{\vec{a} \in \mathbb{R}^d \mid |\vec{a}| \leq 1\}$, as its state space. Every vector $\vec{a}$ in the Bloch ball B^d corresponds to a possible state of the generalized bit. Two-outcome measurements are described by vectors $\vec{b} \in \mathbb{R}^d$ with $|\vec{b}| = 1$, such that the probability of the first outcome if performed on state $\vec{a}$ is $(1+\vec{a}\cdot\vec{b})/2$, and that of the second outcome is $(1-\vec{a}\cdot\vec{b})/2$. In the following, it will be convenient to use the notation $v(\vec{a}) = (1, \vec{a})^\top \in \mathbb{R}^{d+1}$, such that these two probabilities become $\frac{1}{2}v(\vec{a}) \cdot v(\pm\vec{b})$. Reversible transformations of states are given by $\vec{a} \mapsto R\vec{a}$, where $R \in \mathrm{SO}(d)$ is a rotation matrix. These transformations map states to states and can be inverted (by applying R^{-1}), hence we can interpret them as closed-system time evolutions or, equivalently, reversible gates on single generalized bits.

For $d = 3$, this formalism recovers the qubit of standard quantum theory [5]: as is well-known, every 2×2 density matrix ρ can be written in the form

$$\rho = (\mathbf{1} + \vec{a}_\rho \cdot \vec{\sigma})/2,$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ denotes the Pauli matrices. It is automatic in this representation that $\mathrm{tr}\,\rho = 1$, and positivity $\rho \geq 0$ is equivalent to $|\vec{a}_\rho| \leq 1$. Hence the set of states of a quantum bit can be represented by the Bloch ball B^3. This representation has the important property that *statistical mixtures correspond to convex combinations*: if a state ρ is prepared with probability p and another state ρ' is prepared with probability $1-p$, then the total state $p\rho + (1-p)\rho'$ corresponds to the Bloch vector $\vec{a}_{p\rho+(1-p)\rho'} = p\vec{a}_\rho + (1-p)\vec{a}_{\rho'}$. This statistical interpretation of convex mixtures is also taken for balls of other dimensions $d \neq 3$, hence these Bloch balls can be regarded as state spaces of generalized probabilistic theories [11].

In the $d = 3$ case, projective measurements are represented by unit vectors $\vec{b}$, $|\vec{b}| = 1$, with outcome probabilities $(1 \pm \vec{a} \cdot \vec{b})/2$ as described above. Unitary transformations U on states, acting as $\rho \mapsto U\rho U^\dagger$, are described in the Bloch ball picture by orthogonal maps R_U, $R_U^\top R_U = \mathbf{1}$, such that $\vec{a}_{U\rho U^\dagger} = R_U\vec{a}_\rho$. More general measurements (positive operator-valued measures) or transformations (completely positive maps) can also be described in the Bloch ball representation, but they are not needed in what follows and therefore omitted.

The simplest case of $d = 1$ corresponds to the classical bit: there are two possible configurations, $\vec{a} = +1$ and $\vec{a}' = -1$, and further states that represent classical uncertainty about the configuration. Namely, if we have $+1$ with probability p (and thus -1 with probability $1-p$), this corresponds to the state $p\vec{a} + (1-p)\vec{a}'$ in the interior the one-dimensional "Bloch ball".

There is one peculiarity in the $d = 1$ case: instead of $\mathrm{SO}(1) = \{\mathbf{1}\}$, we should allow the group $\mathrm{O}(1) = \{-\mathbf{1}, \mathbf{1}\}$ as Bloch ball transformations such that also the bit flip is allowed.

What is the significance of the d-dimensional Bloch balls if d is neither one nor three? These gbits have appeared in various places in quantum information theory and the foundations of quantum mechanics. Historically, they have first shown up as precisely those two-level state spaces that can be described as (formally real, irreducible) Jordan algebras [22], a natural algebraic generalization of standard quantum theory. In fact, quantum theory with real amplitudes, i.e. over the field $\mathbb{R}$ instead of $\mathbb{C}$, has a $(d = 2)$-dimensional Bloch ball as its "quantum bit", and the bits of quaternionic and octonionic quantum theory correspond to B^d for $d = 5$ and $d = 9$ respectively. Furthermore, the fact that a two-level system should have a Euclidean ball state space can be derived from a variety of different sets of natural assumptions. In many reconstructions of quantum theory from physical or information-theoretic principles [25–32, 34–36], this fact is derived as a first step. For example, postulating that the group of reversible transformations acts transitively on the pure states implies that the pure states must all lie on the unit hypersphere of an invariant inner product. If some points on the sphere were *not* valid states, then there would exist additional measurements that would violate further natural postulates like Hardy's [25] "Subspaces" axiom. This argumentation or others along similar lines [25–36] lead to Euclidean balls as the most natural state spaces of a generalized bit.

A more geometrical motivation can be found by considering spin-$\frac{1}{2}$ particles (compare e.g. to [23]): under rotations $\mathrm{SO}(3)$, they transform via $\mathrm{SU}(2)$. The density matrix transforms under the adjoint representation, which means that the Bloch vectors transform via the same rotation as in physical space. Therefore, the Bloch vector $\vec{b}$ can be seen as defining an oriented axis in physical space. The model considered in this paper is a direct generalization of the Bloch ball and this interpretation to arbitrary spatial dimensions. Indeed, the possibility that space might have more than three dimensions has appeared in a large variety of physical theories, see e.g. [37–42]. It has also been argued that these generalized bits can be interpreted as "information quasiparticles" in some sense [43]. In summary, these gbits are among the simplest and most natural generalizations of the classical bit and the qubit of quantum mechanics.

B. Several gbits and computation

To describe circuit computation, we need to define the state space, measurements, and transformations of several gbits. In standard quantum theory, where the gbits are qubits, there is a unique definition of these notions:

the states of n qubits are exactly the $(2^n) \times (2^n)$ density matrices, the reversible transformations are the unitaries, and the measurements are described by collections of projection operators. Similar definitions apply to n classical bits. But if the gbits are Bloch balls of dimension $d \notin \{1, 3\}$, then it is apriori unclear what the composite state space should be.

Since we would like to be as general as possible, we will not make any attempt to fix the composite state space from the outset. Instead, we will work with a small set of principles that the composite n-gbit system is supposed to satisfy. While these principles will constrain the n-gbit state space, it is by no means obvious that they determine it uniquely. However, we will show below that they are indeed constraining enough to allow us to derive the full set of states and transformations.

An important principle is the **no-signalling principle** [11]: *the outcome statistics of measurements on any group of gbits does not depend on any other operations (e.g. measurements) that are performed on the remaining gbits.* This is a physically well-motivated constraint that lies at the heart of what we mean by "different wires" (i.e. subsystems) of the circuit in the first place.

This principle is satisfied by classical as well as quantum computation, and so is our second postulate of **tomographic locality** [25, 44]: *every state on n gbits is uniquely characterized by the statistics and correlations of the local gbit measurements.* In other words, a global n-gbit state is nothing but a catalog of probabilities for the outcomes of all the single-gbit measurements and their correlations.

It is not only classical and quantum theory that satisfies the principle of tomographic locality, but also more general probabilistic theories like boxworld [12]. If this principle was violated, then a collection of gbits would in some counterintuitive sense be "more" than a composition of its building blocks. Even though this formulation makes tomographic locality sound very natural, there are simple examples of theories that violate it. One such example is given by quantum theory over the real numbers $\mathbb{R}$ [45, 46]. This is because observables of two single real qubits do not linearly generate all observables of two real qubits. In particular, if σ_y is the Pauli matrix with purely imaginary entries, then σ_y is not a real qubit observable, but $\sigma_y \otimes \sigma_y$ is a real two-qubit observable. Intuitively, it represents a novel "holistic" degree of freedom that cannot be constructed out of local degrees of freedom and their correlations.

Not only is the postulate of tomographic locality very intuitive, but it is also very powerful: it allows us to represent states of n gbits as *tensors* [11]. That is, even if we do not know what the set of n-gbit states is, we know that every such state can be written as an element of the linear space $(\mathbb{R}^{d+1})^{\otimes n}$ (in the quantum case, where $d = 3$, this amounts to the 4^n-dimensional real linear space of Hermitian $(2^n) \times (2^n)$ matrices; for real bits, it is the 2^n-dimensional space that contains the probability vectors over 2^n configurations). In particular, an n-gbit product state with local Bloch vectors $\vec{a}_1, \ldots \vec{a}_n$ is represented by

$$v(\vec{a}_1, \ldots, \vec{a}_n) := (1, \vec{a}_1)^\top \otimes \ldots \otimes (1, \vec{a}_n)^\top,$$

and all other states ω are vectors on the same space (but not of this product form). Tomographic locality then amounts to the fact that all these states are uniquely determined by the numbers

$$2^{-n} v(\vec{b}_1, \ldots, \vec{b}_n)^\top \omega,$$

which are the outcome probabilities of local gbit measurements corresponding to the Bloch vectors $\vec{b}_1, \ldots, \vec{b}_n$ on the state ω. This mathematical property has many intuitively appealing consequences that are not otherwise guaranteed, e.g. the property that products of pure states are pure. It is also the reason why the mathematical literature has focused almost entirely on this notion of composite state space (cf. e.g. [47]): it leads to notions of "tensor products" of ordered linear spaces that allow one to prove general statements that are otherwise unavailable. In the context of this paper, it would seem extremely difficult to make any meaningful statements whatsoever if not even the linear space on which the global states live could be fixed from the outset.

We need one further ingredient to arrive at a model of computation, namely a set of reversible transformations. In analogy to standard quantum computation (where these are the unitaries), we postulate that **the transformations form a closed connected matrix group**, and thus Lie group, $\mathcal{G}$: they form a group since they can be composed; they must be linear maps since if we prepare a state ω with probability p and ω' with probability $(1 - p)$, they must act on the components of the convex combination $p\omega + (1 - p)\omega'$ individually, to be consistent with the probabilistic interpretation [11]. Moreover, it is physically meaningful to model the group as closed since whenever we can approximate a transformation to arbitrary accuracy by gates, it makes sense to declare this transformation as in principle implementable.

This postulate is almost, but not quite, satisfied by classical computation, i.e. the $d = 1$ case. As Bennett has shown [48], classical computation can be made fully reversible, at only marginal cost of space or time resources. There are finite universal gate sets (including e.g. Toffoli gates) that generate the full group of permutations of the 2^n configurations of the n bits. These permutations therefore constitute the reversible transformations of the classical bits, and they form a closed matrix group of linear maps. This group, however, is discrete and not connected.

This discreteness is already reflected in the fact that the one-dimensional "Bloch ball" is discrete, i.e. has only a finite number (two) of pure states. Since the set of classical configurations (pure states) of n bits is discrete, the group of reversible transformations must also be discrete. In the case $d \geq 2$ to which we thus restrict our attention in the following, however, even single bits (Bloch balls) contain a continuous manifold of

pure states. In order to allow every pure state to evolve into every other (which we would expect to be crucial for the exploitation of the full computational potential), it is therefore necessary that the reversible transformations form a continuous group $\mathcal{G}$ — in more detail, that $\mathcal{G}$ is a matrix Lie group such that its connected component at the identity is non-trivial. It then makes sense to consider continuous time evolution that implements elements of this connected component (as it is the case in quantum theory), and to disregard the mathematical possibility of having additional disconnected components. This motivates the assumption that $\mathcal{G}$ is connected.

All gates in a circuit will be elements of $\mathcal{G}$. This group must in particular contain the local qubit rotations: for $R \in \mathrm{SO}(d)$, write $\widehat{R}(1,\vec{a})^\top := (1, R\vec{a})^\top$, then the subgroup of local transformations is

$$\mathcal{G}_{\mathrm{loc}} := \{\widehat{R}_1 \otimes \widehat{R}_2 \otimes \ldots \otimes \widehat{R}_n \mid R_i \in \mathrm{SO}(d)\}.$$

Note that we have used tomographic locality in deriving this prescription: since a local transformation acts like a product of transformations on the product states, it must act like this on all other states too since they live on the vector space that is spanned by the product states. Tomographic locality hence enforces that we can represent any linear map $X : (\mathbb{R}^{(d+1)})^{\otimes n} \to (\mathbb{R}^{(d+1)})^{\otimes n}$ as a tensor with n upper and n lower indices; that is,

$$X^{\alpha_1\alpha_2\ldots\alpha_n}_{\beta_1\beta_2\ldots\beta_n} := (\vec{e}_{\beta_1} \otimes \ldots \otimes \vec{e}_{\beta_n})^\top X (\vec{e}_{\alpha_1} \otimes \ldots \otimes \vec{e}_{\alpha_n}),$$

where $0 \leq \alpha_i, \beta_i \leq d$, and $\vec{e}_\gamma$ denotes the γ-th unit vector, e.g. $\vec{e}_0 = (1, 0, \ldots, 0)^\top$. This is in contrast to Bloch vectors $\vec{b} \in \mathbb{R}^d$, where we use the notation $\mathbb{R}^d \ni \vec{b} = \vec{e}_1 = (1, 0, \ldots, 0)^\top$.

We demand that $\mathcal{G}_{\mathrm{loc}} \subseteq \mathcal{G}$, but do not make any further assumptions on $\mathcal{G}$. In particular, we do not assume that the n gbits play physically identical roles: our assumptions allow in principle composite state spaces of n gbits that are not symmetric with respect to permutations of the gbits. Hence we are also *not* assuming that gbits can be reversibly swapped, or that other natural choices of transformations such as extensions of classical reversible gates (like CNOT) can necessarily be implemented. Therefore, our framework does not rely on the same set of assumptions as the circuit framework of *symmetric monoidal categories* [49] that is often used in the quantum foundations context.

C. The trivial case $\mathcal{G} = \mathcal{G}_{\mathrm{loc}}$

For any Bloch ball dimension d, there is a trivial computational model: namely the choice that $\mathcal{G} = \mathcal{G}_{\mathrm{loc}}$. This describes a theory where the *only* possible reversible transformations are independent local transformations of the single gbits. Such a model does not even allow for classical gates like the CNOT; it only admits gates and computations that evolve the gbits independently from each other without ever correlating them, i.e. products of single-gbit gates. A state space that is compatible with this choice of global transformations is simply

$$\mathrm{conv}\left\{(1,\vec{a}_1)^\top \otimes \ldots \otimes (1,\vec{a}_n)^\top \mid \vec{a}_i \in B^d\right\},$$

i.e. all convex combinations of product states. This is a state space that does not contain entanglement.

D. $d = 3$ equals quantum computation, and relation to earlier work

For the case of the standard qubit, i.e. of $d = 3$, it has been proven in [50] that there is only a single possible non-trivial ($\mathcal{G}_{\mathrm{loc}} \subsetneq \mathcal{G}$) theory that satisfies the assumptions from above: namely, standard quantum theory over n qubits, with the $(2^n)\times(2^n)$ density matrices as the states, and the projective unitary group $\mathcal{G} = \mathrm{PU}(2^n)$ of transformations. That is, the postulates on composition of gbits from above, together with the structure of the single qubit, are sufficient to determine qubit quantum computation uniquely.

While this result is interesting in its own right, it is also the main motivation for the present work: if quantum computation is characterized by such a simple list of principles, then maybe one obtains other interesting models of computation by slightly tweaking one of the postulates. Since large parts of the mathematical structure are determined by the postulates on composition (no-signalling and tomographic locality), the most promising road towards modifying the setup and also keeping important mathematical tools seems to be to modify the structure of the single qubit — and technically as well as conceptually (as explained in Subsection II A), the most natural way to do this is by changing the dimension of the Bloch ball d.

In the special case of $n = 2$ gbits, the consequences of the above postulates have been explored in [51, 52]. There it has been proven that *the only consistent choice of transformations for Bloch ball dimension $d \neq 3$ is given by the trivial choice $\mathcal{G} = \mathcal{G}_{\mathrm{loc}}$*. However, computation is typically taking place on a large number $n \gg 2$ of gbits, and the techniques of [51, 52] cannot readily be generalized to $n > 2$.

In fact, it has been suggested in [23] that it is essential for Bloch ball dimensions $d \geq 4$ to allow for genuine m-partite interaction of the gbits, where $m \geq d - 1 \geq 3$. Without a conclusive proof or explicit construction of the state space, the authors conjectured that interesting multipartite reversible dynamics is possible for such systems. In contrast to quantum theory, this m-partite dynamics would not be decomposable into two-gbit interactions. While tomographic locality has not been assumed in [23], it is an important first step to check their conjecture under this additional assumption. In fact, it has been argued in [53] that in the context of spacetime

physics (the Bloch balls are interpreted in [23] as carrying some sort of d-dimensional spin degrees of freedom), tomographic locality is to be expected due to arguments from group representation theory.

This gives us another, independent motivation to ask the main question of this paper: *if $d \neq 3$ and n is any finite number of gbits, then what are the possible theories that satisfy the assumptions of Subsection II B?*

III. MAIN RESULT

The main result of this work is an answer to the question posed at the end of the previous section:

Theorem 1. *Consider a theory of n gbits, where single gbits are described by a $(d \geq 2)$-dimensional Bloch ball state space, subject to the single-gbit transformation group $\mathrm{SO}(d)$. As described above, let us assume no-signalling, tomographic locality, and that the global transformations form a closed connected matrix group $\mathcal{G}$.*

If $d \neq 3$, then necessarily $\mathcal{G} = \mathcal{G}_{\mathrm{loc}}$, i.e. the only possible gates are (independent combinations of) single-gbit gates. No transformation can correlate gbits that are initially uncorrelated; hence not even classical computation is possible.

We will now prove this result for the case $d \geq 4$. The proof in the $d = 2$ case uses similar techniques, but differs in several details for group-theoretic reasons. It will hence be deferred to the appendix.

As a first step, we will consider the generators of global transformations and show that there exists at least one that is of a certain normal form. This part of the proof is valid for all dimensions $d \geq 2$. A large part of this first step follows the construction in Ref. [50] and extends it to arbitrary dimensions.

A. Generator normal form for all dimensions $d \geq 2$

Let $G \in \mathcal{G}$ be a transformation of the composite system. Suppose we prepare n gbits initially in states with Bloch vectors $\vec{a}_1, \ldots, \vec{a}_n$, evolve the resulting product state via G, and perform a final local n-gbit measurement with Bloch vectors $\vec{b}_1, \ldots, \vec{b}_n$. The probability that the all the n outcomes on the n gbits are "yes" is

$$2^{-n} v(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n)^\top G v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) \in [0, 1].$$

Let us consider a group element $G = e^{\epsilon X}$ with $X \in \mathfrak{g}$ (the corresponding Lie algebra) and $\varepsilon \in \mathbb{R}$ and expand:

$$v(\vec{b}_1, \ldots, \vec{b}_n)^\top \left(\mathbf{1} + \epsilon X + \frac{\epsilon^2}{2} X^2 + \mathcal{O}(\epsilon^3)\right) v(\vec{a}_1, \ldots, \vec{a}_n) \in [0, 2^n].$$

From now on we restrict ourselves to unit length Bloch vectors, i.e. $|\vec{a}_i| = |\vec{b}_j| = 1$ for all i, j. We obtain

$$\mathcal{C}[\vec{a}_1] := v(-\vec{a}_1, \vec{b}_2, ..., \vec{b}_n)^\top X v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) = 0$$

since the zeroth order is zero which is a local minimum as a function of ϵ (see Figure 2 for further explanation). Thus the second order contribution has to be non-negative:

$$v(-\vec{a}_1, \vec{b}_2, \ldots, \vec{b}_n)^\top X^2 v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) \geq 0,$$

or more generally with the roles of qubits 1 and k exchanged,

$$v(\vec{b}_1, \ldots, \vec{b}_{k-1}, -\vec{a}_k, \vec{b}_{k+1}, \ldots \vec{b}_n)^\top X^2 v(\vec{a}_1, \ldots, \vec{a}_n) \geq 0. \tag{1}$$

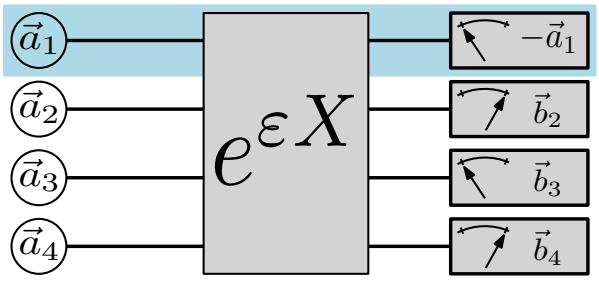


FIG. 2. We are using configurations like this one to derive constraints on the generators $X \in \mathfrak{g}$. In the special case $\varepsilon = 0$, the transformation $\exp(\varepsilon X)$ reduces to the identity. Hence, if we prepare the first wire in the (pure) state with Bloch vector $\vec{a}_1$, and perform a final measurement of that wire with Bloch vector $-\vec{a}_1$, the corresponding outcome will have probability zero, regardless of which local measurements we choose for the other wires. But probability zero is a local minimum, which implies that the derivative of this probability with respect to ε must be zero (yielding $C[\vec{a}_1] = 0$), and the second derivative must be non-negative (yielding constraint (1) in the case $k = 1$).

Other first and second order constraints are

$$v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)^\top X v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) = 0, \tag{2}$$
$$v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)^\top X^2 v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) \leq 0 \tag{3}$$

for analogous reasons as above (since $\vec{b}_j = \vec{a}_j$ for all j yields probability one for $\epsilon = 0$, which is the global and thus a local maximum). For fixed Bloch vectors $\vec{a}_2, \ldots, \vec{a}_n, \vec{b}_2, \ldots, \vec{b}_n$, define W^α_β as

$$\left[\vec{e}_\beta \otimes \begin{pmatrix}1\\ \vec{b}_2\end{pmatrix} \otimes \ldots \otimes \begin{pmatrix}1\\ \vec{b}_n\end{pmatrix}\right]^\top X \left[\vec{e}_\alpha \otimes \begin{pmatrix}1\\ \vec{a}_2\end{pmatrix} \otimes \ldots \otimes \begin{pmatrix}1\\ \vec{a}_n\end{pmatrix}\right]. \tag{4}$$

The equation $\mathcal{C}[\vec{e}_i] = 0$ implies $W^0_0 + W^i_0 - W^0_i - W^i_i = 0$, and $\mathcal{C}[-\vec{e}_i] = 0$ implies $W^0_0 - W^i_0 + W^0_i - W^i_i = 0$. Thus, $W^i_i = W^0_0$ and $W^i_0 = W^0_i$ for all $i \geq 1$. Since the vectors $\begin{pmatrix}1\\ \vec{a}\end{pmatrix}$ linearly span all of $\mathbb{R}^{d+1}$, we get

$$X^{i\ \alpha_2\ \ldots\ \alpha_n}_{i\ \beta_2\ \ldots\ \beta_n} = X^{0\ \alpha_2\ \ldots\ \alpha_n}_{0\ \beta_2\ \ldots\ \beta_n}, \tag{5}$$
$$X^{i\ \alpha_2\ \ldots\ \alpha_n}_{0\ \beta_2\ \ldots\ \beta_n} = X^{0\ \alpha_2\ \ldots\ \alpha_n}_{i\ \beta_2\ \ldots\ \beta_n} \tag{6}$$

for all $i \geq 1$ and all $\alpha_2, \ldots, \alpha_n, \beta_2, \ldots, \beta_n \geq 0$. Similarly, $\mathcal{C}[\frac{1}{\sqrt{2}}(\vec{e}_i + \vec{e}_j)] = 0$ for $i \neq j$, $i, j \geq 1$ yields

$$W_0^0 + \frac{1}{\sqrt{2}}W_0^i + \frac{1}{\sqrt{2}}W_0^j - \frac{1}{\sqrt{2}}W_i^0 - \frac{1}{2}W_i^i - \frac{1}{2}W_i^j - \frac{1}{\sqrt{2}}W_j^0 - \frac{1}{2}W_j^i - \frac{1}{2}W_j^j = 0.$$

Using the results on W_i^i and W_i^0 further above, this reduces to $-\frac{1}{2}W_i^j - \frac{1}{2}W_j^i = 0$, and thus

$$X_{j\ \beta_2\ \ldots\ \beta_n}^{i\ \alpha_2\ \ldots\ \alpha_n} = -X_{i\ \beta_2\ \ldots\ \beta_n}^{j\ \alpha_2\ \ldots\ \alpha_n} \tag{7}$$

for all $i, j \geq 1$ and $\alpha_2, \ldots, \alpha_n, \beta_2, \ldots, \beta_n \geq 0$. While we have derived (5), (6) and (7) for the first gbit, analogous equations hold for all other gbits with labels $2, \ldots, n$.

Let us denote by $\mathcal{A}$ the antisymmetric $(d+1) \times (d+1)$-matrices of the form

$$\mathcal{A} := \left\{ A_{\overline{A}} = \left(\begin{array}{c|c} 0 & \vec{0}^\top \\ \hline \vec{0} & \overline{A} \end{array}\right) \;\middle|\; \overline{A}^\top = -\overline{A} \right\},$$

and by $\mathcal{B}$ the symmetric $(d+1) \times (d+1)$-matrices of the form

$$\mathcal{B} := \left\{ B_{\vec{b}} = \left(\begin{array}{c|c} 0 & \vec{b}^\top \\ \hline \vec{b} & 0 \end{array}\right) \;\middle|\; \vec{b} \in \mathbb{R}^d \right\}.$$

Furthermore, let $\mathcal{I} := \mathbb{R} \cdot \mathbf{1}$, i.e. all multiples of the $(d+1) \times (d+1)$ identity matrix. The sets $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{I}$ are real linear matrix subspaces. Note that these three spaces are pairwise orthogonal with respect to the Hilbert-Schmidt inner product $\langle X, Y \rangle := \mathrm{tr}(X^\top Y)$. The matrix W defined in (4) must then be an element of $\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}$ due to the identities for its components that we have derived above. More generally, since the same identities hold for every index $i \in \{1, \ldots, n\}$ for the tensor X, we obtain $X \in (\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I})^{\otimes n}$. Since $X \in \mathfrak{g}$ was arbitrary, this tells us that

$$\mathfrak{g} \subset (\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I})^{\otimes n}.$$

The Lie algebra of the local transformations is

$$\mathfrak{g}_{\mathrm{loc}} = \mathcal{A}\otimes\mathbf{1}\otimes\ldots\otimes\mathbf{1} + \mathbf{1}\otimes\mathcal{A}\otimes\mathbf{1}\otimes\ldots\otimes\mathbf{1} + \ldots + \mathbf{1}\otimes\mathbf{1}\otimes\ldots\otimes\mathbf{1}\otimes\mathcal{A},$$

writing "+" instead of "$\oplus$" for readability. We can write the space $(\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I})^{\otimes n}$ in a somewhat different form. To this end, consider strings of symbols $x \in \{A, B, I\}^n$, for example $x = ABAI$ (if $n = 4$), and denote the corresponding tensor product matrix spaces by S_x; for this example, $S_x = \mathcal{A} \otimes \mathcal{B} \otimes \mathcal{A} \otimes \mathcal{I}$. Then $S_x \perp S_y$ for $x \neq y$ (with respect to the Hilbert-Schmidt inner product), and

$$(\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I})^{\otimes n} = \bigoplus_{x \in \{A,B,I\}^n} S_x.$$

Now let $X \in \mathfrak{g} \setminus \mathfrak{g}_{\mathrm{loc}}$ be an arbitrary generator which is not in the local Lie algebra (here we explicitly make the assumption that such an X exists). Since $X \neq 0$, there must exist x such that $\Phi_x(X) \neq 0$ for the orthogonal projection Φ_x into S_x, and since $X \notin \mathfrak{g}_{\mathrm{loc}}$, at least one of those x must satisfy

$$x \notin \{AI \ldots I, IAI \ldots I, \quad \ldots, \quad I \ldots IA\}.$$

Reordering the gbits, we may assume that $x = A^{n_A} B^{n_B} I^{n_I}$, where $n_A + n_B + n_I = n$ and one of the following three cases applies:

(i) $n_A = 0$,

(ii) $n_A = 1$ and $n_B \geq 1$,

(iii) $n_A \geq 2$.

Since S_x has an orthonormal basis of matrices of the form $A_{\overline{A}_1} \otimes \ldots \otimes A_{\overline{A}_{n_A}} \otimes B_1 \otimes \ldots \otimes B_{n_B} \otimes \mathbf{1}^{\otimes n_I}$, where all $A_{\overline{A}_i} \in \mathcal{A}$ and $B_i \in \mathcal{B}$, there must exist some matrix $\widetilde{M}_x$ of that form (i.e. $\widetilde{M}_x \in S_x$) such that $\langle X, \widetilde{M}_x \rangle \neq 0$. By moving constant scalar factors into the A-terms, we may assume that there are unit vectors $\vec{b}_i$ such that $B_i = B_{\vec{b}_i}$ for $i = 1, \ldots, n_B$. But since $\widehat{R} B_{\vec{b}} \widehat{R}^\top = B_{R\vec{b}}$ for all $R \in \mathrm{SO}(d)$, there are orthogonal matrices $\widehat{R}_i$ such that $R_i \vec{b}_i = \vec{e}_1 = (1, 0, \ldots, 0)^\top$ for all i, and the local transformation $T := \mathbf{1}^{\otimes n_A} \otimes \widehat{R}_1 \otimes \ldots \otimes \widehat{R}_{n_B} \otimes \mathbf{1}^{\otimes n_I}$ satisfies

$$\begin{aligned} M'_x &:= T\widetilde{M}_x T^{-1} = T\widetilde{M}_x T^\top \\ &= A_{\overline{A}_1} \otimes \ldots \otimes A_{\overline{A}_{n_A}} \otimes B^{\otimes n_B} \otimes \mathbf{1}^{\otimes n_I}, \end{aligned}$$

where $B := B_{\vec{e}_1}$. Set $X' := TXT^{-1}$, then since $T \in \mathcal{G}_{\mathrm{loc}} \subset \mathcal{G}$ and since the adjoint action of $\mathcal{G}_{\mathrm{loc}}$ preserves $\mathfrak{g}_{\mathrm{loc}}$, we have $X' \in \mathfrak{g} \setminus \mathfrak{g}_{\mathrm{loc}}$, and $\langle X', M'_x \rangle = \mathrm{tr}(TXT^{-1}T\widetilde{M}_x T^{-1}) = \langle X, \widetilde{M}_x \rangle \neq 0$. Similar argumentation allows us to bring the $A_{\overline{A}_i}$ into a standard form. Since the $d \times d$-matrices $\overline{A}_i$ are antisymmetric, one can infer from the results in [66] that there are orthogonal transformations $R_i \in \mathrm{SO}(d)$ such that

$$R_i \overline{A}_i R_i^\top = \begin{pmatrix} 0 & \lambda_1^{(i)} & & & & & \\ -\lambda_1^{(i)} & 0 & & & & & \\ & & 0 & \lambda_2^{(i)} & & & \\ & & -\lambda_2^{(i)} & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & \lambda_{d/2}^{(i)} \\ & & & & & -\lambda_{d/2}^{(i)} & 0 \end{pmatrix} \quad (d \text{ even}),$$

$$\begin{pmatrix} 0 & & & & & & & \\ & 0 & \lambda_1^{(i)} & & & & & \\ & -\lambda_1^{(i)} & 0 & & & & & \\ & & & 0 & \lambda_2^{(i)} & & & \\ & & & -\lambda_2^{(i)} & 0 & & & \\ & & & & & \ddots & & \\ & & & & & & 0 & \lambda_{\frac{d-1}{2}}^{(i)} \\ & & & & & & -\lambda_{\frac{d-1}{2}}^{(i)} & 0 \end{pmatrix} \quad (d \text{ odd}).$$

To save space, we will use the following notation in the remainder of the paper, where $\sigma = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$:

$$R_i \overline{A}_i R_i^\top = \begin{cases} \lambda_1^{(i)}\sigma \oplus \lambda_2^{(i)}\sigma \oplus \ldots \oplus \lambda_{d/2}^{(i)}\sigma & (d \text{ even}), \\ 0_{1\times 1} \oplus \lambda_1^{(i)}\sigma \oplus \lambda_2^{(i)}\sigma \oplus \ldots \oplus \lambda_{\frac{d-1}{2}}^{(i)}\sigma & (d \text{ odd}). \end{cases}$$

Now consider the corresponding $(d+1)\times(d+1)$-matrices $A_{R_i\overline{A}_iR_i^\top}$, for which we will introduce the following notation. By A_j, denote the matrix for which only the j-th block is non-zero, with $\lambda_j = 1$. That is, for even d, we have the $(d+1)\times(d+1)$-matrices

$$\begin{aligned} A_1 &= 0_{1\times 1} \oplus \sigma \oplus 0_{2\times 2} \oplus \ldots \oplus 0_{2\times 2}, \\ A_2 &= 0_{1\times 1} \oplus 0_{2\times 2} \oplus \sigma \oplus 0_{2\times 2} \oplus \ldots \oplus 0_{2\times 2}, \\ &\vdots \\ A_{d/2} &= 0_{1\times 1} \oplus 0_{2\times 2} \oplus \ldots \oplus 0_{2\times 2} \oplus \sigma, \end{aligned}$$

and for odd d, we have an extra initial zero, namely

$$\begin{aligned} A_1 &= 0_{2\times 2} \oplus \sigma \oplus 0_{2\times 2} \oplus \ldots \oplus 0_{2\times 2}, \\ A_2 &= 0_{2\times 2} \oplus 0_{2\times 2} \oplus \sigma \oplus 0_{2\times 2} \oplus \ldots \oplus 0_{2\times 2}, \\ &\vdots \\ A_{(d-1)/2} &= 0_{2\times 2} \oplus 0_{2\times 2} \oplus \ldots \oplus 0_{2\times 2} \oplus \sigma. \end{aligned}$$

The local transformation $\widetilde{T} := \widehat{R}_1 \otimes \ldots \otimes \widehat{R}_{n_A} \otimes \mathbf{1}^{\otimes n_B} \otimes \mathbf{1}^{\otimes n_I}$ satisfies

$$\begin{aligned} M_x := \widetilde{T} M_x' \widetilde{T}^{-1} = \widetilde{T} M_x' \widetilde{T}^\top \\ = \left(\sum_j \lambda_j^{(1)} A_j\right) \otimes \ldots \otimes \left(\sum_j \lambda_j^{(n_A)} A_j\right) \otimes B^{\otimes n_B} \otimes \mathbf{1}^{\otimes n_I} \quad (8) \end{aligned}$$

where the $\lambda_j^{(i)}$ are real numbers. Set $X'' := \widetilde{T} X' \widetilde{T}^{-1}$, then since $\widetilde{T} \in \mathcal{G}_{\text{loc}} \subset \mathcal{G}$, we have $X'' \in \mathfrak{g} \setminus \mathfrak{g}_{\text{loc}}$, and $\langle X'', M_x\rangle = \operatorname{tr}(\widetilde{T}X'\widetilde{T}^{-1}\widetilde{T}M_x'\widetilde{T}^{-1}) = \langle X', M_x'\rangle \neq 0$.

In summary, we have shown that if there exist any nonlocal generators at all, then there is one (denoted X'') that has non-zero overlap with a matrix $M_x \in S_x$ of the simple form (8).

Next we will show that this implies that $\mathfrak{g} = \mathfrak{g}_{\text{loc}}$ for all Bloch ball dimensions $d \geq 4$.

B. Proof of Theorem 1 for $d \geq 4$

We now use Schur's Lemma to construct orthogonal projectors (with respect to the Hilbert-Schmidt inner product) onto the subspaces of $\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}$. First, define

$$\Phi_I[M] := \int_{\mathrm{SO}(d)} \widehat{R} M \widehat{R}^{-1}\, dR \qquad (M \in \mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}),$$

then $\Phi_I[M] = 0$ for all $M \in \mathcal{A} \oplus \mathcal{B}$ and $\Phi_I[M] = M$ for all $M \in \mathcal{I}$. Since these subspaces are orthogonal with respect to the Hilbert-Schmidt inner product, Φ_I is the orthogonal projector onto the subspace $\mathcal{I}$ of $\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}$ (we are not interested in its action on matrices that are not in the space $\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}$).

For $j = 1, \ldots, d$, consider the stabilizer subgroup

$$\mathcal{G}_j := \{R \in \mathrm{SO}(d) \mid R\vec{e}_j = \vec{e}_j\},$$

where $\vec{e}_j$ denotes the jth standard unit vector in $\mathbb{R}^d$. Every $\mathcal{G}_j$ is isomorphic to $\mathrm{SO}(d-1)$ whose fundamental representation is irreducible (note that this is not true for $d = 3$; this causes the crucial difference to Ref. [50]). Set

$$\Phi_{\vec{e}_j}[M] := \int_{\mathcal{G}_j} \widehat{R} M \widehat{R}^{-1}\, dR \qquad (M \in \mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}),$$

then $\Phi_{\vec{e}_1}[M] = \int_{\mathrm{SO}(d-1)} \begin{pmatrix} \mathbf{1}_2 & \\ & S \end{pmatrix} M \begin{pmatrix} \mathbf{1}_2 & \\ & S^{-1} \end{pmatrix} dS,$ and, similarly as above, Schur's Lemma implies that $\Phi_{\vec{e}_1}$ is the orthogonal projector onto $\mathrm{span}(B) \oplus \mathcal{I}$. Hence $\Phi_B := \Phi_{\vec{e}_1} - \Phi_I$ is the orthogonal projector onto $\mathrm{span}(B)$.

Finally, we will construct the orthogonal projector onto $\mathcal{A}_{\text{blocks}} := \mathrm{span}\{A_1, \ldots, A_z\}$, where $z = d/2$ if d is even and $z = (d-1)/2$ if d is odd. To this end, define the SO(2)-matrix $R(\theta) := \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, and set

$$\widehat{R}(\theta_1, \theta_2, \ldots, \theta_z) := \begin{pmatrix} \mathbf{1}_y & & & \\ & R(\theta_1) & & \\ & & \ddots & \\ & & & R(\theta_z) \end{pmatrix},$$

where $y = 1$ if d is even and $y = 2$ if d is odd. Furthermore, define $\Phi'[M]$ as

$$\int_0^{2\pi} \frac{d\theta_1}{2\pi} \int_0^{2\pi} \frac{d\theta_2}{2\pi} \ldots \int_0^{2\pi} \frac{d\theta_z}{2\pi} \widehat{R}(\theta_1, \ldots, \theta_z) M \widehat{R}(\theta_1, \ldots, \theta_z)^{-1}.$$

Using the identities

$$\int_0^{2\pi} R(\theta)\frac{d\theta}{2\pi} = 0, \qquad \int_0^{2\pi} R(\theta)\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} R(-\theta)\frac{d\theta}{2\pi} = \frac{1}{2}\begin{pmatrix} m_{11}+m_{22} & m_{12}-m_{21} \\ -m_{12}+m_{21} & m_{11}+m_{22} \end{pmatrix} =: \Psi\left[\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}\right],$$

we can evaluate the action of Φ' as follows. First, any given $(d+1)\times(d+1)$-matrix M can be written in the block matrix form

$$M = \begin{pmatrix} M_{0,0} & \dots & M_{0,z} \\ \vdots & \ddots & \vdots \\ M_{z,0} & \dots & M_{z,z} \end{pmatrix}$$

where $M_{0,0}$ is a $y\times y$-matrix, all $M_{i,j}$ for $i,j\geq 1$ are 2×2-matrices, and the other matrices are $y\times 2$ and $2\times y$-matrices. Then, the action of Φ' becomes

$$\Phi'[M] = \begin{pmatrix} M_{0,0} & 0 & \dots & 0 \\ 0 & \Psi[M_{1,1}] & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \Psi[M_{z,z}] \end{pmatrix}.$$

Hence Φ' is an orthogonal projection that acts as the identity on $\mathcal{I}$ (i.e. $\Phi'(\mathbf{1})=\mathbf{1}$), and it projects $\mathcal{A}$ into its subspace $\mathcal{A}_{\text{blocks}}$. Furthermore, if d is even, then Φ' annihilates $\mathcal{B}$, and if d is odd, then Φ' projects $\mathcal{B}$ into its subspace $\text{span}(B)$. Thus, for d even, the orthogonal projector onto $\mathcal{A}_{\text{blocks}}$ is $\Phi_A := \Phi' - \Phi_I$, and for d odd, it is $\Phi_A := \Phi' - \Phi_I - \Phi_B$. Note that all these statements are only claimed to hold for the case that the maps are applied to operators in $\mathcal{A}\oplus\mathcal{B}\oplus\mathcal{I}$.

The projectors Φ_I, Φ_B and Φ_A map the Lie algebra $\mathfrak{g}$ into itself, if we apply different products of those projectors to the n sites. For example, consider the special case $n=1$. Then $Z\in\mathfrak{g}$ implies $\Phi_I[Z]\in\mathfrak{g}$ since $\mathfrak{g}$ is closed with respect to conjugations by elements of $\mathcal{G}$ and integrals. Similarly, $\Phi_{\vec{e}_1}[Z]\in\mathfrak{g}$, and since $\mathfrak{g}$ is a linear space, we also have $\Phi_B[Z] = \Phi_{\vec{e}_1}[Z] - \Phi_I[Z]\in\mathfrak{g}$, and similarly for the projector Φ_A. If $n\geq 2$, then we can successively apply the projectors to one of the sites, using the fact that tensoring local rotations with identities gives local transformations in $\mathcal{G}_{\text{loc}}$. Thus, if we define

$$\Phi := \Phi_A^{\otimes n_A}\otimes\Phi_B^{\otimes n_B}\otimes\Phi_I^{\otimes n_I},$$

then $Y := \Phi[X'']$ is another valid generator, $Y\in\mathfrak{g}$. Furthermore, $\Phi[M_x] = M_x$, hence

$$0\neq\langle X'', M_x\rangle = \langle X'', \Phi[M_x]\rangle = \langle\Phi[X''], M_x\rangle = \langle Y, M_x\rangle \tag{9}$$

and thus $Y\neq 0$ (we have used that Φ is an orthogonal projection and thus in particular self-adjoint with respect to the Hilbert-Schmidt inner product). In particular, $Y\in\text{Im}(\Phi) = \mathcal{A}_{\text{blocks}}^{\otimes n_A}\otimes\text{span}(B)^{\otimes n_B}\otimes\mathcal{I}^{\otimes n_I}$. Consequently, there are real numbers $\lambda_{j_1,\dots,j_{n_A}}$ such that

$$Y = \sum_{j_1,\dots,j_{n_A}=1}^{z}\lambda_{j_1,\dots,j_{n_A}} A_{j_1}\otimes\dots\otimes A_{j_{n_A}}\otimes B^{\otimes n_B}\otimes\mathbf{1}^{\otimes n_I}.$$

Now we apply the identities $A_jA_k = -\delta_{jk}P_j$ and $B^2 = P_B$, where

$$\begin{aligned} P_B &= \mathbf{1}_{2\times 2}\oplus 0_{(d-1)\times(d-1)}, \\ P_1 &= 0_{y\times y}\oplus\mathbf{1}_{2\times 2}\oplus 0_{2(z-1)\times 2(z-1)}, \\ P_2 &= 0_{y\times y}\oplus 0_{2\times 2}\oplus\mathbf{1}_{2\times 2}\oplus 0_{2(z-2)\times 2(z-2)} \end{aligned}$$

and so on, up to P_z. This gives us

$$Y^2 = (-1)^{n_A}\sum_{j_1,\dots,j_{n_A}}\lambda^2_{j_1,\dots,j_{n_A}} P_{j_1}\otimes\dots\otimes P_{j_{n_A}}\otimes P_B^{\otimes n_B}\otimes\mathbf{1}^{\otimes n_I}. \tag{10}$$

Suppose that n_A is even so that $(-1)^{n_A}=1$. We will now show that constraint (3) gets violated. To this end, fix some $j_1^{\mathbf{0}},\dots,j_{n_A}^{\mathbf{0}}$ such that $\lambda_{j_1^{\mathbf{0}},\dots,j_{n_A}^{\mathbf{0}}}\neq 0$. For $i=1,\dots,n_A$, choose some unit vector $\vec{a}_i\in\mathbb{R}^d$ such that $\begin{pmatrix}1\\ \vec{a}_i\end{pmatrix}^\top P_{j_i^{\mathbf{0}}}\begin{pmatrix}1\\ \vec{a}_i\end{pmatrix} > 0$; for all other j_i, we automatically get $\begin{pmatrix}1\\ \vec{a}_i\end{pmatrix}^\top P_{j_i}\begin{pmatrix}1\\ \vec{a}_i\end{pmatrix}\geq 0$. For $i=n_A+1,\dots,n_A+n_B$, set $\vec{a}_i := \vec{e}_1$, then $\begin{pmatrix}1\\ \vec{a}_i\end{pmatrix}^\top P_B\begin{pmatrix}1\\ \vec{a}_i\end{pmatrix} = 2$. Finally, for $i\geq n_A+n_B+1$, choose $\vec{a}_i$ arbitrarily such that $\begin{pmatrix}1\\ \vec{a}_i\end{pmatrix}^\top\mathbf{1}\begin{pmatrix}1\\ \vec{a}_i\end{pmatrix} = 2$. Altogether, we obtain

$$v(\vec{a}_1,\dots,\vec{a}_n)^\top Y^2\, v(\vec{a}_1,\dots,\vec{a}_n) > 0$$

which violates constraint (3). Thus n_A must be odd, and $(-1)^{n_A}=-1$.

Recall constraint (1) in the special case $k=2$:

$$v(\vec{b}_1,-\vec{a}_2,\vec{b}_3,\dots,\vec{b}_n)^\top Y^2\, v(\vec{a}_1,\vec{a}_2,\dots,\vec{a}_n)\geq 0 \tag{11}$$

for all unit vectors $\vec{a}_i,\vec{b}_j\in\mathbb{R}^d$. For all $i\in[n_A+n_B+1,n]\setminus\{2\}$, choose $\vec{a}_i,\vec{b}_i$ such that $\begin{pmatrix}1\\ \vec{b}_i\end{pmatrix}^\top\mathbf{1}\begin{pmatrix}1\\ \vec{a}_i\end{pmatrix} > 0$ (simply avoid the choice $\vec{a}_i=-\vec{b}_i$). Similarly, for all $i\in[n_A+1,n_A+n_B]\setminus\{2\}$, choose $\vec{a}_i,\vec{b}_i$ such that $\begin{pmatrix}1\\ \vec{b}_i\end{pmatrix}^\top P_B\begin{pmatrix}1\\ \vec{a}_i\end{pmatrix} > 0$. We will now distinguish two cases for n_A.

First, consider the case $n_A=1$. Since our original generator X was chosen nonlocal, it follows that $n_B\geq 1$, as explained in Subsection III A. Thus, the second tensor factor in (10) must be P_B. We will now choose $\vec{a}_2=\vec{e}_2$ which implies that $\begin{pmatrix}1\\ -\vec{a}_2\end{pmatrix}^\top P_B\begin{pmatrix}1\\ \vec{a}_2\end{pmatrix} = 1$. But then we may still choose $\vec{b}_1,\vec{a}_1$ arbitrarily, and by choosing these two unit vectors suitably from the subspace $\text{Im}(P_{j_1})$, we may generate an arbitrary sign for $\begin{pmatrix}1\\ \vec{b}_1\end{pmatrix}^\top P_{j_1}\begin{pmatrix}1\\ \vec{a}_1\end{pmatrix}$. Thus, we can break constraint (11) by a suitable choice of these two unit vectors, which yields a contradiction.

Second, suppose that $n_A\geq 3$ (we already know that n_A must be odd). Then we can choose $\vec{a}_2$ such that $\begin{pmatrix}1\\ -\vec{a}_2\end{pmatrix}^\top P_{j_2}\begin{pmatrix}1\\ \vec{a}_2\end{pmatrix} = -1$. We have even more freedom

than in the previous case: for all $i \in [1, n_A] \setminus \{2\}$, we can choose $\vec{b}_i, \vec{a}_i$ from the subspace $\mathrm{Im}(P_{i_j})$ such that we get an arbitrary sign for every $\begin{pmatrix} 1 \\ \vec{b}_i \end{pmatrix}^\top P_{j_i} \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix}$. This also leads to a violation of constraint (11), and we obtain a contradiction as well.

This means that our initial assumption must have been wrong — namely, that there exists a generator in $\mathfrak{g} \setminus \mathfrak{g}_{\mathrm{loc}}$. We conclude that instead this set must be empty, hence $\mathfrak{g} = \mathfrak{g}_{\mathrm{loc}}$. But since $\mathcal{G}$ is compact and connected, it follows from [54, Theorem VII.2.2 (v)] that $\mathcal{G}$ cannot be larger than $\mathcal{G}_{\mathrm{loc}}$. This proves our main result, Theorem 1, for Bloch ball dimensions $d \geq 4$. The proof for $d = 2$ is given in Appendix A.

IV. CONCLUSIONS

Given a few simple properties that turn out to characterize qubit quantum computation, we have considered a natural modification: allowing the single bits to have more or less than the qubit's $d = 3$ degrees of freedom. We have analyzed the set of possible reversible transformations in the resulting theories, under the conjecture [23] (and in hopes) of discovering novel computational models that differ in interesting ways from quantum computation. Unfortunately, it turns out that the resulting models do not allow for any non-trivial reversible gates whatsoever. This reinforces earlier intuition [24] that quantum theory, or in this context quantum computation, is an "island in theoryspace".

While we have made an effort to be as careful and parsimonious in our assumptions as possible, it is still interesting to ask whether there are any remaining "loopholes" that could in principle leave some wiggle room for non-trivial beyond-quantum computation: can any of the assumptions of Subsection II B be dropped or weakened, while insisting that single bits are described by Bloch balls? We discuss several options in Appendix B; in short, the most promising (but difficult) approaches would be to drop tomographic locality, and/or to drop reversibility or continuity of transformations. Both options present formidable mathematical challenges and are therefore deferred to future work.

The "rigidity" of quantum theory, i.e. the difficulty of modifying it in consistent ways, has been recognized in different contexts for a long time, see e.g. Weinberg's proposal of a nonlinear modification of quantum mechanics [55], and Gisin's subsequent discovery [56] that this modification allows for superluminal signalling. The research presented in this paper and in other work (like [57, 58]) makes this intuition more rigorous by specifying which combinations of principles already enforce the familiar behavior of quantum theory. These insights also illuminate our understanding of quantum computation, since they tell us which physical principles enforce its properties, and/or which other theoretical models of computation are plausibly conceivable.

Finally, it is interesting to speculate that the result of this paper is indirectly related to spacetime physics. After all, it is the fact that a qubit is represented as a 3-ball B^3, with $\mathrm{SO}(3)$ as its transformation group, which allows for spin-$1/2$ particles that couple to rotations in three-dimensional space. Given the popularity of approaches in which spacetime emerges in some way from an underlying quantum theory [59–61], this observation can perhaps be regarded as more than a coincidence. In fact, it has been argued more rigorously that the structures of quantum theory and spacetime mutually constrain each other [23, 53, 62–64]. This suggests a slogan that also fits some other ideas from quantum information [65]: the limits of computation are the limits of our world.

Acknowledgments. We thank Časlav Brukner, Borivoje Dakić, Philip Goyal, Lluís Masanes, and Jochen Rau for discussions. M.K. acknowledges support from the Austrian Science Fund (FWF) through the Doctoral Programme CoQuS (W1210). This publication was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

[1] P. W. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, SIAM J. Sci. Statist. Comput. **26**, 1484 (1997).
[2] S. Aaronson, *Quantum computing since Democritus*, Cambridge University Press, New York, 2013.
[3] D. Gottesman, talk at the International Conference on Group Theoretic Methods in Physics (1998); see also arXiv:quant-ph/9807006.
[4] S. Aaronson and D. Gottesman, *Improved simulation of stabilizer circuits*, Phys. Rev. A **70**, 052328 (2004).
[5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.
[6] M. Howard, J. Wallman, V. Veitch and J. Emerson, *Contextuality supplies the 'magic' for quantum computation*, Nature **510**, 351–355 (2014).
[7] D. Gottesman, *An Introduction to Quantum Error Correction and Fault-Tolerant Quantum Computation*, arXiv:0904.2557.
[8] D. S. Abrams and S. Lloyd, *Nonlinear Quantum Mechan-*

ics Implies Polynomial-Time Solution for NP-Complete and Sharp-P Problems, Phys. Rev. Lett. **81**, 3992 (1998).

[9] M. Sipser, *Introduction to the Theory of Computation*, PWS Publishing, 1997.

[10] S. Aaronson and J. Watrous, *Closed timelike curves make quantum and classical computing equivalent*, Proc. R. Soc. A **465**, 631–647 (2009).

[11] J. Barrett, *Information processing in generalized probabilistic theories*, Phys. Rev. A **75**, 032304 (2007).

[12] D. Gross, M. Müller, R. Colbeck, and O. C. O. Dahlsten, *All Reversible Dynamics in Maximally Nonlocal Theories are Trivial*, Phys. Rev. Lett. **104**, 080402 (2010).

[13] S. W. Al-Safi and A. J. Short, *Reversible dynamics in strongly non-local Boxworld systems*, J. Phys. A: Math. Theor. **47**, 325303 (2014).

[14] J. G. Richens, J. H. Selby, and S. W. Al-Safi, *Entanglement is Necessary for Emergent Classicality in All Physical Theories*, Phys. Rev. Lett. **119**, 080503 (2017).

[15] C. M. Lee and J. Barrett, *Computation in generalised probabilistic theories*, New J. Phys. **17**, 083001 (2015).

[16] J. Barrett, N. de Beaudrap, M. J. Hoban, and C. M. Lee, *The computational landscape of general physical theories*, arXiv:1702.08483.

[17] C. M. Lee and J. H. Selby, *Deriving Grover's lower bound from simple physical principles*, New J. Phys. **18**, 093047 (2016).

[18] H. Barnum, C. M. Lee, and J. H. Selby, *Oracles and query lower bounds in generalised probabilistic theories*, Found. Phys. **48**, 954 (2018).

[19] A. J. P. Garner, *Interferometric computation beyond quantum theory*, Found. Phys. **48**, 886 (2018).

[20] C. M. Lee and J. H. Selby, *Generalised phase kick-back: the structure of computational algorithms from physical principles*, New J. Phys. **18**, 033023 (2016).

[21] C. M. Lee and M. J. Hoban, *Bounds on the power of proofs and advice in general physical theories*, Proc. R. Soc. A **472** (2190), 20160076 (2016).

[22] P. Jordan, J. von Neumann, and E. Wigner, *On an Algebraic Generalization of the Quantum Mechanical Formalism*, Annals of Mathematics **35**(1), 29–64 (1934).

[23] B. Dakić and Č. Brukner, *The classical limit of a physical theory and the dimensionality of space*, in "Quantum Theory: Informational Foundations and Foils", eds. G. Chiribella and R. W. Spekkens, Springer, 2016.

[24] S. Aaronson, *Is Quantum Mechanics An Island In Theoryspace?*, arXiv:quant-ph/0401062.

[25] L. Hardy, *Quantum Theory From Five Reasonable Axioms*, arXiv:quant-ph/0101012.

[26] B. Dakić and Č. Brukner, *Quantum theory and beyond: is entanglement special?*, in "Deep Beauty: Understanding the Quantum World through Mathematical Innovation", ed. H. Halvorson, Cambridge University Press, Cambridge, 2011.

[27] Ll. Masanes and M. P. Müller, *A derivation of quantum theory from physical requirements*, New J. Phys. **13**, 063001 (2011).

[28] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Informational derivation of quantum theory*, Phys. Rev. A **84**, 012311 (2011).

[29] L. Hardy, *Reformulating and Reconstructing Quantum Theory*, arXiv:1104.2066.

[30] P. A. Höhn, *Toolbox for reconstructing quantum theory from rules on information acquisition*, Quantum **1**, 38 (2017).

[31] P. A. Höhn and A. Wever, *Quantum theory from questions*, Phys. Rev. A **95**, 012102 (2017).

[32] P. A. Höhn, *Quantum theory from rules on information acquisition*, Entropy **19**(3), 98 (2017).

[33] P. Goyal, *Derivation of quantum theory from Feynman's rules*, Phys. Rev. A **89**, 032120 (2014).

[34] A. Wilce, *A Royal Road to Quantum Theory (or Therebouts)*, arXiv:1606.09306.

[35] J. H. Selby, C. M. Scandolo, and B. Coecke, *Reconstructing quantum theory from diagrammatic postulates*, arXiv:1802.00367.

[36] J. van de Wetering, *Reconstruction of quantum theory from universal filters*, arXiv:1801.05798.

[37] L. Randall, *Warped Passages: Unraveling the Mysteries of the Universe's Hidden Dimensions*, (Harper Perennial, 2006).

[38] K. Agashe and A. Pomarol, *Focus on Extra Space Dimensions*, New J. Phys. **12**, 075010 (2010).

[39] T. Kaluza, *Zum Unitätsproblem der Physik*, Akad. Wiss. Berlin (Math. Phys.), 966—972 (1921).

[40] O. Klein, *Quantentheorie und fünfdimensionale Relativitätstheorie*, Zeitschrift für Physik A **37**(12), 895—906 (1926).

[41] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, *The Hierarchy problem and new dimensions at a millimeter*, Phys. Lett. B **429**, 263—272 (1998).

[42] I. Antoniadis, *A possible new dimension at a few TeV*, Phys. Lett. B **246**, 377-–384 (1990).

[43] M. Pawlowski and A. Winter, *Hyperbits: the information quasiparticles*, Phys. Rev. A **85**, 022331 (2012).

[44] H. Araki, *On a Characterization of the State Space of Quantum Mechanics*, Commun. Math. Phys. **75**, 1–24 (1980).

[45] H. Barnum, M. A. Graydon, and A. Wilce, *Some Nearly Quantum Theories*, EPTCS **195**, 59–70 (2015).

[46] H. Barnum, M. A. Graydon, and A. Wilce, *Composites and Categories of Euclidean Jordan Algebras*, arXiv:1606.09331.

[47] I. Namioka and R. R. Phelps, *Tensor products of compact convex sets*, Pac. J. Math. **31**(2), 469–480 (1969).

[48] C. H. Bennett, *Logical reversibility of computation*, IBM Journal of Research and Development **17**(6), 525–532 (1973).

[49] B. Coecke and A. Kissinger, *Picturing Quantum Processes – A First Course in Quantum Theory and Diagrammatic Reasoning*, Cambridge University Press, Cambridge, 2017.

[50] G. de la Torre, Ll. Masanes, A. J. Short, and M. P. Müller, *Deriving quantum theory from its local structure and reversibility*, Phys. Rev. Lett. **109**, 090403 (2012).

[51] Ll. Masanes, M. P. Müller, R. Augusiak, and D. Pérez-García, *Existence of an information unit as a postulate of quantum theory*, Proc. Natl. Acad. Sci. USA **110**(41), 16373 (2013).

[52] Ll. Masanes, M. P. Müller, R. Augusiak, and D. Pérez-García, *Entanglement and the three-dimensionality of the Bloch ball*, J. Math. Phys. **55**, 122203 (2014).

[53] M. P. Müller and Ll. Masanes, *Three-dimensionality of space and the quantum bit: an information-theoretic approach*, New J. Phys. **15**, 053040 (2013).

[54] B. Simon, *Representations of Finite and Compact Groups*, American Mathematical Society, 1996.

[55] S. Weinberg, *Testing Quantum Mechanics*, Ann. Phys. **194**, 336–386 (1989).

[56] N. Gisin, *Weinberg's nonlinear quantum mechanics and supraluminal communications*, Phys. Lett. A **143**, 1–2 (1990).

[57] T. D. Galley and Ll. Masanes, *Classification of all alternatives to the Born rule in terms of informational properties*, Quantum **1**, 15 (2017).

[58] T. D. Galley and Ll. Masanes, *Impossibility of mixed-state purification in any alternative to the Born rule*, arXiv:1801.06414.
[59] R. Jacobson, *Thermodynamics of Spacetime: the Einstein Equation of State*, Phys. Rev. Lett. **75**, 1260 (1995).
[60] S. Ryu and T. Takayanagi, *Holographic Derivation of Entanglement Entropy from the anti–de Sitter Space/Conformal Field Theory Correspondence*, Phys. Rev. Lett. **96**, 181602 (2006).
[61] P. A. Höhn, *Reflections on the information paradigm in quantum and gravitational physics*, J. Phys.: Conf. Series **880**(1) (2017).
[62] P. A. Höhn and M. P. Müller, *An operational approach to spacetime symmetries: Lorentz transformations from quantum communication*, New J. Phys. **18**, 063026 (2016).
[63] M. P. Müller, S. Carrozza, and P. Höhn, *Is the local linearity of space-time inherited from the linearity of probabilities?*, J. Phys. A: Math. Theor. **50**, 054003 (2017).
[64] A. J. P. Garner, M. P. Müller, and O. C. O. Dahlsten, *The complex and quaternionic quantum bit from relativity of simultaneity on an interferometer*, Proc. R. Soc. A **473**, 20170596 (2017).
[65] S. Aaronson, *NP-complete problems and physical reality*, ACM SIGACT News **36**(1), 30–52 (2005).
[66] G. Thompson, *Normal Forms for Skew-Symmetric Matrices and Hamiltonian Systems with First Integrals Linear in Momenta*, Proc. Am. Math. Soc. **104**(3), 910–916 (1988).
[67] M. A. Graydon, *Quaternions and quantum theory*, Master's thesis, University of Waterloo, Waterloo, Canada, 2011.
[68] G. Chiribella, G, M. D'Ariano, and P. Perinotti, *Probabilistic theories with purification*, Phys. Rev. A **81**, 062348 (2010).
[69] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Quantum theory, namely the pure and reversible theory of information*, Entropy **14**(10), 1877–1893 (2012).
[70] H. Barnum, C. M. Lee, C. M. Scandolo, and J. H. Selby, *Ruling out higher-order interference from purity principles*, Entropy **19**(6), 253 (2017).

Appendix A: Proof of Theorem 1 for $d = 2$

Due to different group-theoretic properties, we now have less freedom to construct projectors by integrating over conjugations with local transformations. A first difference to the case $d \geq 4$ appears already in

$$\Phi_{AI}[M] := \int_{\mathrm{SO}(2)} \widehat{R} M \widehat{R}^{-1}\, dR \qquad (M \in \mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}).$$

It turns out that this map leaves not only $\mathcal{I}$ but also $\mathcal{A}$ (which is now one-dimensional) invariant. Since it still annihilates $\mathcal{B}$, it is the orthogonal projector onto $\mathcal{A} \oplus \mathcal{I}$. We can still use $\Phi_B := \mathbf{1} - \Phi_{AI}$ as the projector onto $\mathcal{B}$, but we cannot construct a projector onto $\mathrm{span}(B)$ in a similar way. Now set $n_{AI} := n_A + n_I$, and reorder the gbits such that A comes first, and then I, and then B (in contrast to the previous subsections). Next define the orthogonal projector

$$\Phi := \Phi_{AI}^{\otimes n_{AI}} \otimes \Phi_B^{\otimes n_B},$$

then $Y := \Phi[X'']$ is another valid generator, i.e. $Y \in \mathfrak{g}$, and $Y \in (\mathcal{A} \oplus \mathcal{I})^{\otimes n_{AI}} \otimes \mathcal{B}^{\otimes n_B}$. Since $\Phi[M_x] = M_x$, the calculation (9) proves that $Y \neq 0$. It also follows that $Y \in \mathfrak{g} \setminus \mathfrak{g}_{\mathrm{loc}}$ since Y has non-zero overlap with M_x which in turn is orthogonal onto $\mathfrak{g}_{\mathrm{loc}}$. Defining $A^{(0)} := \mathbf{1}$ and $A^{(1)} := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$ (which spans the one-dimensional space $\mathcal{A}$), $B_0 := \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $B_1 := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, the generator Y can be written in the form

$$Y = \sum_{k_1,\ldots,k_n=0}^{1} \alpha_{k_1,\ldots,k_n} A^{(k_1)} \otimes A^{(k_2)} \otimes \ldots \otimes A^{(k_{n_{AI}})} \otimes B_{k_m} \otimes \ldots \otimes B_{k_n}, \tag{A1}$$

where the $\alpha_{k_1,\ldots,k_n}$ are real numbers and $m := n_{AI} + 1$.

Now we will apply the first-order constraint (2) for some special choice of unit vectors $\vec{a}_i$. First, fix $j_1, j_2, \ldots, j_n \in \{0, 1\}$ arbitrarily. For $i \leq n_{AI}$ set $\vec{a}_i := \vec{e}_1$, and for $i \geq m$ set

$$\vec{a}_i := \begin{cases} \vec{e}_1 & \text{if } j_i = 0 \\ \vec{e}_2 & \text{if } j_i = 1. \end{cases}$$

We obtain the following two equations

$$\begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix}^{\top} A^{(k_i)} \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix} = 2\delta_{k_i,0} \quad (i = 1, \ldots, n_{AI}), \qquad \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix}^{\top} B_{k_i} \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix} = 2\delta_{j_i,k_i} \quad (i = m, \ldots, n),$$

and substituting them into constraint (2) yields

$$0 = v(\vec{a}_1,\dots,\vec{a}_n)^\top Y v(\vec{a}_1,\dots,\vec{a}_n) = 2^{n_{AI}} \sum_{k_m,\dots,k_n=0}^{1} \alpha_{0,\dots,0,k_m,\dots,k_n} \prod_{\ell=m}^{n} \begin{pmatrix}1\\ \vec{a}_\ell\end{pmatrix}^\top B_{k_\ell} \begin{pmatrix}1\\ \vec{a}_\ell\end{pmatrix} = 2^n \alpha_{0,\dots,0,j_m,\dots,j_n}.$$

Thus $\alpha_{0,\dots,0,j_m,\dots,j_n} = 0$, i.e. every non-vanishing summand in (A1) contains at least one $A^{(1)}$-term. Furthermore, in the special case that $n_B = 0$, all summands with a single $A^{(1)}$-term are themselves elements of $\mathfrak{g}_{\text{loc}}$, and by subtracting those elements, we obtain another non-zero generator (which now also call Y) for which every non-vanishing summand has at least *two* $A^{(1)}$-terms.

Next we slightly generalize constraint (1):

Lemma 2. *The constraint*

$$v(\vec{b}_1,\vec{b}_2,\dots,\vec{b}_{k-1},-\vec{a}_k,\vec{b}_{k+1},\dots,\vec{b}_n)^\top X^2 v(\vec{a}_1,\vec{a}_2,\dots,\vec{a}_n) \geq 0 \tag{A2}$$

also holds if we replace one or more of the unit vectors $\vec{b}_j, \vec{a}_j$, but not $\pm\vec{a}_k$, by the zero vector.

Proof. We start with constraint (1), where all vectors are assumed to be unit vectors. To replace, for example, $\vec{b}_j$ (for $j \neq k$) by $\vec{0}$, consider (1) and its version with $\vec{b}_j$ replaced by $-\vec{b}_j$. Adding up the two inequalities (and dividing the result by two) proves (A2) for $\vec{b}_j = 0$. We can similarly replace any of the $\vec{a}_j$ (for $j \neq k$) by $\vec{0}$, and do so recursively. □

Now we are ready to state and prove the main result of the appendix:

Lemma 3. *If $d = 2$ then $\mathcal{G} = \mathcal{G}_{\text{loc}}$, i.e. the only reversible transformations are the local transformations.*

Proof. Our strategy is to prove the following claim:

Claim: Let $0 \leq \ell \leq n_{AI}$ be an integer. Then Y does not contain any summand in (A1) which has exactly ℓ occurrences of $A^{(0)}$. In more formal words, if $j_1,\dots,j_n$ has the property that $\#\{i \in [1,n_{AI}] \mid j_i = 0\} = \ell$ then $\alpha_{j_1,\dots,j_n} = 0$.

This claim will then imply that $Y = 0$, which is a contradiction (we have shown further above that $Y \neq 0$).

We will prove this claim for two different cases separately; in both cases, our proof will be by induction. Note that we have already shown the claim above for $\ell = n_{AI}$ (since there must be at least one $A^{(1)}$-term in every summand).

Case 1: $n_B = 0$ (such that $n_{AI} = n$).

Induction start: We know the claim is true for $\ell = n$. Furthermore, since $n_B = 0$, we have constructed Y such that no summand contains exactly one $A^{(1)}$-term, hence the claim is also true for $\ell = n-1$.

Induction hypothesis: Consider an arbitrary integer ℓ with $0 \leq \ell \leq n-2$. Let us assume that for any integer ℓ' with $0 \leq \ell' \leq n$ and $\ell' > \ell$ we know that Y contains no summand with exactly ℓ' occurrences of $A^{(0)}$.

Induction step: Using the induction hypothesis, we will now show that the Claim also holds for ℓ itself.

We do so by contradiction. Suppose there was at least one non-vanishing summand in Y with exactly ℓ occurrences of $A^{(0)}$. That is, there exist $j_1^{\mathbf{0}},\dots,j_n^{\mathbf{0}}$ such that $\alpha_{j_1^{\mathbf{0}},\dots,j_n^{\mathbf{0}}} \neq 0$ and exactly ℓ of the $j_i^{\mathbf{0}}$ are equal to zero. We will apply constraint (A2) for some choice of vectors $\vec{a}_i, \vec{b}_i$. To this end, for every i with $j_i^{\mathbf{0}} = 0$ set $\vec{a}_i := \vec{0}$. For those i, it follows that $A^{(j_i)}A^{(k_i)} \begin{pmatrix}1\\ \vec{a}_i\end{pmatrix} = \delta_{j_i,0}\delta_{k_i,0} \begin{pmatrix}1\\0\\0\end{pmatrix}$. Now Y^2 is of the form

$$Y^2 = \sum_{j_1,\dots,j_n=0}^{1} \sum_{k_1,\dots,k_n=0}^{1} \alpha_{j_1,\dots,j_n}\alpha_{k_1,\dots,k_n} (A^{(j_1)}A^{(k_1)}) \otimes \dots \otimes (A^{(j_n)}A^{(k_n)}).$$

Now consider $w := Y^2 v(\vec{a}_1,\dots,\vec{a}_n)$. If a summand of Y^2 has less than ℓ indices k_i with $k_i = 0$ then it does not contribute to w; also, there are no summands with more than ℓ indices k_i with $k_i = 0$. Among those summands with exactly ℓ indices k_i with $k_i = 0$, these indices must occur in exactly those places i where $j_i^{\mathbf{0}} = 0$, otherwise those summands do not contribute to w. But this enforces that only the summand with $(k_1,\dots,k_n) = (j_1,\dots,j_n) = (j_1^{\mathbf{0}},\dots,j_n^{\mathbf{0}})$ contributes to w, and we get

$$Y^2 v(\vec{a}_1,\dots,\vec{a}_n) = \alpha^2_{j_1^{\mathbf{0}},\dots,j_n^{\mathbf{0}}} \bigotimes_{z=1}^{n} \left(A^{(j_z^{\mathbf{0}})}\right)^2 \begin{pmatrix}1\\ \vec{a}_z\end{pmatrix}.$$

There are at least two indices z with $j_z^{\mathbf{0}} = 1$; let k be one of those indices, and define $\vec{a}_k := \vec{e}_1$. Then $\begin{pmatrix} 1 \\ -\vec{a}_k \end{pmatrix}^\top \left(A^{(j_k^{\mathbf{0}})}\right)^2 \begin{pmatrix} 1 \\ \vec{a}_k \end{pmatrix} = 1$. Among the remaining places z with $j_z^{\mathbf{0}} = 1$, we can choose $\vec{a}_z$ and $\vec{b}_z$ such that $\begin{pmatrix} 1 \\ \vec{b}_z \end{pmatrix}^\top \left(A^{(j_z^{\mathbf{0}})}\right)^2 \begin{pmatrix} 1 \\ \vec{a}_z \end{pmatrix}$ takes any sign we like. This will allow is to violate constraint (A2), and we have a contradiction.

Case 2: $n_B \geq 1$.

Induction start: We have already shown the claim for $\ell = n_{AI}$.

Induction hypothesis: Consider an arbitrary integer ℓ with $0 \leq \ell \leq n_{AI} - 1$. Let us assume that for any integer ℓ' with $0 \leq \ell' \leq n_{AI}$ and $\ell' > \ell$ we know that Y contains no summand with exactly ℓ' occurrences of $A^{(0)}$.

Induction step: We proceed similarly as in Case 1. Using the induction hypothesis, we will now show that the Claim also holds for ℓ itself.

We do so by contradiction. Suppose there was at least one non-vanishing summand in Y with exactly ℓ occurrences of $A^{(0)}$. That is, there exist $j_1^{\mathbf{0}}, \ldots, j_n^{\mathbf{0}}$ such that $\alpha_{j_1^{\mathbf{0}},\ldots,j_n^{\mathbf{0}}} \neq 0$ and exactly ℓ of the $j_i^{\mathbf{0}}$ among $i \in [1, n_{AI}]$ are equal to zero. We will apply constraint (A2) for some choice of vectors $\vec{a}_i, \vec{b}_i$. To this end, for every i with $j_i^{\mathbf{0}} = 0$ set $\vec{a}_i := \vec{0}$ and choose $\vec{b}_i$ arbitrarily. For those i, it follows that

$$\begin{pmatrix} 1 \\ \vec{b}_i \end{pmatrix}^\top A^{(j_i)} A^{(k_i)} \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix} = \delta_{j_i,0}\delta_{k_i,0}. \tag{A3}$$

In Case 2, Y^2 is of the form

$$Y^2 = \sum_{j_1,\ldots,j_n=0}^{1} \sum_{k_1,\ldots,k_n=0}^{1} \alpha_{j_1,\ldots,j_n}\alpha_{k_1,\ldots,k_n} (A^{(j_1)}A^{(k_1)}) \otimes \ldots \otimes (A^{(j_{n_{AI}})}A^{(k_{n_{AI}})}) \otimes (B_{j_m}B_{k_m}) \otimes \ldots \otimes (B_{j_n}B_{k_n}).$$

Again, we have to choose which place corresponds to the k in constraint (A2). This time, we will choose $k = m$, and set $\vec{a}_k = \vec{e}_1$ if $j_k^{\mathbf{0}} = 1$ resp. $\vec{a}_k = \vec{e}_2$ if $j_k^{\mathbf{0}} = 0$, which implies

$$\begin{pmatrix} 1 \\ -\vec{a}_k \end{pmatrix}^\top B_{j_k}B_{k_k} \begin{pmatrix} 1 \\ \vec{a}_k \end{pmatrix} = \delta_{j_k,j_k^{\mathbf{0}}}\delta_{k_k,j_k^{\mathbf{0}}}.$$

For all other $i \in [m,n] \setminus \{k\}$ we make the following choice. If $j_i^{\mathbf{0}} = 1$ we set $\vec{b}_i = -\vec{e}_1$ and $\vec{a}_i = \vec{e}_1$, and if $j_i^{\mathbf{0}} = 0$ we set $\vec{b}_i = -\vec{e}_2$ and $\vec{a}_i = \vec{e}_2$. This enforces

$$\begin{pmatrix} 1 \\ \vec{b}_i \end{pmatrix}^\top B_{j_i}B_{k_i} \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix} = \delta_{j_i,j_i^{\mathbf{0}}}\delta_{k_i,j_i^{\mathbf{0}}} \qquad (i \in [m,n] \setminus \{k\}).$$

Regardless of how we choose the remaining $\vec{a}_i$, we obtain

$$\begin{aligned} &v(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_{k-1}, -\vec{a}_k, \vec{b}_{k+1}, \ldots, \vec{b}_n)^\top Y^2 v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) = \\ &= \sum_{j_1,\ldots,j_n=0}^{1} \sum_{k_1,\ldots,k_n=0}^{1} \alpha_{j_1,\ldots,j_n}\alpha_{k_1,\ldots,k_n} \prod_{z=0}^{n_{AI}} \left[\begin{pmatrix} 1 \\ \vec{b}_z \end{pmatrix}^\top A^{(j_z)}A^{(k_z)} \begin{pmatrix} 1 \\ \vec{a}_z \end{pmatrix}\right] \prod_{z \geq m, z \neq k} \left[\begin{pmatrix} 1 \\ \vec{b}_z \end{pmatrix}^\top B_{j_z}B_{k_z} \begin{pmatrix} 1 \\ \vec{a}_z \end{pmatrix}\right] \times \\ &\qquad \times \left[\begin{pmatrix} 1 \\ -\vec{a}_k \end{pmatrix}^\top B_{j_k}B_{k_k} \begin{pmatrix} 1 \\ \vec{a}_k \end{pmatrix}\right] \\ &= \sum_{j_1,\ldots,j_{n_A}=0}^{1} \sum_{k_1,\ldots,k_{n_A}=0}^{1} \alpha_{j_1,\ldots,j_{n_A},j_m^{\mathbf{0}},\ldots,j_n^{\mathbf{0}}}\alpha_{k_1,\ldots,k_{n_A},j_m^{\mathbf{0}},\ldots,j_n^{\mathbf{0}}} \prod_{z=0}^{n_{AI}} \left[\begin{pmatrix} 1 \\ \vec{b}_z \end{pmatrix}^\top A^{(j_z)}A^{(k_z)} \begin{pmatrix} 1 \\ \vec{a}_z \end{pmatrix}\right]. \end{aligned} \tag{A4}$$

Consider the different possibilities for $k_1, \ldots, k_n$ for which $\alpha_{k_1,\ldots,k_{n_{AI}},j_m^{\mathbf{0}},\ldots,j_n^{\mathbf{0}}} \neq 0$. There are less than or equal to ℓ many occurrences of k_i $(1 \leq i \leq n_{AI})$ with $k_i = 0$. If there are less, then (A3) implies that the final product in (A4) vanishes, hence the corresponding summand does not contribute to (A4). On the other hand, if there are exactly ℓ

many, then (A3) implies that this product vanishes unless the occurrences of $k_i = 0$ agree with the occurrences of $j_i^{\mathbf{0}} = 0$. Similar argumentation works for the $j_1, \ldots, j_n$, and if we also use (A2), we finally get

$$0 \leq \underbrace{\alpha^2_{j_1^{\mathbf{0}},\ldots,j_n^{\mathbf{0}}}}_{\neq 0} \prod_{1\leq z\leq n_{AI}: j_z^{\mathbf{0}}=1} \left[\begin{pmatrix} 1 \\ \vec{b}_z \end{pmatrix}^{\top} \left(A^{(1)}\right)^2 \begin{pmatrix} 1 \\ \vec{a}_z \end{pmatrix} \right].$$

The product runs over $n_{AI} - \ell$ many indices, so there is at least one z such that $1 \leq z \leq n_{AI}$ and $j_z^{\mathbf{0}} = 1$. We have not yet chosen the corresponding $\vec{b}_z$ and $\vec{a}_z$; it is easy to see that we can choose them so that the terms in the product in the previous expression attain any sign we want. This produces a contradiction, like in Case 1. □

Appendix B: Which assumptions could possibly be dropped or weakened?

One candidate assumption that one might consider to weaken is the assumption that the group of single-gbit reversible transformations must be $\mathrm{SO}(d)$. It is natural to assume that this group must be able to map every pure gbit state to any other (and thus be transitive on the $(d-1)$-sphere). In fact, for odd $d \neq 7$, this demand already singles out $\mathrm{SO}(d)$. However, if d is even or $d = 7$, then there are other transitive groups (such as $\mathrm{SU}(2)$ for $d = 4$), and the analysis of the present paper is in principle applicable to this more general situation. The case of $n = 2$ gbits has been treated in this more general setting in [52]. There it was shown that these other groups do not work either in the two-gbit case. It seems reasonable to conjecture from our results and the results in [52] that also for more gbits groups other than $\mathrm{SO}(d)$ fail to yield any non-trivial solution. Furthermore, $\mathrm{SO}(d)$ is the natural choice for generalizing the geometrical meaning of the Bloch ball for spin-$\frac{1}{2}$ particles to higher spatial dimensions, namely that the Bloch vector defines a direction in physical space.

Another route might be to drop tomographic locality, as in [23]. In fact, the $d = 2$ Bloch ball corresponds to the quantum bit over the real numbers, and if we simply define the corresponding n-gbit state space to be the 2^n-level quantum states over the reals, then this defines a model with interesting computational power (namely, equal to standard quantum computation), albeit one that does not satisfy tomographic locality. A similar construction can be performed for the $d = 5$ case of quaternionic quantum theory [67] (but see the subtleties pointed out in [45, 46]). The problem is, however, that these two cases are extremely special: building the composite state space uses the postulate that the result is supposed to be a Euclidean Jordan algebra. This assumption is not consistent with any of the other cases $d \notin \{2, 3, 5\}$.

Furthermore, tomographic locality is a very natural postulate: it formalizes the idea that the whole is just composed of its parts and the relations between them. Other forms of state space composition would have to violate this intuition. Furthermore, they would have to violate the fact that states of composite systems can be described by tensors, a fundamental structural property of quantum theory with a myriad of physical consequences.

A possible way to drop tomographic locality despite these problems would be to instead assume (some version of) *purification* [28, 68, 69]. While purification has been very successful as a postulate of quantum theory, in particular, by illuminating how several characteristic properties of quantum theory can be understood directly via diagrammatic reasoning [68], it is also very strong as a postulate. In fact, it is so strong that it is currently not clear whether there are any theories other than standard complex quantum theory and some of its subtheories [70] that satisfy it. A potential alternative can be found in the work by Galley and Masanes [57, 58] who have pioneered an approach to construct composite state spaces directly in terms of group representations, without assuming tomographic locality.

We have made the implicit assumption that computations are carried out in the following way: first, the input is encoded into the initial state; then the actual computation is performed fully reversibly; and finally, the output is read out by a measurement. While this is arguably a natural standard scenario in the *reversible* context, one might ask whether allowing measurements at any point during the computation could increase the computational capabilities of a theory. This is not the case in standard quantum mechanics, where all measurements can be modelled as unitary transformations on the system and an ancilla. But in principle, it might be true for other computational probabilistic theories.

Finally, one could drop the assumption of reversibility and/or connectedness of the groups, and consider transformations that are elements of some semigroup or finite group. However, dropping connectedness means giving up continuous (time) evolution, a large step away from our current conception of physics. Similarly, dropping reversibility means a substantial departure from our current understanding of fundamental physics: it would mean to give up conservation of information at the fundamental level.

Remark:
As mentioned before, de la Torre et al. (2012) analyzed the three-dimensional case. To prove that only quantum theory allows for interactions of three-dimensional balls given by reversible time evolution, de la Torre et al. had to assume additionally that Bloch balls can be exchanged, i.e. permuted.

Chapter 7

Quantum Darwinism and the spreading of classical information in non-classical theories

Authors: Roberto D. Baldijão*, Marius Krumm*, Andrew J. P. Garner, Markus P. Müller

Submitted to *Quantum* for review (27th of Dec., 2020).
Preprint available as: arXiv:2012.06559
https://arxiv.org/abs/2012.06559

Contribution: Roberto D. Baldijão and I share first authorship of the paper. I made essential contributions to all parts of the research. I provided the leading input for the sufficient conditions for GPT Darwinism. I massively simplified Roberto D. Baldijão's argument for GPT Darwinism in Spekken's toy model, leading to the fanout gate and proof that it achieves GPT Darwinism in Spekken's toy model as presented in the thesis.

Abstract: Quantum Darwinism posits that the emergence of a classical reality relies on the spreading of classical information from a quantum system to many parts of its environment. But what are the essential physical principles of quantum theory that make this mechanism possible? We address this question by formulating the simplest instance of Darwinism – CNOT-like *fan-out* interactions – in a class of probabilistic theories that contain classical and quantum theory as special cases. We determine necessary and sufficient conditions for any theory to admit such interactions. We find that every non-classical theory that admits this spreading of classical information must have both entangled states and entangled measurements. Furthermore, we show that Spekkens' toy theory admits this form of Darwinism, and so do all probabilistic theories that satisfy principles like strong symmetry, or contain a certain type of decoherence processes. Our result suggests the counterintuitive general principle that in the presence of local non-classicality, a classical world can only emerge if this non-classicality can be "amplified" to a form of entanglement.

I. INTRODUCTION

Quantum Darwinism [1–12] addresses one of the toughest questions raised by quantum theory: If the universe is fundamentally described by quantum mechanics, how does an objective classical world arise? At the heart of this question is a tension between the microscopic quantum realm, in which systems happily exist in states of super-imposed possibility, and the macroscopic world of "classical" systems (such as the pointer needle of a read-out gauge), which are only ever observed in definite objective states. Several mechanisms and formalisms have been proposed which intend to provide a bridge between the quantum and classical realms, including the formal limit of $\hbar \to 0$ [13], saddle point approximations to the path integral [14], and the process of environment-induced decoherence [1, 15].

Quantum Darwinism identifies key prerequisite for such a bridge to arise: there must be a mechanism by which some aspect of a quantum system can be spread out to many parts of its environment. Particularly, since the no–cloning theorem [16] forbids the copying of quantum information, this means some classical information from the system must be copied into its environment in such a way that given long enough (and enough of the environment), this information can be learned through enough measurements on the environment.

Here we ask: What are the essential features of quantum theory that enable this spreading of classical information in the first place? Certainly, this is possible in Quantum Theory's rich mathematical structure of complex Hilbert spaces, but can we identify a selective subset of more physically–motivated principles that similarly enable this Darwinistic emergence of classical reality? To approach this, we adopt the minimal–assumptions framework of generalized probabilistic theories (GPTs) [17, 18]. These encompass a wide class of operational scenarios, in which a physical system is entirely characterized by its experimental statistics resulting from preparation and subsequent measurement procedures. The GPT approach has thus far enjoyed particular success in identifying which operational features are necessary or sufficient for quantum phenomena like teleportation [19], no-cloning [20], entanglement [18], phase and interference [21, 22], or decoherence [23]. With this article, we aim to extend this canon to include Quantum Darwinism.

We begin by recalling the essential features of Quantum Darwinism (section II A), and providing a brief overview of the GPT framework (section II B). We then proceed to the results of the article: an operational formulation of Quantum Darwinism (section III A), followed by necessary (section III B) and sufficient (section III C) conditions for such to exist. Particularly, we show that both entangled states and entangled measurements are necessary features in any non-classical theory that exhibits Darwinism, suggesting the counterintuitive general principle that in the presence of local non-classicality, a classical world can only emerge if this non-classicality can be "amplified" to a form of entanglement. We then identify how other physically–motivated features, such as the no-restriction hypothesis [24, 25] and strong symmetry [26], or the existence of decoherence [23], are sufficient to imply the presence of Darwinism. Finally (section III D), we give a concrete example of a non-classical theory other than quantum theory that admits Darwinism: we show its existence in Spekkens' Toy Model [27] and its convex extensions [21, 25].

II. BACKGROUND

A. Quantum Darwinism

The typical setting of Quantum Darwinism (QD) [1–3] consists of a central system $\mathcal{S}$ interacting with a multipartite environment $\mathcal{E}_1, \dots \mathcal{E}_N$. This is similar to the setting in which *decoherence* is studied (e.g. [15]), but rather than focusing on the change in $\mathcal{S}$'s state, QD is concerned with the information that fragments of the environment can learn about $\mathcal{S}$.

Not everything about $\mathcal{S}$ can be spread to the environment – for instance, sharing arbitrary quantum information would violate the *no cloning principle* [16]. Nonetheless, something can still be learned about $\mathcal{S}$ – perhaps because the interaction induces certain quantum states on system and environment such that measurements made on $\mathcal{S}$ and $\{\mathcal{E}_i\}$ in the right choice of basis yield correlated outcomes. This interaction must also preserve some aspect of the initial state of $\mathcal{S}$, so that what the environment learns can be considered as being about $\mathcal{S}$.

In the ideal scenario, we would like to extract as much classical information from any $\mathcal{E}_i$ about $\mathcal{S}$ as we could from $\mathcal{S}$ directly. Holevo's theorem [28] tells us that the most information that can possibly be shared with each environmental system is upper-bounded by that directly obtainable from a single measurement on $\mathcal{S}$. This can be realized as follows, when $\mathcal{S}$ and all of $\mathcal{E}_i$ $(i = 1 \dots N)$ are d-dimensional quantum systems. Let $\mathcal{M} := \{|0\rangle, \dots |d-1\rangle\}$ be some orthonormal basis. Suppose $\mathcal{S}$ is initially in a pure state $|\psi\rangle_S = \sum \alpha_k |k\rangle$, and each environmental system starts in a pure basis state $|j_i\rangle_i \in \mathcal{M}$. Consider the following *fan-out* gate (a generalization of control-NOT / control-shift gates, see fig. 1):

$$\mathrm{FAN}\left(|k\rangle_{\mathcal{S}} \otimes \bigotimes_{i=1}^{N} |j_i\rangle\right) := |k\rangle_{\mathcal{S}} \otimes \bigotimes_{i=1}^{N} |j_i \oplus k\rangle\,, \qquad (1)$$

such that

$$\mathrm{FAN}\left(|\psi\rangle_{\mathcal{S}} \otimes \bigotimes_{i=1}^{N} |j_i\rangle\right) = \sum_k \alpha_k \mathrm{FAN}\left(|k\rangle_{\mathcal{S}} \otimes \bigotimes_{i=1}^{N} |j_i\rangle\right)$$
$$= \sum_k \alpha_k |k\rangle_{\mathcal{S}} \otimes \bigotimes_{i=1}^{N} |j_i \oplus k\rangle\,, \qquad (2)$$

where $j_i \oplus k$ indicates addition modulo d.

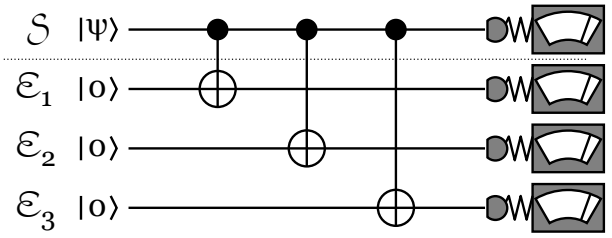


Figure 1. **Ideal Quantum Darwinism: fan-out gate.** The fan-out gate (eq. (2)) is realized for the case $N = 3$, $d = 2$ by three consecutive CNOT gates. After this process, the statistics of the computational-basis measurement Z on all of the environmental subsystems ($\mathcal{E}_1$, $\mathcal{E}_2$ and $\mathcal{E}_3$) agrees with those of the the main system ($\mathcal{S}$). Meanwhile, the statistics of this measurement on the system are the same as if it had been directly made on $|\psi\rangle$. As such, the classical information about Z in $\mathcal{S}$ has been spread to its environment.

It is clear that fan-out realizes the above ideals, perfectly broadcasting classical information about $\mathcal{M}$ on $\mathcal{S}$ to every environment system, while preserving the outcome probabilities of $\mathcal{M}$ on $\mathcal{S}$. First, if $|\psi\rangle_{\mathcal{S}} \in \mathcal{M}$ (as in eq. (1)), then it remains unchanged after the interaction; that is, $\mathcal{M}$ is the *pointer basis* selected by FAN. Moreover, if $|\psi\rangle_{\mathcal{S}} \in \mathcal{M}$, this element can be perfectly identified by simply measuring any of $\mathcal{E}_i$ with $\mathcal{M}$ and applying the appropriate relabelling (subtraction of j_i modulo d) – as a consequence of the so-called *einselection process* [1]. Furthermore, when $|\psi\rangle_{\mathcal{S}}$ is a superposition of multiple states in $\mathcal{M}$, the resulting entangled state now has the property that whatever the outcome of $\mathcal{M}$ on $\mathcal{S}$, the *same* outcome will be obtained by making $\mathcal{M}$ on $\mathcal{E}_i$ (again, via subtraction of j_i modulo d). Finally, the statistics of measuring $\mathcal{M}$ on $\mathcal{S}$ before and after the fan-out are identical. Thus, such a fan-out implements an *ideal Darwinism process*.

In this idealized setting, *any* state in $\mathcal{M}$ represents a valid initial state of an environment $\mathcal{E}_i$ for which FAN can register information about $\mathcal{S}$'s pointer basis. This multiplicity of "good registers" makes the process more robust to modifications in the initial state of the environment subsystems – generically reducing 'misalignment', in the language of Zwolak et al. [11]. In addition, this type of interaction aligns with physically-motivated models of Darwinism [1, 8, 10, 12, 29–31].

QD can also encompass more complicated scenarios [3, 8–12], where only partial information is spread (typically quantified through mutual–information quantities – though the efficacy of this is of debate [6, 7]). For instance, pointer states may not be perfectly robust to interaction, the information may not be perfectly registered in the environment [8, 10, 12], or a more general class of measurement than projection onto the pointer basis may be used [4, 5]. However, for the purpose of this article (and preempting the need to cast the scenario in the operational language of GPTs), we restrict our discussion here to the idealized case described above.

B. The GPT framework

Generalized probabilistic theories (GPTs) are a minimal–assumptions framework in which a physical theory is specified by the statistics of every experiment that could be conducted within it. The fundamental elements of a GPT correspond to laboratory operations, such as state preparations, and measurement outcomes. In addition to the aforementioned isolation of quantum features [18–23], this broad operational approach makes the GPT framework well-suited for attempts to reconstruct quantum theory either from experimental data [32] or from sets of reasonable physically–motivated axioms [17, 26, 33, 34]. Theories such as quantum theory (QT) and classical probability theory (CPT) are GPTs, but the framework also admits more exotic theories such as "boxworld" [18] or higher-dimensional Bloch ball state spaces [35].

In this section, we briefly review the aspects of the GPT framework that are relevant for our discussion. Readers who are familiar with the GPT framework may wish to skip to the summary of assumptions at the end of the section. For more detailed and pedagogical introductions to the GPT framework, see e.g. [17, 18, 36].

1. Single Systems

The primitive elements of a GPT are the *states* that one can prepare, and the outcomes of measurements (known as *effects*) that one can make on a given physical system. Mathematically, states (not necessarily normalized) are given by the elements of a closed subset A_+ of some finite-dimensional real vector space A. With a slight abuse of notation, the physical system will also be denoted A. This subset A_+ is assumed to be a *cone*, meaning that $\varphi, \omega \in A_+$ and $\lambda \geq 0$ imply that $\lambda\varphi \in A_+$ and $\varphi + \omega \in A_+$. Furthermore, A_+ is assumed to be *generating*, i.e. $\mathrm{span}(A_+) = A$, and *pointed*, i.e. $A_+ \cap (-A_+) = \{0\}$. (For the example of QT, this is the cone of positive semidefinite matrices, see example 1 below.)

Effects correspond to elements in a generating cone $E_A \subseteq A^*$, where A^* is A's dual space of linear functionals. The probability of observing effect $e \in E_A$ given a preparation $\omega \in A_+$ is given by $e(\omega)$. Since this must be non-negative, we must have $E_A \subseteq A_+^*$, where $A_+^* := \{e \in A^* \mid e(\omega) \geq 0 \text{ for all } \omega \in A_+\}$ is the *dual cone* of A_+ [37]. We assume the existence of a distinguished *unit effect* $u_A \in E_A$ such that for all $a \in E_A$ there is some $\lambda > 0$ with $a \leq \lambda u_A$ (where $a \leq b$ if and only if there exists some $c \in E_A$ such that $a + c = b$). The *measurements* of a theory correspond to collections of effects $\{e_i\}_{i=1\ldots N}$ that sum to u_A – each constituent effect corresponds to one mutually exclusive outcome. Since $\sum_i e_i(\omega) = u_A(\omega)$, we can interpret $u_A(\omega)$ as the

normalization of the state ω — that is, the total probability to obtain any outcome if the measurement is performed on the corresponding physical system. We say an effect is *valid* if it can be part of a measurement (i.e. $e \in E_A$ and $e \leq u_A$).

If $E_A = A_+^*$, we say that the system is *unrestricted*, or that it *satisfies the no-restriction hypothesis* [24, 25]. From u_A and A_+, one can infer its compact convex set of *normalized states* $\Omega_A := \{\omega \in A_+ \,|\, u_A(\omega) = 1\} \subset A_+$. An example is sketched in fig. 2.

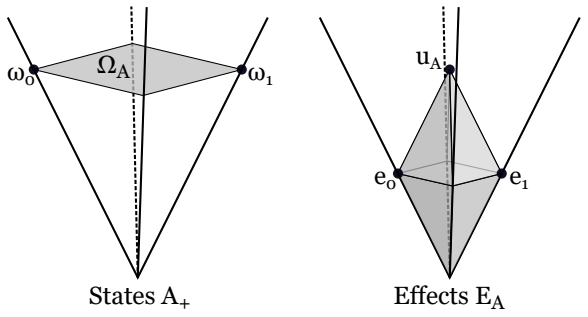


Figure 2. **Geometric picture of a GPT.** An example state space A_+ (LHS) and effect space E_A (RHS) of a GPT with $A = \mathbb{R}^3$ is drawn. On the RHS, the *unit effect* u_A is labeled, and all effects on or within the shaded octahedron are *valid* in that $e \leq u_A$. Two *pure* effects $\{e_0, e_1\}$ that satisfy $e_0 + e_1 = u_A$ and hence form a *refined measurement* are labelled. On the LHS, the convex set of normalized states $u_A(\omega) = 1$ is shaded as Ω_A. Within it, a *maximal frame* of two pure states $\{\omega_0, \omega_1\}$ is labelled.

The convexity of A_+ and E_A amounts to the assumption that statistical fluctuations can always be introduced into an experiment. Consider measurement outcome $e \in E_A$ on one of two preparations ω_1 or ω_2, with respective statistics $e(\omega_1)$ and $e(\omega_2)$. If ω_1 is prepared with probability p and ω_2 otherwise, then this preparation procedure should be representable by the single state ω whose statistics satisfy: $e(\omega) = pe(\omega_1) + (1-p)e(\omega_2) = e(p\omega_1 + (1-p)\omega_2)$. It then follows that $\omega = p\omega_1 + (1-p)\omega_2$. A similar interpretation of convex combinations applies to the effects.

An effect $e \in E_A$ is said to be pure [33, 38] if $e = \sum f_i$, with $f_i \in E_A$, implies $f_i \propto e$ for all i (see also fig. 2). Pure effects cannot be obtained from (non-trivial) coarse-graining of other effects. A collection of pure effects that sum to u_A with no effects proportional to any other in the set is known as a *refined measurement*. A *pure state* is defined to be a normalized state that is extremal in Ω_A, i.e. that cannot be written as a non-trivial convex combination of other normalized states. A *frame* is a collection of pure states $\{\omega_j\}$ that can be perfectly distinguished in a single measurement: i.e. there is at least one measurement $\{e_i\}$ such that $e_i(\omega_j) = \delta_{ij}$. A *maximal frame* is a frame with the largest number of distinguishable states for that system.

Dynamics in GPTs are described by linear maps $T : A \to A$ known as *transformations*. Transformations T must map states to states, i.e. $T(A_+) \subseteq A_+$, and effects to effects, in the sense that if $e \in E_A$ is a valid effect, then $e \circ T$ must also be a valid effect. (The latter corresponds to an outcome where transformation T has been applied before the measurement.) Motivated by the intuition to consider only closed-system dynamics in which all environments are explicitly modelled, we will in the following restrict our attention to *reversible transformations*. These are transformations T that are invertible as a linear map and whose inverse T^{-1} is also a transformation. Since transformations can be composed, it follows that the reversible transformations of any GPT system A form a group $\mathcal{T}_A$. Furthermore, they map the set Ω_A of normalized states onto itself.

In summary, a GPT system A is defined by a tuple $(A, A_+, E_A, u_A, \mathcal{T}_A)$ of a real vector space, the state and effect cones, the unit effect, and the group of reversible transformations. Let us illustrate this framework with two familiar examples:

Example 1 (Quantum theory (QT))**.** *An n-level quantum system corresponds to the GPT system $(A^{(n)}, A_+^{(n)}, E_A^{(n)}, u_A^{(n)}, \mathcal{T}_A^{(n)})$ with*

$$A^{(n)} = \mathbf{H}_n(\mathbb{C}), \quad A_+^{(n)} = \mathbf{H}_n^+(\mathbb{C}) \simeq E_A^{(n)}, \quad u_A^{(n)} = \mathbf{1}_n,$$
$$\mathcal{T}_A^{(n)} = \{\rho \mapsto U\rho U^\dagger \mid U^\dagger U = \mathbf{1}_n\},$$

where $\mathbf{H}_n(\mathbb{C})$ is the real vector space of $n \times n$ complex Hermitian matrices, and $\mathbf{H}_n^+(\mathbb{C})$ the subset of positive semidefinite matrices. Via the Hilbert-Schmidt inner product, $\langle X, Y\rangle := \mathrm{tr}(XY)$, we can identify $A^{(n)}$ with its dual space such that the effects are also Hermitian matrices. For example, $u_A^{(n)}(\rho) = \mathrm{tr}(\rho)$ can be written $\langle \mathbf{1}_n, \rho\rangle$, hence we can identify $u_A^{(n)} = \mathbf{1}_n$.

The measurements $\{E_i\}_{i=1,\ldots,N}$ thus correspond to POVMs (positive operator-valued measures), i.e. $E_i \geq 0$ and $\sum_i E_i = \mathbf{1}_n$. The normalized states $\Omega_A^{(n)}$ are the (unit-trace) density matrices, and the maximal frames correspond to the various n-element orthonormal bases of the Hilbert space $\mathbb{C}^n$. The reversible transformations are the unitary conjugations. Pure effects correspond to rank-1 POVM elements.

Example 2 (Classical probability theory (CPT))**.** *A classical random variable that can take n different values corresponds to the GPT system $(B^{(n)}, B_+^{(n)}, E_B^{(n)}, u_B^{(n)}, \mathcal{T}_B^{(n)})$ with*

$$B^{(n)} = \mathbb{R}^n, \quad B_+^{(n)} = \{x \in \mathbb{R}^n \mid \textit{all } x_i \geq 0\} \simeq E_B^{(n)},$$
$$u_B^{(n)} = (1, 1, \ldots, 1)^{\mathrm{T}}, \quad \mathcal{T}_B^{(n)} \simeq S_n.$$

In this notation, we have identified $\mathbb{R}^n$ with its dual space via the usual dot product $x \cdot y = \sum_i x_i y_i$. The unit effect is thus $u_B^{(n)} \cdot p = \sum_{i=1}^n p_i$, and so $\Omega_B^{(n)}$ is the simplex of n-dimensional probability vectors, i.e. $\Omega_B^{(n)} = \{p \in \mathbb{R}^n \mid p_i \geq 0, \sum_i p_i = 1\}$. The reversible transformations are the permutations of the entries:

$p_i \mapsto p_{\pi(i)}$, *with* π *some permutation of* $\{1, 2, \ldots, n\}$. *Thus, the group of reversible transformations is a representation of the permutation group* S_n.

A crucial signature of classicality is that CPT has (up to relabelling) only a single *refined measurement* $\{e_i\}$. *Its effects are* $e_i(p) = p_i$, *and it can be interpreted as asking which of the n possible configurations is actually the case. It distinguishes the (up to relabelling) unique maximal frame* $\{\omega_j\}$, *where* $\omega_j := (0, \ldots, 0, \underbrace{1}_{j}, 0, \ldots, 0)^{\mathrm{T}}$.

Both QT and CPT are *unrestricted* and *self-dual* [39], i.e. there is some inner product according to which $A_+ = E_A$. Note that GPTs will in general satisfy neither of these two properties.

2. Maximal classical information (MCI) frames

Our goal is to generalize ideal Quantum Darwinism – in particular, the mechanism for perfect spreading of classical information via fan-out gates – to GPTs. As a first step, we have to identify the analogue of the pointer states and the measurements that read out their encoded classical information. We will focus on Darwinism generalizations that allow one to extract the maximal amount of classical information. In the quantum case, such classical information is encoded onto an orthonormal basis $\{|j\rangle\}$. The natural analogue of this in a GPT is a maximal frame $\{\omega_j\}$.

Let us consider the measurements that could extract this classical information. As seen in example 1, QT enjoys a strong form of duality that allows one to treat the pure states $\omega_j = |j\rangle\langle j|$ and the corresponding rank-1 projective measurements $e_j(\bullet) = \mathrm{Tr}[|j\rangle\langle j| \ \bullet]$ as the "same" objects, and it is exactly this dual set of rank-1 projectors that form the measurement that extracts the maximal amount of information out of the system.

In general, GPTs do not have such an automatic duality between states and effects. Moreover, measurements that distinguish the elements of a maximal frame do not even need to be refined. However, since we are interested in the idealized case, where one spreads the maximal classical information contained in some system, we will here focus on maximal frames that can be distinguished by a refined measurement:

Definition 1 (Maximal classical information in GPTs)**.** *A maximal frame* $\omega_1, ..., \omega_n$ *is called a* ***maximal classical information frame*** *(MCI-frame) if there is a refined measurement* $\{e_i\} \subset E_A$ *which discriminates the states* ω_j, *i.e.* $e_j(\omega_k) = \delta_{jk}$.

Many GPTs contain MCI-frames: quantum theory certainly does (in the form of orthonormal bases), and so do quantum theory over the real numbers and over the quaternions, and d-ball state spaces. As expected, classical theories in all dimensions also have MCI-frames. Furthermore, so-called "dichotomic" systems as defined in Ref. [40] contain MCI-frames, which includes unrestricted systems whose sets of normalized states are regular polygons with an even number of vertices, or a d-cube or d-octoplex for $d \geq 3$.

In appendix A, we give an example of a state space that does *not* have an MCI-frame: the pentagon. This example illustrates the counterintuitive properties of such systems: the pentagon has at most *two* perfectly distinguishable states, but one can in some sense encode *more* than one bit of information into such a system [41]. That is, any classical bit that sits inside this state space does not represent the maximal amount of information that can be encoded into the system. For the remainder of this work, we will thus exclude such systems and focus on state spaces that contain MCI-frames.

3. Composite systems

Darwinism is inherently linked to composition of subsystems; therefore, we need to understand how to treat composition in GPTs. There are several approaches to this [24, 42], including category-theoretic formulations [43]. Here, we will motivate and state a list of minimal assumptions on a state space AB, composed of two state spaces A and B, that allows us to formulate a generalization of Darwinism. For the case of more than two subsystems, we assume that the joint state satisfies all desiderata on all pairs of subsystems.

First, we demand that the combined state space AB has a notion of independent parallel preparation. This means that given some state φ^A on A and some state ω^B on B, there should be a state of AB (denoted $\varphi^A \odot \omega^B$) that represents the state obtained by the independent local preparation of the two states on A and on B. Since statistical mixtures of local preparations must lead to statistical mixtures of the corresponding global state, the map $\odot$ must be bilinear.

As pure states can be interpreted as states of maximal knowledge, we assume that independent parallel preparations of pure states lead to global pure states, i.e. we demand that if φ^A and ω^B are pure then so is $\varphi^A \odot \omega^B$ [38]. Likewise, there should exist a notion of parallel implementation of measurements on the systems. For this we require another bilinear function (also denoted by $\odot$) that maps effects $e^A \in E_A$, $f^B \in E_B$ to effects $e^A \odot f^B \in E_{AB}$. Furthermore, if one performs a parallel implementation of two local measurements on a composite state whose parts were prepared independently in parallel, then the probabilities should factorize in the sense that $e^A_j \odot f^B_k(\varphi^A \odot \omega^B) = e^A_j(\varphi^A) f^B_k(\omega^B)$. In other words, independent local procedures lead to statistical independence. The bilinearity of $\odot$ ensures the validity of the no-signalling principle: the choice of local measurement on B does not affect the outcome probabilities of local measurements in A (and vice-versa). Indeed, $\sum_j e^A_i \odot e^B_j = e^A_i \odot u^B$, for all effects e^A_i and any measurement $\{e^B_j\}$.

Similarly as for states, we assume that the composition of pure effects results in a pure effect.

Finally, we must ensure that the global structure is consistent with the local structure. Consider a valid composite effect $e^{AB} \in E^{AB}$ and a normalized state $\omega^B \in \Omega_B$. Then the effect $\tilde{e}^A$ defined by $\tilde{e}^A(\varphi^A) := e^{AB}(\varphi^A \odot \omega^B)$ should be valid effect on A, i.e. $\tilde{e}^A \in E_A$ and $\tilde{e}^A \leq u_A$: it can be implemented by preparing ω^B on B and then measuring e^{AB} on AB.

Similarly, consider a global state $\omega^{AB} \in \Omega_{AB}$ shared between two parties A and B. Imagine that one of the parties, say B, implements a local measurement $\{f_k^B\}_k \subset E_B$ and tells the other party the outcome k. Then agent A holds a conditional state, which should be a (subnormalized) element of the state space of A. More specifically, for an effect $e^A \in E_A$, the probability for both f_k^B and e^A to be obtained is given by $e^A \odot f_k^B(\omega_{AB})$. This implicitly defines a subnormalized state $\tilde{\omega}^A$ on A via $e^A(\tilde{\omega}^A) = e^A \odot f_k^B(\omega^{AB})$, which must thus be an element of A_+. In the special case of the trivial measurement $f_k^B = u^B$, the state $\tilde{\omega}^A$ becomes the reduced state on A. A similar condition should hold if the roles of A and B are interchanged.

Together, we will call these assumptions the *minimal assumptions on composition*.

Definition 2 (Minimal assumptions on composition)**.** *A composition of GPT systems A and B is a GPT system AB together with two bilinear maps $A \times B \to AB$ and $A^* \times B^* \to (AB)^*$, both denoted by $\odot$, satisfying the following:*

i. All product states are allowed and normalized: if $\omega^A \in \Omega_A$ and $\omega^B \in \Omega_B$ then $\omega^A \odot \omega^B \in \Omega_{AB}$.

ii. All products of valid effects are valid effects: if $e^A \in E_A$ and $e^B \in E_B$ then $e^A \odot e^B \in E_{AB}$. In particular, local measurements cannot lead to probabilities larger than 1: $u^A \odot u^B \leq u^{AB}$.

iii. Local measurements on product states yield statistically independent outcomes: $e^A \odot f^B(\omega^A \odot \omega^B) = e^A(\omega^A) f^B(\omega^B)$.

iv. Products of pure states (effects) are pure states (effects).

v. Conditional effects: for all effects $e^{AB} \in E_{AB}$ and all normalized states $\varphi^A \in \Omega_A$ and $\omega^B \in \Omega_B$, also $e^{AB}(\varphi^A \odot \bullet) \in E_B$ and $e^{AB}(\bullet \odot \omega^B) \in E_A$ are effects.

vi. Conditional states: for all states $\omega^{AB} \in \Omega_{AB}$ and all effects $e^A \in E_A$, $f^B \in E_B$, the vectors $\tilde{\omega}_A$, $\tilde{\omega}_B$ which are implicitly defined via

$$\tilde{e}^A(\tilde{\omega}^A) = \tilde{e}^A \odot f^B(\omega^{AB})$$
$$\tilde{f}^B(\tilde{\omega}^B) = e^A \odot \tilde{f}^B(\omega^{AB})$$

must be states, i.e. $\tilde{\omega}_A \in A_+$, $\tilde{\omega}_B \in B_+$.

While these assumptions imply the no-signalling principle, we do *not* demand the popular principle of "tomographic locality" [17], i.e. that the $\omega^A \odot \omega^B$ span all of AB. Thus, the $\odot$ operation cannot in general be identified mathematically with the tensor product operation. The above minimal assumptions are also compatible, for example, with QT over the real numbers [17].

As we know from QT, a striking feature of composite systems in non-classical theories is entanglement. Having a definition of composite systems at hand, we are in place to define entangled states and effects in GPTs [19]:

Definition 3 (Entangled states)**.** *Consider a composite system $A = A_1 A_2 \dots A_N$. States $\omega^A \in \Omega_A$ which can be written as*

$$\omega^A = \sum_i p_i \, \omega_i^{A_1} \odot \omega_i^{A_2} \odot \dots \odot \omega_i^{A_N} \tag{3}$$

with $\omega^{A_i} \in \Omega_{A_i}$ and $\{p_i\}$ a probability distribution, are called separable*. States which cannot be written in this form are called* entangled*.*

Definition 4 (Entangled effects)**.** *Effects $e^A \in E_A$ which can be written as*

$$e^A = \sum_i e_i^{A_1} \odot e_i^{A_2} \odot \dots \odot e_i^{A_N} \tag{4}$$

with $e_i^{A_j} \in E_{A_j}$ are called separable*. Effects which cannot be written in this form are called* entangled*.*

A pure effect is separable if and only if it is a product of pure effects (see, e.g. appendix B).

Summary of assumptions. We consider theories that satisfy:

- For a single system $(A, A_+, E_A, u_A, \mathcal{T}_A)$: A is finite-dimensional. We do *not* assume the no-restriction hypothesis.
- For pairs of systems: composition satisfies all conditions of definition 2. In particular, products of pure states (or effects) are pure, but we do *not* assume tomographic locality.
- For three or more systems: composition satisfies all conditions of definition 2 on all subsystem pairs. For example, a quadripartite system $ABCD$ is a valid composition of AB and of CD, with subsystem ABC being a valid composition of B and AC, and so forth.

Unless otherwise stated, all introduced states are normalized, and all introduced effects are valid ($e \leq u_A$).

III. RESULTS

A. A definition of Darwinism in GPTs

With all these ingredients we can now ask: is the ideal mechanism for Darwinism present in GPTs other than quantum theory? To answer this, we must first formulate the features of the ideal Darwinism process in an *operational* way – that is, in terms of experimental statistics.

To this end, recall the scenario of ideal *Quantum* Darwinism (section II A). The desire is to broadcast some classical information encoded within $\mathcal{S}$ to the environment, say, relating to pointer measurement $\mathcal{M} := \{|k\rangle\langle k|\}_{k=0\ldots d-1}$. Let each environment system begin in an eigenstate of $\mathcal{M}$ (for system i, labeled $|j_i\rangle$), then after the fan-out operation T (eq. (2)), the outcome probabilities when measuring $\mathcal{M}$ on any environment $\mathcal{E}_i$ will satisfy

$$\mathrm{P}_{\mathcal{S}}(\mathcal{M}=k) = |\alpha_k|^2 = \mathrm{P}_{\mathcal{E}_i}(\mathcal{M}=j_i+k) \quad \text{for all } k. \tag{5}$$

Moreover, if one makes the joint measurement $\mathcal{M}^{\otimes(N+1)} = \{|k_0\rangle\langle k_0|\otimes\cdots\otimes|k_N\rangle\langle k_N|\}$ on the entire composite system, the probability of outcome $(k_0,...,k_N)$ is

$$p(k_0,\ldots k_N) = |\alpha_k|^2\delta_{k_{\mathbf{0}},k_1-j_1}\ldots\delta_{k_{\mathbf{0}},k_N-j_N}. \tag{6}$$

This is the sense in which objectivity can emerge under Quantum Darwinism: when this mechanism succeeds, all independent observers can learn about the same (maximal) classical information and agree about their findings. Moreover, $\mathrm{P}_{\mathcal{S}}(\mathcal{M}=k)$ is the same before and after the fan-out is performed.

To generalize Darwinism to the GPT framework, we must capture the same operational behaviour on the level of probabilities. First, we need an analogue of pointer states – a set of distinguishable states corresponding to the classical information to be broadcast. As mentioned in section II B 2, this role is played by an MCI-frame $\{\omega_j^{(0)}\}_{j=0,\ldots d-1}$ of $\mathcal{S}$ and its corresponding refined measurement $\{e_k^{(0)}\}$ with the distinguishing property $e_k^{(0)}(\omega_j^{(0)}) = \delta_{jk}$. Again, we assume the main system is in some pure state ν, that may not be an element of $\{\omega_j^{(0)}\}_{j=0,\ldots d-1}$. Lacking the mathematical structure of a Hilbert space, we cannot so easily express ν as a superposition of frame elements. Nonetheless, we may readily recover the outcome probabilities when ν is measured by $\mathcal{M} := \{e_k^{(0)}\}$:

$$\mathrm{P}_{\mathcal{S}}(\mathcal{M}=k) = e_k^{(0)}(\nu). \tag{7}$$

In the special case when ν *is* a member of the MCI-frame, $\nu = \omega_{j_{\mathbf{0}}}^{(0)}$, we have $\mathrm{P}_{\mathcal{S}}(\mathcal{M}=k) = \delta_{j_{\mathbf{0}}k}$.

To carry the d outcomes of the MCI-frame measurement spread from system $\mathcal{S}$, we assume that each environment system (labeled by $i \in \{1\ldots N\}$) contains an MCI-frame $\{\omega_j^{(i)}\}_{j=0,\ldots d-1}$, distinguished by some refined measurement $\{e_k^{(i)}\}$. Like qubits in quantum theory, the $\mathcal{E}_i$ are not necessarily standalone systems like single particles, but they can correspond to effective subsystems of larger environmental systems, picked out by the specific form of the interaction with $\mathcal{S}$. Let us briefly consider the simplest case with just a single environment, initially in the first state ω_0 of the frame $\{\omega_j^{(1)}\}$. Then, to exhibit the same operational behaviour as eq. (6) (via eq. (7)), the joint probability of any pair of outcomes j_0 and j_1 on $\mathcal{S}$ and $\mathcal{E}_1$ should satisfy

$$(e_{j_{\mathbf{0}}}^{(0)}\odot e_{j_1}^{(1)})\,[T(\nu\odot\omega_0)] = e_{j_{\mathbf{0}}}^{(0)}(\nu)\delta_{j_{\mathbf{0}},j_1}. \tag{8}$$

In this way, the distribution $\{e_{j_{\mathbf{0}}}(\nu)\}_{j_{\mathbf{0}}}$ is broadcast to the environment, as in eq. (5). Crucially, eq. (8) implies that the system and environment will agree on the outcome of $\mathcal{M}$ on $\mathcal{S}$. Moreover, the probabilities of such an outcome when directly measuring $\mathcal{S}$ are not affected by the transformation T, making T a member of the phase group [21] of this pointer measurement. This can be seen by summing eq. (8) over j_1.

The same operational desiderata easily extend to the more general case of N environmental systems, each now starting in an arbitrary frame state $\omega_{k_i}^{(i)}$. We summarize this with the following definitions:

Definition 5. *A composition of GPT system $\mathcal{S}$ and environments $\mathcal{E}_1,\ldots,\mathcal{E}_N$ is said to admit an* **ideal Darwinism process** *if*

(a) $\mathcal{S}$ has a d–state MCI-frame $\{\omega_k^{(0)}\}$, discriminated by a refined measurement $\{e_j^{(0)}\}$, and

(b) each $\mathcal{E}_i$ has a d-state MCI-frame $\{\omega_j^{(i)}\}$ discriminated by a refined measurement $\{e_j^{(i)}\}$, such that

(c) there exists a reversible ("fan-out") transformation $T \in \mathcal{T}_{\mathcal{S}\mathcal{E}_1\ldots\mathcal{E}_N}$ that satisfies

$$(e_{j_{\mathbf{0}}}^{(0)}\odot e_{j_1}^{(1)}\odot\ldots\odot e_{j_N}^{(N)})[T(\nu\odot\omega_{k_1}^{(1)}\odot\ldots\odot\omega_{k_N}^{(N)})] = \delta_{j_1,j_{\mathbf{0}}+k_1}\ldots\delta_{j_N,j_{\mathbf{0}}+k_N}e_{j_{\mathbf{0}}}^{(0)}(\nu) \tag{9}$$

for all $k_1,\ldots k_N, j_0, j_1,\ldots j_N$ and all $\nu\in\Omega_{\mathcal{S}}$, where addition is modulo d.

Definition 6. *If for a collection of MCI-frames $\{\omega_j^{(i)}\}$ that satisfy items (a) and (b) of definition 5, a reversible transformation $T \in \mathcal{T}_{\mathcal{S}\mathcal{E}_1\ldots\mathcal{E}_N}$ satisfies*

$$T(\omega_{j_{\mathbf{0}}}^{(0)}\odot\omega_{j_1}^{(1)}\odot\ldots\odot\omega_{j_N}^{(N)}) = \omega_{j_{\mathbf{0}}}^{(0)}\odot\omega_{j_{\mathbf{0}}+j_1}^{(1)}\odot\ldots\odot\omega_{j_{\mathbf{0}}+j_N}^{(N)}, \tag{10}$$

then we say that T **robustly spreads classical information**.

Definition 6 demands that the system and environment behave in some sense like classical information

registers: if, for example, $j_1 = \ldots = j_N = 0$, the transformation T copies the classical information in $\mathcal{S}$ to the environments, directly on the level of states. In quantum theory, such robust spreading of classical information is sufficient for Darwinism: the pointer basis of $\mathcal{S}$ spans the system's Hilbert space, and so eq. (10) implies eq. (9) due to the state vector linearity of unitary maps. More generally, Definitions 5 and 6 are equivalent in quantum theory, in the sense that ideal Quantum Darwinism processes are exactly those that robustly spread classical information.

However, this equivalence does not hold for arbitrary GPTs, since eq. (10) will not in general imply eq. (9). Even if definition 5 holds, definition 6 can put additional constraints on both the system and the environment. With respect to the system, one needs to consider the possibility of a T that preserves the statistics of $\{e_j^{(0)}\}$ on $\mathcal{S}$, but still changes the state of $\mathcal{S}$, even if $\mathcal{S}$ is prepared in one of the frame states $\omega_j^{(0)}$. This is impossible in quantum theory, since every rank-1 quantum projector $E_j^{(0)}$ has a unique normalized and pure state $\omega_j^{(0)}$ that satisfies $\mathrm{tr}\left(E_j^{(0)}\omega_j^{(0)}\right) = 1$. However, many GPT systems (such as gbits [18]) violate the analogous operational condition on MCI-frames, which can in some cases be traced back to the fact that GPTs need not obey the usual quantum uncertainty principles [22]. With respect to the environment, definition 6 precludes the possibility that T creates exotic correlations between the $\mathcal{E}_i$ while preserving the statistics of the product measurements $e_{j_1}^{(1)} \odot \ldots \odot e_{j_N}^{(N)}$.

Thus, definition 5 captures the essential features for ideal Darwinism on the operational level, while definition 6 further requests classical features from the frame states themselves.

B. Necessary features for Darwinism in GPTs

In QT, the fan-out gate (eq. (2)) can create entanglement whenever the system is not initialized to a pointer state. The first main results of this paper are to show that entanglement–creation is a necessary property of *any* generalized ideal Darwinism process. We begin by showing that preventing a Darwinism process from creating entangled states puts a very strong constraint on the theory.

Theorem 1. *Suppose that we have an ideal Darwinism process for which the fan-out transformation T maps separable states to separable states. Then, for every pure state $\nu \in \Omega_{\mathcal{S}}$, we have $e_i^{(0)}(\nu) = 0$ or $e_i^{(0)}(\nu) = 1$ for all i. That is, the system $\mathcal{S}$ cannot have pure states that do not lead to deterministic outcomes on measurement $\{e_i^{(0)}\}_i$.*

Remark. This conclusion is valid also for non-ideal Darwinism processes that, instead of definition 5, satisfy the weaker condition

$$\left(e_i \odot e_{j_1}^{(1)} \odot \ldots \odot e_{j_N}^{(N)}\right) T \left(\nu \odot \omega^{(1)} \odot \ldots \odot \omega^{(N)}\right) = e_i(\nu)\delta_{i,j_1}\delta_{i,j_2}\ldots\delta_{i,j_N} \quad (11)$$

for all states $\nu \in \Omega_{\mathcal{S}}$, where $\omega^{(1)}, \ldots, \omega^{(N)}$ is an *arbitrary* fixed set of pure states and the $\{e_i\}_i$ and $\{e_{j_i}^{(j)}\}_i$ are *arbitrary* fixed measurements (as opposed to MCI–frames and refined measurements).

Proof. Since T is a reversible transformation, it maps pure states to pure states. Hence, if it also preserves separability, then there are pure states $\varphi^{(0)}, \ldots, \varphi^{(N)}$ (which may all depend on ν) such that

$$T\left(\nu \odot \omega^{(1)} \odot \ldots \odot \omega^{(N)}\right) = \varphi^{(0)} \odot \varphi^{(1)} \odot \ldots \odot \varphi^{(N)}. \quad (12)$$

Since T satisfies eq. (11), we obtain

$$e_i(\nu)\delta_{i,j_1}\ldots\delta_{i,j_N} = e_i(\varphi^{(0)})e_{j_1}^{(1)}(\varphi^{(1)})\ldots e_{j_N}^{(N)}(\varphi^{(N)}). \quad (13)$$

Summing over all $j_1, \ldots, j_N$ yields $e_i(\nu) = e_i(\varphi^{(0)})$ for all i.

Now suppose that i^* is an outcome label such that $e_{i^*}(\varphi^{(0)}) = 0$, then $e_{i^*}(\nu) = 0$. On the other hand, consider the case that $e_{i^*}(\varphi^{(0)}) \neq 0$. If at least one of the j_k is different from i^*, then setting $i = i^*$ in eq. (13) yields

$$0 = \underbrace{e_{i^*}(\varphi^{(0)})}_{\neq 0} e_{j_1}^{(1)}(\varphi^{(1)})\ldots e_{j_N}^{(N)}(\varphi^{(N)}),$$

hence $e_{j_1}^{(1)}(\varphi^{(1)})\ldots e_{j_N}^{(N)}(\varphi^{(N)}) = 0$. But since $\sum_{j_1,\ldots,j_N} e_{j_1}^{(1)}(\varphi^{(1)})\ldots e_{j_N}^{(N)}(\varphi^{(N)}) = 1$, we must have $e_{i^*}^{(1)}(\varphi^{(1)})\ldots e_{i^*}^{(N)}(\varphi^{(N)}) = 1$, and so $e_{i^*}^{(j)}(\varphi^{(j)}) = 1$ for all j. Recalling eq. (13) we therefore see that $e_i(\nu) = 0$ for all $i \neq i^*$, and so $e_{i^*}(\nu) = 1$.

In summary, we obtain $e_{i^*}(\nu) \in \{0, 1\}$ for all i^*. □

Thus, for all GPT systems $\mathcal{S}$ that contain pure states on which the MCI-frame measurement gives nondeterministic outcomes, the corresponding ideal Darwinism processes (if they exist) must create entangled states. While this property will be satisfied for typical GPT systems, we cannot immediately conclude that a system satisfying $e(\nu) = 0$ or 1 must be classical. For instance, a GPT system with a cubic state space (i.e. gbits in a theory called "boxworld" [18]) and the full dual octahedral effect space will satisfy $e(\nu) = 0$ or 1 for every pair of pure state ν and pure effect e – but is evidentally nonclassical. However, as we shall see in following theorem, Darwinism in boxworld (among a wider class of theories) can be ruled out by another necessary condition: this time, on the measurements.

In particular, let us focus on GPT systems $\mathcal{S}$ that are non-classical in the following sense: in addition to

the refined measurement $e_0^{(0)}, ..., e_{d-1}^{(0)}$ that reads out the MCI-frame $\{\omega_k^{(0)}\}$, there is at least one other refined measurement $\tilde{e}_0^{(0)}, ..., \tilde{e}_{d-1}^{(0)}$ that is not just a relabelling of the measurement $\{e_j^{(0)}\}$, i.e. at least one of the $\tilde{e}_j^{(0)}$ is not equal to any of the $e_k^{(0)}$. (In quantum theory, this would correspond to projective measurements in different bases, with all projectors rank-one.)

Theorem 2. *Suppose that we have an ideal Darwinism process such that the system $\mathcal{S}$ is non-classical in the sense described above. Then the fan-out transformation T must map some pure product effects to entangled effects.*

Proof. It will be useful to use the notation $\|e\| := \max_{\omega\in\Omega} e(\omega)$ for effects e. Suppose that T maps all pure product effects to separable effects. Then, since T is reversible and preserves purity, Lemma 9 (appendix B) implies that T maps pure product effects to pure product effects. Hence, due to eq. (9), for every j_0 and for every $\mathbf{j} = (j_1, \ldots, j_N)$ there are effects $h_{j_0,\mathbf{j}}^{(0)}, \ldots, h_{j_0,\mathbf{j}}^{(N)}$ such that

$$\left(e_{j_0}^{(0)} \odot e_{j_1}^{(1)} \odot \ldots \odot e_{j_N}^{(N)}\right) T = h_{j_0,\mathbf{j}}^{(0)} \odot h_{j_0,\mathbf{j}}^{(1)} \odot \ldots \odot h_{j_0,\mathbf{j}}^{(N)}. \tag{14}$$

Due to multilinearity, we can move any multiplicative constant into the zeroth factor, and in this way choose the effects such that $\|h_{j_0,\mathbf{j}}^{(i)}\| = 1$ for all $i \in \{1, \ldots, N\}$. If we had $\|h_{j_0,\mathbf{j}}^{(0)}\| < 1$, then the right-hand side could never attain the value 1 on product states, but we know that it does due to definition 5. Thus, $\|h_{j_0,\mathbf{j}}^{(0)}\| = 1$.

Substituting eq. (14) into eq. (9) and noting that the result is valid for every state $\nu \in \Omega_{\mathcal{S}}$, we obtain

$$h_{j_0,\mathbf{j}}^{(0)}\, p_{j_0,\mathbf{j},\mathbf{k}} = \delta_{j_1,j_0+k_1} \ldots \delta_{j_N,j_0+k_N} e_{j_0}^{(0)},$$

where $p_{j_0,\mathbf{j},\mathbf{k}} := h_{j_0,\mathbf{j}}^{(1)}(\omega_{k_1}^{(1)}) \cdot \ldots \cdot h_{j_0,\mathbf{j}}^{(N)}(\omega_{k_N}^{(N)}) \geq 0$. The special case of $\mathbf{k} = \mathbf{j} - j_0 := (j_1 - j_0, \ldots, j_N - j_0)$ yields $e_{j_0}^{(0)} = p_{j_0,\mathbf{j},\mathbf{j}-j_0} h_{j_0,\mathbf{j}}^{(0)}$ But since $\|e_{j_0}^{(0)}\| = 1 = \|h_{j_0,\mathbf{j}}^{(0)}\|$, this implies that $h_{j_0,\mathbf{j}}^{(0)} = e_{j_0}^{(0)}$ for all $\mathbf{j}$.

Since $\mathcal{S}$ is non-classical, there is another refined measurement $\{\tilde{e}_j^{(0)}\}_j$ which is not just a relabelling (i.e. permutation) of $\{e_i^{(0)}\}_i$. Using again our assumption that T maps products of pure effects to product effects, we obtain

$$\left(\tilde{e}_{j_0}^{(0)} \odot e_{j_1}^{(1)} \odot \ldots \odot e_{j_N}^{(N)}\right) T = \tilde{h}_{j_0,\mathbf{j}}^{(0)} \odot \tilde{h}_{j_0,\mathbf{j}}^{(1)} \odot \ldots \odot \tilde{h}_{j_0,\mathbf{j}}^{(N)} \tag{15}$$

for some suitable effects $\tilde{h}_{j_0,\mathbf{j}}^{(0)}, \ldots, \tilde{h}_{j_0,\mathbf{j}}^{(N)}$. Again, we define the effects such that $\|\tilde{h}_{j_0,\mathbf{j}}^{(i)}\| = 1$ for all $i \in \{1, \ldots, N\}$ (the case $i = 0$ will be discussed later). Summing over j_0, using that $\sum_{j_0} \tilde{e}_{j_0}^{(0)} = u_{\mathcal{S}} = \sum_{j_0} e_{j_0}^{(0)}$, yields

$$\sum_{j_0} e_{j_0}^{(0)} \odot h_{j_0,\mathbf{j}}^{(1)} \odot \ldots \odot h_{j_0,\mathbf{j}}^{(N)} = \sum_{j_0} \tilde{h}_{j_0,\mathbf{j}}^{(0)} \odot \tilde{h}_{j_0,\mathbf{j}}^{(1)} \odot \ldots \odot \tilde{h}_{j_0,\mathbf{j}}^{(N)}. \tag{16}$$

Applying both sides to the product state $\nu \odot \omega_{k_1}^{(1)} \odot \ldots \odot \omega_{k_N}^{(N)}$ and recalling eq. (9), we obtain

$$\begin{aligned} &\sum_{j_0} \tilde{h}_{j_0,\mathbf{j}}^{(0)}(\nu) \tilde{h}_{j_0,\mathbf{j}}^{(1)}(\omega_{k_1}^{(1)}) \ldots \tilde{h}_{j_0,\mathbf{j}}^{(N)}(\omega_{k_N}^{(N)}) \\ &= \sum_{j_0} e_{j_0}^{(0)}(\nu) \delta_{j_1,j_0+k_1} \ldots \delta_{j_N,j_0+k_N}. \end{aligned} \tag{17}$$

So far, $\mathbf{j}$ and $\mathbf{k}$ are arbitrary, but now set $j_i := k_i + l$ for all i, where l is fixed (we abbreviate this by $\mathbf{j} = \mathbf{k} + l$). We obtain

$$e_l^{(0)}(\nu) = \sum_{j_0} q_{j_0,\mathbf{k},l} \tilde{h}_{j_0,\mathbf{k}+l}^{(0)}(\nu), \tag{18}$$

where $q_{j_0,\mathbf{k},l} := \tilde{h}_{j_0,\mathbf{k}+l}^{(1)}(\omega_{k_1}^{(1)}) \ldots \tilde{h}_{j_0,\mathbf{k}+l}^{(N)}(\omega_{k_N}^{(N)}) \in [0,1]$. Since this is true for all states $\nu \in \Omega_{\mathcal{S}}$, we may again drop the ν and read it as an equality between effects. Since $e_l^{(0)} \neq 0$, for every l and for every $\mathbf{k}$ there must be some j_0 such that $q_{j_0,\mathbf{k},l} \neq 0$. Since $e_l^{(0)}$ is pure, this implies that $e_l^{(0)} \propto \tilde{h}_{j_0,\mathbf{k}+l}^{(0)}$. Now fix an arbitrary $\mathbf{j}$, and consider the special case $\mathbf{k} := \mathbf{j} - l$. It follows that for all l, there exists at least one j_0 such that $e_l^{(0)}$ is a scalar multiple of $\tilde{h}_{j_0,\mathbf{j}}^{(0)}$. There are d different linearly independent $e_l^{(0)}$ (labelled by l), and there are d different $\tilde{h}_{j_0,\mathbf{j}}^{(0)}$, labelled by j_0. Thus, to every l there is a unique j_0 such that $e_l^{(0)} = q_{j_0,\mathbf{j}-l,l} \tilde{h}_{j_0,\mathbf{j}}^{(0)}$. We have

$$1 = \|e_l^{(0)}\| = \underbrace{q_{j_0,\mathbf{j}-l,l}}_{\leq 1} \underbrace{\|\tilde{h}_{j_0,\mathbf{j}}^{(0)}\|}_{\leq 1}, \tag{19}$$

hence $\|\tilde{h}_{j_0,\mathbf{j}}^{(0)}\| = 1$, and so $e_l^{(0)} = \tilde{h}_{j_0,\mathbf{j}}^{(0)}$. We can rephrase this as follows. For every $\mathbf{j}$ there is a permutation π of the indices such that $\tilde{h}_{j_0,\mathbf{j}}^{(0)} = e_{\pi(j_0)}^{(0)}$ for all j_0.

Now fix some $\mathbf{j}$. Let us return to eq. (16) and apply it to $\omega_{\pi(j_0)}^{(0)} \odot \omega$, where π is the permutation corresponding to $\mathbf{j}$, and ω is an arbitrary global state of the N environments. Using the identities that we have just derived and $e_{j_0}^{(0)}(\omega_i^{(0)}) = \delta_{j_0,i}$, we obtain

$$\tilde{h}_{j_0,\mathbf{j}}^{(1)} \odot \ldots \odot \tilde{h}_{j_0,\mathbf{j}}^{(N)} = h_{\pi(j_0),\mathbf{j}}^{(1)} \odot \ldots \odot h_{\pi(j_0),\mathbf{j}}^{(N)}. \tag{20}$$

Recalling eqs. (14) and (15), it follows that

$$\begin{aligned} \left(\tilde{e}_{j_0}^{(0)} \odot e_{j_1}^{(1)} \odot \ldots \odot e_{j_N}^{(N)}\right) T &= \tilde{h}_{j_0,\mathbf{j}}^{(0)} \odot \tilde{h}_{j_0,\mathbf{j}}^{(1)} \odot \ldots \odot \tilde{h}_{j_0,\mathbf{j}}^{(N)} \\ &= e_{\pi(j_0)}^{(0)} \odot h_{\pi(j_0),\mathbf{j}}^{(1)} \odot \ldots \odot h_{\pi(j_0),\mathbf{j}}^{(N)} \\ &= \left(e_{\pi(j_0)}^{(0)} \odot e_{j_1}^{(1)} \odot \ldots \odot e_{j_N}^{(N)}\right) T. \end{aligned} \tag{21}$$

Since T is reversible, the terms in the brackets must be identical. Consider the case $j_1 = j_2 = \ldots = j_N = 0$ and environment states $\varphi^{(1)}, \ldots, \varphi^{(N)}$ with $e_0^{(k)}(\varphi^{(k)}) = 1$ for all $k = 1, \ldots, N$. Then applying the above brackets to the product state $\nu \odot \varphi^{(1)} \odot \ldots \odot \varphi^{(N)}$ yields $\tilde{e}^{(0)}_{j_\mathbf{0}}(\nu) = e^{(0)}_{\pi(j_\mathbf{0})}(\nu)$. Since this is true for all $\nu \in \Omega_\mathcal{S}$, we obtain $\tilde{e}^{(0)}_{j_\mathbf{0}} = e^{(0)}_{\pi(j_\mathbf{0})}$. This contradicts our assumption that $\{\tilde{e}^{(0)}_j\}_j$ is not just a permutation of the $\{e^{(0)}_i\}_i$. □

Thus, a reversible transformation T that implements an ideal Darwinism process will create entangled effects. An important consequence is that GPTs without entangled effects, such as those constructed by taking the maximal tensor product in the context of tomographic locality, cannot admit such a process. In particular, this rules out Darwinism in boxworld [18] (a theory containing the aforementioned gbits) or any dichotomic maximally nonlocal theory. For these specific examples, one could also infer this from Refs. [40, 44], but here we have shown it without having to determine the complete structure of the reversible transformations.

Interestingly, entanglement for states is also needed in general physical theories if one imposes another condition of relevance for the classical limit: the existence of a decoherence map [23]. However, theories that have an ideal Darwinism process – and by our results need entangled states and measurements – may not contain such a decoherence map, as we shall show in section III D. Therefore, our results provide not only alternative proofs but are complementary to that of Richens et al. [23]: together they support the idea that this non-classical feature must be present for a locally non-classical theory to admit a meaningful classical limit.

C. Sufficient features for Darwinism in GPTs

Let us now determine sufficient conditions that guarantee that Quantum Darwinism can be generalized into a theory. In particular, we are interested in which operationally well-motivated postulates that have already appeared in the GPT literature can lead to such Darwinism. In this spirit, we will see how a framework that admits decoherence also admits Darwinism.

We will first determine sufficient structure in GPTs to allow for the robust spreading of classical information (in the manner of definition 6), before determining which additional postulates can be added to guarantee the existence of an ideal Darwinism process (definition 5) that additionally broadcasts classical information to the environment even when the system is not in a MCI–frame state.

Recall that both, the spreading of classical information and the ideal Darwinism processes, require the system to have an MCI-frame (playing the role of pointer states) that defines the classical information to be spread to the environment (definition 5(a)). Likewise the environments must admit MCI-frames on which to receive this classical information (definition 5(b)). Even though a theory admitting such frames may arguably be said to contain classical information (i.e. admitting "registers" that can encode the appropriate values), it may not generally admit all (or even any!) classical information processing – that is, there is no guarantee that the theory admits sufficient dynamics to satisfy definition 6. In the following, we will consider what physical characteristics *do* ensure that the theory has enough classical information processing power to implement a fan-out gate in the manner of eq. (10).

The first possible characteristic is to demand that composite systems satisfy *strong symmetry* [45]:

Definition 7. *A GPT system with group of reversible transformations $\mathcal{T}$ satisfies **strong symmetry (on states)** if for all $n \in \mathbb{N}$ and for all pairs of frames $\omega_1, ..., \omega_n$ and $\nu_1, ..., \nu_n$, there exists some $T \in \mathcal{T}$ with $T\omega_j = \nu_j$ for all j.*

Strong symmetry says that all ways of encoding classical information are computationally equivalent. In particular, it implies that classical reversible computation can be performed on the MCI-frames of system and environment: since the set of states $\omega_{j_\mathbf{0},\ldots,j_N} := \omega^{(0)}_{j_\mathbf{0}} \odot ... \odot \omega^{(N)}_{j_N}$ constitutes a frame of the composite system, strong symmetry implies that we can perform arbitrary classical reversible gates (and thus arbitrary permutations) of those frame elements. This immediately gives us the following result:

Lemma 3. *Consider GPT systems $\mathcal{S}, \mathcal{E}_1, \ldots, \mathcal{E}_N$ that carry d-outcome MCI-frames. Every composition $\mathcal{S}\mathcal{E}_1 \ldots \mathcal{E}_N$ that satisfies strong symmetry (on states) admits the robust spreading of classical information.*

While strong symmetry on states implies the robust spreading of classical information in the sense of definition 6, we do not know whether this property implies the existence of an ideal Darwinism process in the sense of definition 5. Interestingly, the existence of such a process follows if we consider a dual notion of strong symmetry on the *measurements*:

Definition 8. *A GPT system with group of reversible transformations $\mathcal{T}$ satisfies **strong symmetry (on effects)** if the following holds for all $n \in \mathbb{N}$: If $(e_1, \ldots, e_n)$ is a collection of pure effects that perfectly distinguishes some frame, and so is $(f_1, \ldots, f_n)$, then there exists a $T \in \mathcal{T}$ with $e_j = f_j \circ T$ for all j.*

If this property holds, we can show the following:

Lemma 4. *Consider again GPT systems $\mathcal{S}, \mathcal{E}_1, \ldots, \mathcal{E}_N$ that carry d-outcome MCI-frames. Every composition $\mathcal{S}\mathcal{E}_1 \ldots \mathcal{E}_N$ that satisfies strong symmetry (on effects) admits an ideal Darwinism process.*

Proof. The $e_{j_\mathbf{0},\ldots,j_N}$ are pure effects which perfectly distinguish the frame $\omega_{j_\mathbf{0},\ldots,j_N}$. Thus, strong symmetry on effects implies that there is some $T \in \mathcal{T}_{\mathcal{S}\mathcal{E}_1 \ldots \mathcal{E}_N}$ with

$$e_{j_\mathbf{0},j_1,\ldots,j_N} \circ T = e_{j_\mathbf{0},j_1-j_\mathbf{0},\ldots,j_N-j_\mathbf{0}}, \tag{22}$$

where subtraction is modulo d. One can check directly that this map T satisfies eq. (9). □

Thus, the version of Darwinism that is guaranteed to hold (according to definition 5 or 6) depends on whether we demand strong symmetry on the states or on the effects. Is there a way to guarantee it on both? Indeed, it turns out that the no-restriction hypothesis is sufficient for this:

Theorem 5. *Consider GPT systems $\mathcal{S}, \mathcal{E}_1, \ldots, \mathcal{E}_N$ that carry d-outcome MCI-frames. Every unrestricted composition $\mathcal{S}\mathcal{E}_1 \ldots \mathcal{E}_N$ that satisfies strong symmetry (on states) has a transformation $T \in \mathcal{T}_{\mathcal{S}\mathcal{E}_1 \ldots \mathcal{E}_N}$ that robustly spreads classical information* and *that generates an ideal Darwinism process.*

Proof. For unrestricted systems A with strong symmetry on states, it was shown in Ref. [39] that there is a particularly strong duality between states and effects: there is an inner product $\langle \cdot, \cdot \rangle$ on A such that frames $\omega_1, \ldots, \omega_n$ correspond to orthonormal systems, and the corresponding pure effects with $e_i(\omega_j) = \delta_{ij}$ must be given by $e_i(\omega) = \langle \omega_i, \omega \rangle$. Moreover, all $T \in \mathcal{T}_A$ are orthogonal with respect to this inner product. If $f_1, \ldots, f_n$ is any other collection of pure effects that distinguish a frame (say, $\nu_1, \ldots, \nu_n$), then strong symmetry on states says that there is some $T \in \mathcal{T}_A$ with $T\omega_j = \nu_j$, and so

$$e_j \circ T^{-1}(\omega) = \langle \omega_j, T^{-1}\omega \rangle = \langle T\omega_j, \omega \rangle = \langle \nu_j, \omega \rangle = f_j(\omega).$$

Consequently, A also satisfies strong symmetry on effects. Now, choose T as in eq. (22), then we already know that it generates an ideal Darwinism process. Moreover, we have just seen that T^{-1} maps the corresponding frame elements onto each other, i.e.

$$T^{-1}\omega_{j_\mathbf{0}, j_1, \ldots, j_N} = \omega_{j_\mathbf{0}, j_1 - j_\mathbf{0}, \ldots, j_N - j_\mathbf{0}}.$$

Applying T to both sides shows that T robustly spreads classical information in the sense of definition 6. □

A second path to this spreading of classical information arises from decoherence theory. In quantum theory, decoherence plays an important role in Quantum Darwinism by explaining in some sense why we see classical probabilities instead of superposition states. Recently, a decoherence formalism for GPTs was developed [23], and we shall here see that it enables Darwinism in GPTs as well. We adapt the decoherence formalism of Richens et al. [23] to our setting:

Definition 9 (Decoherence maps)**.** *Consider any GPT system A. A linear map $D : A \to A$ is called a* decoherence map *if the following properties hold:*

1. *The image of A_+ under D is isomorphic to a classical state space, i.e. there exists a frame $\omega_0, \ldots, \omega_{d-1} \in \Omega_A$ such that $D(\Omega_A) = \mathrm{conv}\{\omega_0, \ldots, \omega_{d-1}\}$ (i.e. the convex hull of the $\{\omega_i\}$). Consequently, D is normalization-preserving, i.e. $u_A \circ D = u_A$.*

2. *D is idempotent, i.e. $D \circ D = D$.*

3. *For every classical reversible transformation $T_\mathrm{C} : D(A) \to D(A)$ there is a reversible transformation $T \in \mathcal{T}_A$ that implements T_C, i.e. $T(\omega) = T_C(\omega)$ for all $\omega \in D(A_+)$. Not only does this map T preserve the classical state space $D(A_+)$, but it also preserves the corresponding classical effect space $E_A \circ D$.*

Furthermore, if we have a composite GPT system $A = A_1 A_2 \ldots A_N$ with decoherence maps $D_1, \ldots, D_N$,

4. *A has a decoherence map $D_{1\ldots N}$ that acts as*

$$D_{1\ldots N}(\nu_1 \odot \ldots \odot \nu_N) = D_1(\nu_1) \odot \ldots \odot D_N(\nu_N).$$

Richens et al. [23] additionally assume that D is physically implementable, but we do not assume this here.

In the following, we will need a simple property of decoherence maps:

Lemma 6. *Consider a GPT system A with decoherence map D, and T any reversible transformation that implements some classical transformation in the sense of definition 9 item 3. Then $DT = TD$.*

Proof. Let $e \in E_A$ and $\varphi \in A_+$, then $f := e \circ D$ is an element of the classical effect space $E_A \circ D$, and so is $f' := f \circ T$, hence $f' = f' \circ D$. Thus, we have

$$e \circ DT\varphi = f \circ T\varphi = f'(\varphi) = f' \circ D\varphi = e \circ DTD\varphi.$$

Since A_+ and E_A span A and A^*, respectively, it follows that $DT = DTD$. But T preserves $D(A) = \mathrm{span}(D(A_+))$, hence $DTD = TD$. □

In analogy with how quantum systems decohere to mixtures of pointer states, it is natural to consider Darwinism for frames that can result from decoherence processes.

Definition 10. *Consider any GPT system A. We say that an MCI-frame $\{\omega_i\} \subset \Omega_A$ together with a corresponding refined measurement $\{e_i\} \subset E_A$* arises from decoherence *if there is a decoherence map $D : A \to A$ such that $D(A_+) = \mathrm{cone}\{\omega_i\}$ and $E_A \circ D = \mathrm{cone}\{e_i\}$.*

In this definition, $\mathrm{cone}\{\omega_i\}$ denotes the set of non-negative linear combinations of the ω_i, i.e. the convex cone of unnormalized states generated by the MCI-frame (similarly for the $\{e_i\}$).

Let $\{\omega_j^{(0)}\}_{j=0}^{d-1}$ be an MCI-frame of the main system $\mathcal{S}$ that arises from decoherence map D_0, and similarly let $\{\omega_j^{(i)}\}_{j=0}^{d-1}$, $i = 1, \ldots, N$, be MCI-frames of the environmental systems $\mathcal{E}_1, \ldots, \mathcal{E}_N$ that arise from decoherence maps $D_1, \ldots, D_N$. Then requirement 4 of definition 9 implies that there is a decoherence map $D_{0\ldots N}$ with

$$D_{0\ldots N}(\omega_{j_\mathbf{0}}^{(0)} \odot \ldots \odot \omega_{j_N}^{(N)}) = D_0(\omega_{j_\mathbf{0}}^{(0)}) \odot \ldots \odot D_N(\omega_{j_N}^{(N)}).$$

Since each D_i is a projection map and since every $\omega_j^{(i)}$ is in its image, we have $D_i(\omega_j^{(i)}) = \omega_j^{(i)}$, and hence

$$D_{0\ldots N}(\omega_{j_\mathbf{0}}^{(0)} \odot \ldots \odot \omega_{j_N}^{(N)}) = \omega_{j_\mathbf{0}}^{(0)} \odot \ldots \odot \omega_{j_N}^{(N)}.$$

Requirement 3 for decoherence maps implies that the classical transformation defined by eq. (10) (a particular permutation of the classical pure states) is implemented as a reversible transformation $T \in \mathcal{T}_A$ on the composite GPT system $A := \mathcal{S}\mathcal{E}_1 \ldots \mathcal{E}_N$. This transformation hence robustly spreads classical information in the sense of definition 6.

Furthermore, consider any state $\nu \in \Omega_\mathcal{S}$, and let $\nu_0 := D_0\nu$. Since the MCI-frame of $\mathcal{S}$ arises from D_0, there is a convex decomposition $\nu_0 = \sum_{i=0}^{d-1} \lambda_i \omega_i^{(0)}$ with $\lambda_i \geq 0$, $\sum_{i=0}^{d-1} \lambda_i = 1$. Using lemma 6, we thus obtain

$$\begin{aligned}
&(e_{j_\mathbf{0}}^{(0)} \odot e_{j_1}^{(1)} \odot \ldots \odot e_{j_N}^{(N)}) T(\nu \odot \omega_{k_1}^{(1)} \odot \ldots \odot \omega_{k_N}^{(N)}) \\
&\quad = e_{j_\mathbf{0},\ldots,j_N} \circ D_{0\ldots N} T(\nu \odot \omega_{k_1}^{(1)} \odot \ldots \odot \omega_{k_N}^{(N)}) \\
&\quad = e_{j_\mathbf{0},\ldots,j_N} \circ T D_{0\ldots N}(\nu \odot \omega_{k_1}^{(1)} \odot \ldots \odot \omega_{k_N}^{(N)}) \\
&\quad\quad = e_{j_\mathbf{0},\ldots,j_N} \circ T(\nu_0 \odot \omega_{k_1}^{(1)} \odot \ldots \odot \omega_{k_N}^{(N)}) \\
&\quad = \sum_{i=0}^{d-1} \lambda_i e_{j_\mathbf{0},\ldots,j_N} \circ T(\omega_i^{(0)} \odot \omega_{k_1}^{(1)} \odot \ldots \odot \omega_{k_N}^{(N)}) \\
&\quad\quad = \sum_{i=0}^{d-1} \lambda_i e_{j_\mathbf{0},\ldots,j_N}(\omega_{i,i+k_1,\ldots,i+k_N}) \\
&\quad\quad\quad = \lambda_{j_\mathbf{0}} \delta_{j_1,j_\mathbf{0}+k_1} \ldots \delta_{j_N,j_\mathbf{0}+k_N}.
\end{aligned}$$

Furthermore, $e_{j_\mathbf{0}}^{(0)}(\nu) = e_{j_\mathbf{0}}^{(0)} \circ D_0(\nu) = \lambda_0$. This proves that T generates an ideal Darwinism process.

We summarize our findings in the following theorem:

Theorem 7. *Consider a composition $\mathcal{S}\mathcal{E}_1 \ldots \mathcal{E}_N$ of GPT systems $\mathcal{S}, \mathcal{E}_1, \ldots, \mathcal{E}_N$ that carry d-outcome MCI-frames arising from decoherence. This composite system admits a transformation $T \in \mathcal{T}_{\mathcal{S}\mathcal{E}_1\ldots\mathcal{E}_N}$ that robustly spreads classical information* and *that generates an ideal Darwinism process.*

Composite systems in quantum theory are unrestricted and satisfy strong symmetry (on states and effects). Furthermore, they admit MCI-frames arising from decoherence in the way specified above. Thus, the existence of an ideal Darwinism process and the robust spreading of classical information in quantum theory follow both as special cases of theorem 5 and theorem 7.

D. Darwinism in Spekkens' Toy Model

If one identifies too many specific restrictions on a GPT, it raises the natural question: "is quantum theory the only physical theory that allows for Darwinism?" We answer this in the negative by providing an example that admits Darwinism, but is not quantum theory: Spekkens' Toy Model (STM) [27].

STM satisfies many of the same restrictions as quantum theory, such as no-signalling and no-cloning, and emulates many quantum behaviours such as complementary measurements, interference, entanglement (and monogamy thereof), and teleportation [27]. Despite this, it is very different from quantum theory: both mathematically and conceptually, since at its core it is a classical hidden-variable model. What enables this quantum-like behaviour is that the states of maximum knowledge of the system are subject to the *epistemic restriction* that one knows only half of the possible information about the hidden *ontic* variable, along with a measurement-update rule that ensures that this restriction is maintained even when one makes sequential measurements on the system.

A more detailed description of STM and its extension into the GPT framework is given in appendix C. For now, it suffices to remark that the composition of such systems is achieved by composing the underlying hidden classical variable (i.e. by Cartesian product) and applying the epistemic restriction to both the composite system and every subsystem thereof.

As observed by Pusey [46] (and recounted in appendix C 3), the states within STM may be treated very similarly to the stabilizer subset of quantum theory (for a single system, the state spaces are isomorphic). In particular, a single elementary STM system admits three "toy observables" X, Y and Z which act on the state to produce outputs $+1$ or -1 – and there is one pure state for each of these six possibilities ($|x\pm\rangle, |y\pm\rangle, |z\pm\rangle$) and no other pure states. When the "wrong" observable acts on a pure state (e.g. acting on $|z+\rangle$ with X), outcomes $+1$ and -1 occur with equal probability. In this language, one can define the CNOT analogue for two STM bits "control" C and "target" T:

$$\begin{aligned}
\text{CNOT}: &X_C \mapsto X_C X_T,\ X_T \mapsto X_T, \\
&Z_C \mapsto Z_C,\ Z_T \mapsto Z_C Z_T.
\end{aligned}$$

This can be read as, e.g. $X_C \mapsto X_C X_T$, "The product of the observation of X on C and X on T after the transformation CNOT yields the same outcome statistics as the observation X on C before the transformation."

With this shorthand, we hence specify our candidate for an ideal Darwinism process from main system S onto multiple environments $E_1, ..., E_N$:

$$\begin{aligned}
\text{FAN}: \quad &X_S \mapsto X_S X_{E_1} ... X_{E_n}, \quad \forall k : X_{E_k} \mapsto X_{E_k}, \\
&Z_S \mapsto Z_S, \quad \forall k : Z_{E_k} \mapsto Z_S Z_{E_k}.
\end{aligned} \tag{23}$$

The validity of this, as a transformation in STM, can be verified in one of two ways: the first is to consider a direct implementation of this as a series of pairwise CNOT gates (in the manner of fig. 1), reasoning (e.g. via category theory [47]) that such composition is permissible. The second way is to note that this map is admissible as a transformation on an analogously defined N-bit quantum stabilizer system, and then use the result of Pusey [46] to infer that this makes FAN a valid STM transformation.

Thus, it remains to verify that such a transformation indeed achieves the desired ideal Darwinistic behaviour. Suppose we have an initial state of the form $|\psi\rangle_S \otimes |+z\rangle^{\otimes n}_{E_1 \dots E_N}$ where $|\psi\rangle_S$ is some arbitrary pure STM bit state of the main system, and $|+z\rangle$ corresponds to the state that always gives output $+1$ when measured by toy observable Z. As for each k, FAN maps Z_{E_k} to $Z_S Z_{E_k}$, the final state will always have result $+1$ for joint measurements of $Z_S Z_{E_k}$ – mandating that the results of Z_S and Z_{E_k} are perfectly correlated. (In the case E_j starts at $|-z\rangle$, anti-correlation is established.) Therefore the fan-out results in all observers seeing the same outcome as made on the original system.

Our other requirement for Darwinism is that the outcome probability of Z_S is not changed, and this is also explicitly given by the rule in the map $Z_S \mapsto Z_S$. In particular $|z\pm\rangle$ are the only pure states that have non-zero expectation value for the observable Z, and the map does not take any state of main system stabilized by another observable (i.e. X or Y) to any state stabilized by an expression containing Z. As such, since S can only be in one of these possibilities (or convex combination thereof in the GPT extension) this implies that the statistics of Z_S remain unchanged. We summarize this with our final theorem of the article:

Theorem 8 (STM admits an ideal Darwinism process)**.** *The FAN operation specified in eq.* (23) *implements an ideal Darwinism process, as per definition 5.*

We conclude this section with some remarks on the implications of this example to the theorems of this paper. First, in terms of necessary conditions: STM is nonclassical in the sense that there are more than one set of sufficiently different refined measurements (recall section III B), and indeed also STM has entangled effects as mandated by Theorem 2. Although requiring entangled effects in a non-classical setting, we can further conclude (by counterexample) that the stronger condition of violating of Bell inequalities (see e.g. [48]) is *not* necessary since STM does not violate these. A similar conclusion follows for *contextuality*, which is not present in STM [27] and thus shown to be unnecessary for Darwinism.

Secondly, in terms of the sufficient conditions, STM neither admits a decoherence map, nor is it strongly symmetric (as we show in appendix C 4). This illustrates that the sufficient conditions are not tight – they enable the fan-out dynamic by mandating the existence of *all* classical dynamics within the theory. However, the fan-out operation can be admitted without requiring universal classical computation – indeed, as above for STM, or existing as a member of the (non-universal [49]) Clifford group in the case of quantum stabilizers.

IV. CONCLUSIONS

Quantum Darwinism provides a mechanism through which crucial aspects of classicality can be understood to emerge in the quantum domain [1–5]. In this article, we generalized an ideal notion of Darwinism, where maximal classical information is perfectly broadcast to an environment split into fractions, to the framework of GPTs. We showed that entanglement, in both states and measurements, is a necessary feature for such a process to be present in generalized theories, and demonstrates that some important physical principles – like strong symmetry and decoherence – provide sufficient structure to admit Darwinism. Finally, we described a mechanism for Darwinism in Spekkens Toy Model, showing that such broadcasting of classical information is not unique to quantum theory.

Our results show that objectivity may arise through a Darwinism process in non-classical theories other than quantum – adding to the results of Scandolo et al. [50], which analyzed objectivity through State Spectrum Broadcast in GPTs. Complementing a previous result on decoherence [23], our work also shows the important role of entanglement to allow for emergence of classicality, suggesting the counterintuitive principle that locally non-classical theories must also allow for shared non-classicality to allow for the emergence of classical objectivity. In addition, our results show that strongly symmetric and unrestricted GPTs – that is, those endowed with sufficient structure to allow for reversible classical computation and the encoding and decoding of classical information – have sufficient structure for Darwinism to be present.

Finally, although this work has been presented with a focus on the origins of classical limits, our results also have a bearing on the general foundations of computation [51, 52]. The Darwinism–enabling fan-out transformation (eq. (1)) has its origins in classical logic circuits, connecting the output of one logic gate to the input of many others, and its quantum analogue plays a role in the design of quantum neural networks [53]. The conclusions of this article therefore imply that such computation also necessitates the existence of entanglement, if the theory is not strictly classical – meanwhile identifying potential sufficient structure (e.g. no-restriction and strong symmetry) to guarantee that such computation can be performed.

ACKNOWLEDGMENTS

RDB acknowledges funding by São Paulo Research Foundation – FAPESP, through scholarships no. 2016/24162-8 and no. 2019/02221-0. RDB is also thankful to Marcelo Terra Cunha for insightful ideas and discussions, and to IQOQI for the hospitality during the time as guest researcher. MK acknowledges the support of the Vienna Doctoral School in Physics (VDSP) and the support of the Austrian Science Fund (FWF) through the Doctoral Programme CoQuS. MK, AJPG and MPM thank the Foundational Questions Institute and Fetzer Franklin Fund, a donor advised fund of Silicon Valley Community Foundation, for support via

grant number FQXi-RFP-1815. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

[1] W. H. Zurek. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.*, 75(3): 715–775, 2003. doi:10.1103/revmodphys.75.715.

[2] W. H. Zurek. Relative States and the Environment: Einselection, Envariance, Quantum Darwinism, and the Existential Interpretation. *Pre-print*, arXiv:0707.2832, 2007. URL `https://arxiv.org/abs/0707.2832`.

[3] W. H. Zurek. Quantum Darwinism. *Nature Physics*, 5 (3):181–188, 2009. doi:10.1038/nphys1202.

[4] F. G. S. L. Brandão, M. Piani, and P. Horodecki. Generic emergence of classical features in quantum darwinism. *Nature Communications*, 6(1), 2015. doi: 10.1038/ncomms8908.

[5] P. A. Knott, T. Tufarelli, M. Piani, and G. Adesso. Generic emergence of objectivity of observables in infinite dimensions. *Phys. Rev. Lett.*, 121:160401, 2018. doi:10.1103/PhysRevLett.121.160401.

[6] R. Horodecki, J. K. Korbicz, and P. Horodecki. Quantum origins of objectivity. *Phys. Rev. A*, 91(3):032122, 2015. doi:10.1103/PhysRevA.91.032122.

[7] T. P. Le and A. Olaya-Castro. Objectivity (or lack thereof): Comparison between predictions of quantum darwinism and spectrum broadcast structure. *Phys. Rev. A*, 98(3), 2018. doi:10.1103/physreva.98.032103.

[8] R. Blume-Kohout and W. H. Zurek. Quantum darwinism in quantum brownian motion. *Phys. Rev. Lett.*, 101: 240405, 2008. doi:10.1103/PhysRevLett.101.240405.

[9] R. Blume-Kohout and W. H. Zurek. A simple example of "quantum darwinism": Redundant information storage in many-spin environments. *Foundations of Physics*, 35(11):1857–1876, 2005. doi:10.1007/s10701-005-7352-5.

[10] C. J. Riedel and W. H. Zurek. Quantum darwinism in an everyday environment: Huge redundancy in scattered photons. *Phys. Rev. Lett.*, 105:020404, 2010. doi: 10.1103/PhysRevLett.105.020404.

[11] M. Zwolak, H. T. Quan, and W. H. Zurek. Redundant imprinting of information in nonideal environments: Objective reality via a noisy channel. *Phys. Rev. A*, 81:062110, 2010. doi:10.1103/PhysRevA.81.062110.

[12] M. Zwolak, H. T. Quan, and W. H. Zurek. Quantum darwinism in a mixed environment. *Phys. Rev. Lett.*, 103:110402, 2009. doi: 10.1103/PhysRevLett.103.110402.

[13] P. A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, Oxford, 4th edition, 1958.

[14] J. J. Sakurai and J. Napolitano. *Modern Quantum Mechanics*. Addison-Wesley, 2nd edition, 2011.

[15] M. Schlosshauer. Decoherence, the measurement problem, and interpretations of quantum mechanics. *Rev. Mod. Phys.*, 76(4):1267–1305, 2004. doi: 10.1103/RevModPhys.76.1267.

[16] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, 1982. doi: 10.1038/299802a0.

[17] L. Hardy. Quantum Theory From Five Reasonable Axioms. *Pre-print*, arXiv:quant-ph/0101012, 2001. URL `https://arxiv.org/abs/quant-ph/0101012`.

[18] J. Barrett. Information processing in generalized probabilistic theories. *Phys. Rev. A*, 75:032304, 2007. doi: 10.1103/PhysRevA.75.032304.

[19] H. Barnum, J. Barrett, M. Leifer, and A. Wilce. Teleportation in General Probabilistic Theories. *Pre-print*, arXiv:0805.3553, 2008. URL `https://arxiv.org/abs/0805.3553`.

[20] H. Barnum, J. Barrett, M. Leifer, and A. Wilce. Generalized no-broadcasting theorem. *Phys. Rev. Lett.*, 99 (24), 2007. doi:10.1103/physrevlett.99.240501.

[21] A. J. P. Garner, O. C. O. Dahlsten, Y. Nakata, M. Murao, and V. Vedral. A framework for phase and interference in generalized probabilistic theories. *New J. Phys.*, 15(9):093044, 2013. doi:10.1088/1367-2630/15/9/093044.

[22] O. C. O. Dahlsten, A. J. P. Garner, and V. Vedral. The uncertainty principle enables non-classical dynamics in an interferometer. *Nature Communications*, 5(4592), 2014. doi:10.1038/ncomms5592.

[23] J. G. Richens, J. H. Selby, and S. W. Al-Safi. Entanglement is necessary for emergent classicality in all physical theories. *Phys. Rev. Lett.*, 119:080503, 2017. doi: 10.1103/PhysRevLett.119.080503.

[24] G. Chiribella, Giacomo M. D'Ariano, and P. Perinotti. Probabilistic theories with purification. *Phys. Rev. A*, 81(6), 2010. doi:10.1103/physreva.81.062348.

[25] P. Janotta and R. Lal. Generalized probabilistic theories without the no-restriction hypothesis. *Phys. Rev. A*, 87 (5), 2013. doi:10.1103/physreva.87.052131.

[26] H. Barnum, M. P. Müller, and C. Ududec. Higher-order interference and single-system postulates characterizing quantum theory. *New J. Phys.*, 16(12):123029, 2014. doi:10.1088/1367-2630/16/12/123029.

[27] R. W. Spekkens. Evidence for the epistemic view of quantum states: A toy theory. *Phys. Rev. A*, 75(3): 32110, 2007. doi:10.1103/PhysRevA.75.032110.

[28] A. S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Probl. Peredachi Inf.*, 9(3):177–183, 1973.

[29] M. Zwolak, C. J. Riedel, and W. H. Zurek. Amplification, decoherence and the acquisition of information by spin environments. *Scientific Reports*, 6(1):25277, 2016. doi:10.1038/srep25277.

[30] T. K. Unden, D. Louzon, M. Zwolak, W. H. Zurek, and F. Jelezko. Revealing the emergence of classicality using nitrogen-vacancy centers. *Phys. Rev. Lett.*, 123:140402, 2019. doi:10.1103/PhysRevLett.123.140402.

[31] M. A. Ciampini, G. Pinna, P. Mataloni, and M. Paternostro. Experimental signature of quantum darwinism in photonic cluster states. *Phys. Rev. A*, 98:020101, 2018. doi:10.1103/PhysRevA.98.020101.

[32] M. D. Mazurek, M. F. Pusey, K. J. Resch, and R. W. Spekkens. Experimentally bounding deviations from

quantum theory in the landscape of generalized probabilistic theories. *Pre-print*, arXiv:1710.05948, 2017. URL `https://arxiv.org/abs/1710.05948`.

[33] G. Chiribella, G. M. D'Ariano, and P. Perinotti. Informational derivation of quantum theory. *Phys. Rev. A*, 84(1):012311, 2011. doi:10.1103/PhysRevA.84.012311.

[34] Ll. Masanes and M. P. Müller. A derivation of quantum theory from physical requirements. *New J. Phys.*, 13 (6):063001, 2011. doi:10.1088/1367-2630/13/6/063001.

[35] M. P. Müller and Ll. Masanes. Three-dimensionality of space and the quantum bit: an information-theoretic approach. *New J. Phys.*, 15(5):053040, 2013. doi: 10.1088/1367-2630/15/5/053040.

[36] M. P. Müller. Probabilistic theories and reconstructions of quantum theory. *Les Houches 2019 lecture notes*, arXiv:2011.01286, 2020. URL `https://arxiv.org/abs/2011.01286`.

[37] R. Webster. *Convexity*. Oxford University Press, Oxford, 1994. ISBN 0-19-853147-8.

[38] G. Chiribella and C. M. Scandolo. Operational axioms for diagonalizing states. *Electronic Proceedings in Theoretical Computer Science*, 195:96–115, 2015. doi: 10.4204/EPTCS.195.8.

[39] M. P. Müller and C. Ududec. Structure of reversible computation determines the self-duality of quantum theory. *Phys. Rev. Lett.*, 108:130401, 2012. doi: 10.1103/PhysRevLett.108.130401.

[40] S. W. Al-Safi and J. Richens. Reversibility and the structure of the local state space. *New J. Phys.*, 17(12): 123001, 2015. doi:10.1088/1367-2630/17/12/123001.

[41] S. Massar and M. K. Patra. Information and communication in polygon theories. *Phys. Rev. A*, 89:052124, 2014. doi:10.1103/PhysRevA.89.052124.

[42] P. Janotta and H. Hinrichsen. Generalized probability theories: what determines the structure of quantum theory? *Journal of Physics A: Mathematical and Theoretical*, 47(32):323001, 2014. doi:10.1088/1751-8113/47/32/323001.

[43] B. Coecke and C. Heunen. Pictures of complete positivity in arbitrary dimension. *Information and Computation*, 250:50–58, 2016. doi:10.1016/j.ic.2016.02.007.

[44] D. Gross, M. P. Müller, R. Colbeck, and O. C. O. Dahlsten. All reversible dynamics in maximally nonlocal theories are trivial. *Phys. Rev. Lett.*, 104:080402, 2010. doi:10.1103/PhysRevLett.104.080402.

[45] H. Barnum, M. P. Müller, and C. Ududec. Higher-order interference and single-system postulates characterizing quantum theory. *New J. Phys.*, 16(12):123029, 2014. doi:10.1088/1367-2630/16/12/123029.

[46] M. F. Pusey. Stabilizer Notation for Spekkens' Toy Theory. *Foundations of Physics*, 42(5):688–708, 2012. doi: 10.1007/s10701-012-9639-7.

[47] B. Coecke, B. Edwards, and R. W. Spekkens. Phase Groups and the Origin of Non-locality for Qubits. *Electronic Notes in Theoretical Computer Science*, 270(2): 15–36, 2011. doi:10.1016/j.entcs.2011.01.021.

[48] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. *Reviews of Modern Physics*, 86(2):419–478, 2014. doi:10.1103/RevModPhys.86.419.

[49] D. Gottesman. The Heisenberg Representation of Quantum Computers. In *Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, pages 32–43, 1999. URL `https://arxiv.org/abs/quant-ph/9807006`.

[50] C. M. Scandolo, R. Salazar, J. K. Korbicz, and P. Horodecki. The origin of objectivity in all fundamental causal theories. *Pre-print*, arXiv:1805.12126, 2018. URL `https://arxiv.org/abs/1805.12126`.

[51] C. M. Lee and J. Barrett. Computation in generalised probabilisitic theories. *New J. Phys.*, 17(17), 2015. doi: 10.1088/1367-2630/17/8/083001.

[52] A. J. P. Garner. Interferometric Computation Beyond Quantum Theory. *Foundations of Physics*, 2018. doi: 10.1007/s10701-018-0142-7.

[53] K. H. Wan, O. C. O. Dahlsten, H. Kristjánsson, R. Gardner, and M. S. Kim. Quantum generalisation of feedforward neural networks. *npj Quantum Information*, 3(1), 2017. doi:10.1038/s41534-017-0032-4.

[54] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner. Limits on nonlocal correlations from the structure of the local state space. *New J. Phys.*, 13(6):063024, 2011. doi:10.1088/1367-2630/13/6/063024.

[55] Ll. Masanes, M. P. Müller, R. Augusiak, and D. Pérez-García. Existence of an information unit as a postulate of quantum theory. *Proceedings of the National Academy of Sciences of the United States of America*, 110(41):16373–16377, 2013. doi: 10.1073/pnas.1304884110.

[56] L. Hardy. Disentangling nonlocality and teleportation. *Pre-print*, arXiv:quant-ph/9906123, 1999. URL `https://arxiv.org/abs/quant-ph/9906123`.

[57] A. J. P. Garner. *Phase and interference phenomena in generalised probabilistic theories*. PhD thesis, University of Oxford, 2015. URL `https://ora.ox.ac.uk/objects/uuid:c0017faf-cbe0-4365-a1ff-080fa031d006`.

[58] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000. ISBN 0521635039.

APPENDIX

Appendix A: The pentagon state space

We present an example of a state space [41] (brought to our attention in Janotta et al. [54]) without an MCI-frame (definition 1), and illustrate its counterintuitive properties.

Example 3 (Pentagon state space)**.** *Consider a GPT system with states in $A = \mathbb{R}^3$ such that Ω_A is a regular pentagon (with pure states being the vertices), and with a dual space of effects E_A subject to the no-restriction hypothesis. Such a system admits a* self-dual *identification between A_+ and E_A in the following sense: for each vertex ω_j, there is a unique related effect $e_j \in E_A$ with $e_j \leq u_A$ such that $e_j(\nu) = 1 \Rightarrow \nu = \omega_j$; that is, these effects are in one-to-one correspondence with the vertices – and those are exactly the refined effects.*

Let us label the vertices clockwise. The maximal frame is of size two: any pair of vertices whose absolute difference between indices is 2 (modulo 5) lie on "opposite" sides of the pentagon, and form such a

frame. Then, both $\{\omega^0, \omega^2\}$ and $\{\omega^0, \omega^3\}$ are maximal frames; their states are distinguished, for example, by $M_1 = \{e_0, u_A - e_0\}$. However $u_A - e_0$ is *not refined*: we have that $u_A - e_0 = \alpha(e_2 + e_3)$ with $\alpha > (1/2)$, so one could also perform the refined measurement $M_2 = \{e_0, \alpha e_2, \alpha e_3\}$ to distinguish the states in each frame. (In this case, the equation $e_i(\omega^k) = \delta_{ik}$ is not valid, but the idea of single shot measurements allowing for the detection of the states still holds: if outcome αe_2 or αe_3 is obtained, one knows that ω^0 was not prepared.) A similar conclusion applies to every maximal frame in this theory: they all fail to be MCI-frames. All unrestricted GPT systems built from regular polygon state spaces with an odd number of vertices also lack MCI-frames.

Suppose someone is promised to receive, with probability $\mathrm{P}(i)$, the state ω^i from one of the frames $\{\omega^0, \omega^2\}$ or $\{\omega^0, \omega^3\}$ and should guess the value of i. Then, the probability of success when using measurement $M_1 = \{e_0, u_A - e_0\}$ is given by

$$p^{M_1}_{\text{success}} = \mathrm{P}(0) + (1/2)[\mathrm{P}(2) + \mathrm{P}(3)],$$

since one can always guess correctly if the outcome related to e_0 clicks but must make a random guess for $i = 2$ or $i = 3$ if the other outcome clicks. However, by using $M_2 = \{e_0, \alpha e_2, \alpha e_3\}$ one has

$$p^{M_2}_{\text{success}} = \mathrm{P}(0) + \alpha[\mathrm{P}(2) + \mathrm{P}(3)] > p^{M_1}_{\text{success}},$$

since $\alpha > 1/2$. We see that the refined measurement M_2 allows for a higher probability of distinguishing between a set of states which is larger than the maximal frames. In other words, the refined measurement M_2 can distinguish slightly more than 1 bit, even though the maximal frame has size 2 and this measurement M_2 coarse-grains to the distinguishing measurement M_1. If one understands coarse-graining as erasing of classical information, $p^{M_2}_{\text{success}} > p^{M_1}_{\text{success}}$ suggests that there was more classical information available than can be encoded onto a maximal frame. Such a phenomenon occurs for every unrestricted GPT built from a polygon state space with an odd number of vertices (see also Massar and Patra [41]). This difference between the amount of classical information that can be encoded into a GPT system and the size of a maximal frame is a violation of a principle that has been called "No Simultaneous Encoding" [55]. By explicitly only allowing MCI–frames (definition 1) to characterize the classical information to be spread by an ideal Darwinism process (definition 5), we ensure that no such over–encoding occurs in the systems considered in this article.

Appendix B: Pure separable effects

Lemma 9. *A pure effect is separable if and only if it is a product of pure effects.*

Proof. Only one direction is non-trivial: suppose that the effect $e^{1,2,\ldots,N}$ is separable, then it can be written

$$e^{1,2,\ldots,N} = \sum_i e_i^{(1)} \odot \ldots \odot e_i^{(N)} \tag{B1}$$

where the $e_i^{(j)}$ are suitable local effects. Since $e^{1,2,\ldots,N}$ is pure, we must have $e_i^{(1)} \odot \ldots \odot e_i^{(N)} \propto e^{1,2,\ldots,N}$ for all i. Hence these product effects are all multiples of each other, and $e^{1,2,\ldots,N} = e^{(1)} \odot \ldots \odot e^{(N)}$ for suitable local effects $e^{(j)}$. If we could non-trivially decompose any of the $e^{(j)}$, then we could decompose $e^{1,2,\ldots,N}$, which would contradict its purity. □

Appendix C: Spekkens' Toy Model

In this appendix, we briefly review some details of Spekkens' Toy Model [27] (STM) and its GPT extensions [21, 25, 56].

1. Overview

STM is essentially a classical hidden-variable model on which an *epistemic restriction* is imposed: no more than half the information (as measured in bits) can be known. The simplest (and for our purposes, only) single system in this framework then consists of a so-called *ontic* hidden variable with four possibilities $\{1, 2, 3, 4\}$. Valid questions about such system can only narrow down the state to at best two possibilities (e.g. "is the system in $1 \vee 2$ (read '1 or 2')?") for both affirmative and negative answers to the question. This yields three sets of mutually exclusive questions of the form "is the system in [X]" which we label as follows:

$$\begin{aligned} \langle x^+| &:= 1 \vee 3, & \langle x^-| &:= 2 \vee 4, \\ \langle y^+| &:= 1 \vee 4, & \langle y^-| &:= 2 \vee 3, \\ \langle z^+| &:= 1 \vee 2, & \langle z^-| &:= 3 \vee 4. \end{aligned} \tag{C1}$$

By the rules of STM, whenever such a question is asked, the ontic state must be randomized within the supporting set of states consistent with the answer to the question. For example, an affirmative answer to question $\langle z^+|$ will randomize the ontic state of the system to 1 or 2. This randomization ensures we cannot find the exact ontic state, say, by asking two different questions in a row – while maintaining the property that if we ask the same question twice in a row, we will get the same answer. Thus, one may define a set of maximum–knowledge *epistemic states* in one-to-one correspondence with the affirmative answer to these questions, labeled, e.g., as $|x^+\rangle = 1 \vee 2$. (STM also admits a "unit" question $u :=$ "is the system in $1 \vee 2 \vee 3 \vee 4$?" to which the answer is always affirmative; similarly, there is also a maximally mixed state, in which the ontic state can take any value with the same probability.)

The ontic state of a composite system is formed by taking the Cartesian product of each constituent system's ontic state (written for a and b as ab). The allowed epistemic states in this context then are those that satisfy the epistemic restriction both on the entire system, and also any subsystem thereof. Thus, a two-system epistemic state must admit at least four ontic possibilities. In addition to the Cartesian product of single system states, this also allows for "entangled" states, such as $11 \vee 22 \vee 33 \vee 44$, where even though the local marginal states are maximally mixed, perfect correlation is guaranteed if the same measurement is made on both systems. On the other hand, a state such as $11 \vee 12 \vee 33 \vee 44$ is forbidden. This is because should the $\langle z^+|$ measurement on the second system be answered in the affirmative, then the first system is definitely in state 1, which violates the epistemic restriction. It can thus be seen that STM is *self–dual by construction*: every maximum–knowledge measurement outcome can be uniquely identified with a maximum–knowledge epistemic state [46].

Transformations in the theory are performed by permuting the underlying hidden variable, in such a way that no valid epistemic state is taken to an invalid state. For single systems, every permutation is valid – but this is not the case for multipartite systems. Since these permutations are a finite group, when searching for a transformation that achieves a desired outcome (e.g. exhibits Darwinism), one can (with computer assistance) exhaustively search through possible transformations to find one that achieves the desired aims – or otherwise rule out its existence entirely [57]. However, by formalizing the similarity between STM and the stabilizer subset of quantum mechanics, Pusey [46] enables an elegant sufficient condition for the existence of a transformation, which we will subsequently describe.

2. GPT Extension

First, however, let us remark on the extension of STM into the GPT framework. In particular, STM defines a discrete state space with a finite number of states – so in order to treat it as a GPT, we must make it continuous. This is done in the obvious way: we treat the questions such as "is the system in $1 \vee 2$?" as an effect, and then admit all convex combinations of such effects. A complete (i.e. at least one question answers in the affirmative for any state) and mutually exclusive (i.e. no more than one question answers in the affirmative) set of questions maps to a set of effects that form a normalized measurement (i.e. will sum to the unit effect). Meanwhile, each set of epistemic states of maximum knowledge with no overlap in their ontic variable support (e.g. $\{1 \vee 2, 3 \vee 4\}$) form maximal frames, in which the maximum-knowledge epistemic states are extremal. We then allow convex combinations of such states as "mixed" states, yielding a theory dubbed STM–GPT. The set of allowed transformations on the theory are then defined as exactly those allowed on the (non-GPT) STM, and due to linearity, each of these uniquely extends into a transformation on the STM–GPT state space[1].

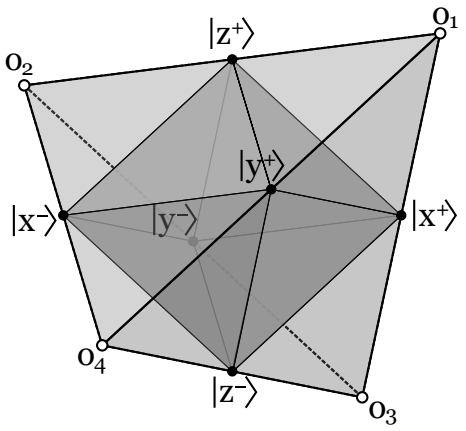


Figure 3. **Normalized states of a Spekkens' bit.** The tetrahedron is the normalized slice of $\mathbb{R}^4$ corresponding to the underlying classical ontic variable, with basis states $\{\vec{o}_1, \vec{o}_2, \vec{o}_3, \vec{o}_4\}$. The pure epistemic states correspond to the half-way points between these ontic states. The valid epistemic states of the theory are these states' octahedral convex hull.

One representation of a single system in STM–GPT in $\mathbb{R}^4$ is to identify each *ontic* state with a Cartesian vector, $\vec{o}_1 := (1,0,0,0)^{\mathrm{T}}$, $\vec{o}_2 := (0,1,0,0)^{\mathrm{T}}$, $\vec{o}_3 := (0,0,1,0)^{\mathrm{T}}$, $\vec{o}_4 := (0,0,0,1)^{\mathrm{T}}$, and then write each epistemic state $x \vee y$ as the vector $\frac{1}{2}(\vec{e}_x + \vec{e}_y)$ (see fig. 3). Here $A = \mathbb{R}^4$ and Ω_A is the convex combination of such (geometrically: this is the octahedron formed by connecting the midpoint of every line in a tetrahedron [21]). As observed in Janotta and Lal [25], the unrestricted dual of this space is cubic (i.e. a gbit) – but STM does not follow the no-restriction hypothesis. Rather, instead the space of effects can be represented by *exactly the same* vector space (carrying forward the self-duality-by-construction of STM), where the self-dualizing inner product $\langle e, \rho \rangle := 2e \cdot \rho$ is directly proportional to the Euclidean inner product on the real vector spaces.

An analogous representation can also be formed for n STM–GPT systems in $\mathbb{R}^{4^n}$. Take the Cartesian product $\{\vec{o}_1, \vec{o}_2, \vec{o}_3, \vec{o}_4\}^{\otimes n}$ to find the set of ontic states, and likewise define the epistemic set as valid (as per above) mixtures thereof. For example, $11 \vee 22 \vee 33 \vee 44$ is represented here as $\frac{1}{4}(\vec{o}_{11} + \vec{o}_{22} + \vec{o}_{33} + \vec{o}_{44})$. Meanwhile,

[1] This implies that not all symmetries of the state space of STM-GPT belong to the group of allowed transformations, $\mathcal{G}$. For instance, the rotation in the z-axis which permutes $|y^+\rangle \mapsto |x^+\rangle \mapsto |y^-\rangle \mapsto |x^-\rangle \mapsto |y^+\rangle$ is a symmetry of the octahedron but is not an allowed transformation in the ontic state space (see figure 3).

product states of lower-dimensional STM–GPT systems are simply found by the tensor product. For example, $1 \vee 2 \otimes 1 \vee 3 \equiv 11 \vee 13 \vee 21 \vee 23$ satisfies $\frac{1}{2}(\vec{o}_1 + \vec{o}_2) \otimes \frac{1}{2}(\vec{o}_1 + \vec{o}_2) = \frac{1}{4}(\vec{o}_{11} + \vec{o}_{13} + \vec{o}_{21} + \vec{o}_{23})$. This also allows for a self-dualizing inner product: $\langle \vec{e}, \vec{\rho} \rangle := 2^n \vec{e} \cdot \vec{\rho}$.

3. Stabilizer Formalism

Stabilizer groups originate in group theory, but have been adapted for use in quantum theory in the context of error-correcting codes and measurement-based quantum computation, as they provide concise ways to describe certain high-dimensional quantum states. Essentially, a transformation T is said to stabilize a state $|\psi\rangle$ if $T|\psi\rangle = |\psi\rangle$ [58]. Listing enough simultaneous stabilizing transformations may be enough to uniquely define a state (up to global phase): for example, the only two qubit state stabilized by both $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$ is the Bell state $|\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. The *stabilizer subset of quantum theory* are exactly the n qubit states that can be so described, when the stabilizers are taken from the Pauli group $P_n := \{\pm\mathbb{1}, \pm\sigma_x, \pm i\sigma_x, \pm\sigma_y, \pm i\sigma_y, \pm\sigma_z, \pm i\sigma_z\}^{\otimes n}$.

STM(–GPT) shares many similarities with (the convex hull of) quantum stabilizer states [46]. For instance, a qubit has six distinct pure qubit stabilizer states (stabilized by the Hermitian elements $\pm\sigma_x$, $\pm\sigma_y$, and $\pm\sigma_z$). Meanwhile, for an STM bit (using the GPT representation above), we can similarly define three "observable" matrices:

$$X := \mathrm{diag}(1, -1, 1, -1),$$
$$Y := \mathrm{diag}(1, -1, -1, 1),$$
$$Z := \mathrm{diag}(1, 1, -1, -1),$$

such that for each measurement, there is a unique (pure) epistemic state corresponding to the 1 and -1 eigenvector from each (e.g. $X|x^+\rangle = |x^+\rangle$) – and this covers all pure epistemic states. We can identify each of X, Y and Z respectively with the ontic state permutations

$$X \leftrightarrow 3412, \quad Y \leftrightarrow 4321, \quad Z \leftrightarrow 2143,$$

along with an identity element $I := \mathrm{diag}(1,1,1,1) \leftrightarrow 1234$. Then $\{I, X, Y, Z\}$ together with matrix multiplication is the Klein four-group V and is isomorphic to the permutation subgroup $\{1234, 3412, 4321, 2143\}$. The Cartesian product of these matrices with $\mathbb{Z}_2 = \{+1, -1\}$ forms the *toy stabilizer group* $G := \mathbb{Z}_2 \otimes V = \{\pm I, \pm X, \pm Y, \pm Z\}$.

Unlike the Pauli group, this group is Abelian with $XZ = ZX = Y$ (cf. $\sigma_x\sigma_z = -\sigma_z\sigma_x = -i\sigma_y$). For n bit systems, we denote the application of $T \in V$ to the k^{th} system as $T_k := I^{\otimes(k-1)} \otimes T \otimes I^{\otimes(n-k)}$. Finally, let us define the map $m : V^n \to P^n$ that makes an obvious identification between STM stabilizers and quantum stabilizers (e.g. $m : I_1 X_2 \mapsto \mathbb{1} \otimes \sigma_x$).

Now we may use the result of Pusey [46]: if a set of independent quantum stabilizers $m(R^1), m(R^2), \ldots m(R^k)$ describes a unique quantum state, then $R^1, R^2, \ldots, R^k$ describes a unique epistemic state in STM. Moreover, if a map on a set of quantum stabilizers $T : m(A^1) \mapsto m(B^1), \ldots, m(A^k) \mapsto m(B^k)$ defines a unitary quantum transformation and $m(A^1) \ldots m(A^k)$ are a *canonical generating set*, then $A^1 \mapsto B^1, \ldots, A^k \mapsto B^k$ defines a valid STM transformation. The full definition of canonical generating set is complicated, but for our purposes, it suffices to note that $\{X_1, \ldots X_k, Z_1, \ldots Z_k\}$ is one such set. With the aid of these sets, we can construct the FAN transformation (defining how it acts on each $\mathcal{X}_k/\mathcal{Z}_k$, that broadcasts information about the measurement $\{\langle z^+|, \langle z^-|\}$ to the environment, (see equation (23)).

4. STM is not strongly symmetric, nor does it have a decoherence map

In this section, we show that stabilizer quantum theory and (GPT-)STM fail to admit a decoherence map (in the sense of Richens et al. [23], as adapted in definition 9), and similarly neither theory obeys strong symmetry.

Lemma 10. *Stabilizer quantum states do not admit a decoherence map.*

Proof. By counterexample. Consider the classical 3-bit control-control-NOT gate that flips the third bit only if the first two bits are in state 1, and otherwise does nothing. This corresponds to a Toffoli gate in the quantum circuit, which is *not* a member of the Clifford group [49], and hence not a valid quantum stabilizer transformation. This violates condition 3 of definition 9: there is a classical reversible transformation that cannot be induced by a transformation in the theory. □

Analogously, there is a classical transformation that cannot be implemented in STM as well:

Lemma 11. *Spekkens' Toy Model does not admit a decoherence map.*

Proof. By counterexample. Consider the classical 3-bit transformation where bits 2 and 3 are swapped if bit 1 is set. This also forms a valid classical transformation in these bits. It is shown exhaustively by Garner [57] that STM does not have a 3-bit controlled SWAP. This then amounts to a valid classical transformation that cannot be induced within STM, violating condition 3 of definition 9. □

As we have argued in the main text, the ability to induce any reversible classical transformation on a frame's states (effects) is a necessary condition for strong symmetry on states (on effects) in a GPT system. Thus, we conclude:

Lemma 12. *Neither stabilizer quantum theory nor Spekkens' toy model satisfy strong symmetry (on states or effects).*

Chapter 8

Conclusions and outlook

In this thesis, we have addressed the foundations of quantum physics and spacetime using methods of quantum information theory. For that purpose we have applied several frameworks that extend the usual setting and formalism of quantum information theory to exotic and extreme regimes. As this approach to physics relies on operational concepts such as agents' interventions, protocols and distinguishability of similar scenarios [1, 2], it offers potential strategies for how to extract observable predictions in uncharted parts of physics.

The framework of *general/operational probabilistic theories* [3–9] allows to formulate theories of physics that are neither quantum nor classical. Within this approach, we investigated generalizations [10] of *Quantum Darwinism* [11, 12], a mechanism that explains objectivity of measurement outcomes in quantum physics. Furthermore, we analyzed a particular modification of quantum theory in which the dimension of the Bloch balls is changed [13]. Generalizing the approach of [14] to arbitrary dimension, we saw that such models do not allow for interactions, no matter how many Bloch balls are considered [15].

The setting of *device-independent information processing* [16–19] reduces physical devices to black boxes that are characterized solely by the statistics of their abstract input values and output values. Nonetheless, there exist spacetime scenarios such as Bell tests [20] for which these statistics can prove the presence of useful quantum correlations. We investigated a modification in which the abstract input values are replaced with spacetime parameters [21], such as angles of polarizers, orientations of extended objects, and time durations of laser pulses. This modification offers a new approach to understand the relation between quantum physics and spacetime.

The *process matrix* framework [22, 23] allows to investigate quantum information processing in so-called *indefinite causal structures* [24, 25]. Such radically new causal structures may appear in quantum gravity when light cones get blurred by quantum fluctuations and superpositions of mass configurations may induce superpositions of spacetimes [24–28]. Despite the crucial importance of the tensor product in quantum information theory, we saw that there cannot exist a universal formalism that describes protocols that apply several process matrices in parallel [29, 30]. Furthermore, based on observations of [31], we developed a systematic extension of the process matrix framework that models the local perception of time via quantum clocks [32].

We considered an operational scenario in which an external observer investigates a composite quantum system. In this scenario, the observer may only have access to an incomplete reference frame, for example it may lack an origin [33]. We saw that the symmetry transformations that such an incomplete reference frame cannot distinguish can be interpreted as quantum reference frame transformations [34–36].

While these frameworks differ in how they relax and generalize the usual settings of physics, they all have in common that quantum physics and spacetime often clash in unconventional ways. Probably the greatest problem of physics is the unification of quantum physics and general relativity into a satisfying theory of quantum gravity [37–39]. Our struggle implies that we have not understood the relation between spacetime and quantum physics well enough to achieve such a unification. Therefore, it might be crucial to challenge our understanding of quantum spacetime in new and exotic scenarios beyond the typical quantum field theory setting and in settings that carry operational meaning. In this thesis we encountered many such testing grounds.

The approach to device-independence with spacetime parameters can be extended in many ways, in particular by considering different causal scenarios [40, 41] or different spacetime parameters. In particular, a long term goal would be to approach relativity in this setting, by considering boxes whose input-output-statistics react to Lorentz-transformations applied to the box.

So far, we have only applied our setting with an outside experimenter who lacks a full reference frame to the simple example of a missing origin. Therefore, there are many choices of incomplete reference frames and spacetime symmetry groups that one can investigate and for which one can analyze the corresponding algebras of observables. Mathematically, we focused on classifying the algebras of states and observables and the embeddings as a whole. Important future work is to apply this formalism to specific examples of physical or practical interest, and in particular include dynamics. In that context, it might be important to better understand the physical difference of the different algebras we discovered, and how to apply them. It might also be fruitful to explore the connection of our formalism to the literature of constraint quantization, in particular in the context of quantum gravity [42–46].

The framework that combines process matrices and quantum clocks also offers important opportunities for further investigations. As the observation that the causal past of an agent is affine-linear in the operations of the other agents indicates, the explicit modeling of local time via quantum clocks induces further compatibility conditions that restrict the physically allowed process matrices. Our calculations during the project indicated that there might be more of such restrictions. In particular, an important open conjecture is that all the causal structures that fit into this physical scenario do not violate causal inequalities. One hint that supports this conjecture is the resynchronization problem that we encountered for the (reversed) Lugano process [47–50]. It is conceivable that the only indefinite causal structures with quantum clocks that can develop back into a standard definite causal structure are those whose indefiniteness relies on some kind of quantum control degree of freedom [51] - perhaps one that can be influenced by an agent to control the order of the other agents.

While the frameworks and scenarios introduced in this thesis offer a lot of opportunities to investigate the operational relation between spacetime and quantum physics, we also encountered no-go-theorems that severely restrict certain important approaches to quantum spacetime.

We considered modifications of quantum theory in which the Bloch ball is assumed to have a dimension that is different from three. As the Bloch vectors for many physical systems couple to rotations in position space, such models can be interpreted as a particular attempt to adapt quantum theory to a space in which space is not three-dimensional [52, 53]. However, we found that interactions are impossible, no matter how many Bloch balls are considered. Strategies to avoid the no-go theorem require giving up some important physical properties - like the fundamental reversibility of time evolution, tomographic locality [4], or that Bloch vectors can couple to all higher-dimensional rotations. It seems that Bloch balls have to be three-dimensional, even if space has more than three dimensions and Bloch vectors couple to rotations.

Furthermore, we investigated quantum information processing in scenarios that try to take advantage of parallel implementations of indefinite causal structures. While it was known before that the tensor product formalism fails [29], we saw that there cannot exist a satisfying replacement either. This severely restricts the applicability of many methods of quantum information theory, since one often considers the "asymptotic limit" of many copies that get combined in parallel [2]. During the development of the proof by my coworkers and me, we learned that the essence of the no-go theorem is the problem to combine incompatible causal orders. That is the same problem one encounters with the tensor product itself, although the contradiction is more difficult to isolate for general notions of parallel composition. Let us consider the physical target system that travels through a parallel implementation of indefinite causal structures, for example two quantum switches. If one imagines following the quantum system that the combined agents act on through the combined indefinite causal structure, it seems that one can only get a consistent picture if the local causal orders of the agents in the different processes that get combined are compatible.

At last, let us consider our extension of quantum Darwinism beyond quantum theory. We only generalized the ideal quantum Darwinism process and therefore there is a lot of quantum Darwinism literature left whose applicability beyond quantum theory is yet to be investigated. Nonetheless, let us focus on one particular aspect: In our Darwinism considerations, spacetime only played a small role. However, Darwinism describes the spreading of classical information from a central system to its environment. In our approach, we focussed on the original uncorrelated situation and the final situation in which perfect correlation with the full environment has been established via the fanout gate. But the region in between is of great interest, because it is here that the perfect correlation is established dynamically. This phenomenology of post-quantum Darwinism creates a specific scenario in post-quantum theories in which one can investigate the observable impact of a finite speed of information propagation in GPT analogues of quantum many body systems [54, 55].

In conclusion, while the approach to use methods and concepts of quantum

information theory to understand the foundations of physics is rather young, we have seen that it is very promising in the context of quantum foundations and spacetime. In particular, we identified several operational scenarios and no-go-theorems that challenge our understanding of spacetime in quantum physics. Mastering these challenges may not just help us in our strife for quantum gravity, but their focus on information processing and communication may also lead to new or improved technologies.

8.1 Bibliography of Chapter 8 (*Conclusions and outlook*)

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*.
Cambridge University Press, 2010.

[2] M. M. Wilde, *Quantum Information Theory*.
Cambridge University Press, 2 ed., 2017.

[3] M. Plávala, "General probabilistic theories: An introduction."
arXiv:2103.07469, 2021.

[4] J. Barrett, "Information processing in generalized probabilistic theories,"
Physical Review A **75**, 032304 (2007).

[5] L. Hardy, "Quantum Theory from Five Reasonable Axioms."
arXiv:quant-ph/0101012, 2001.

[6] B. Dakic and C. Brukner, "Quantum Theory and Beyond: Is Entanglement Special?." In "Deep Beauty: Understanding the Quantum World through Mathematical Innovation", Ed. H. Halvorson, Cambridge University Press, Cambridge, 2011.

[7] L. Masanes and M. P. Müller, "A derivation of quantum theory from physical requirements," *New Journal of Physics* **13**, 063001 (2011).

[8] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Informational derivation of quantum theory," *Physical Review A* **84**, 012311 (2011).

[9] G. Chiribella, G. M. D' Ariano, and P. Perinotti, "Quantum theory, namely the pure and reversible theory of information," *Entropy* **14**, 1877–1893 (2012).

[10] R. D. Baldijão, M. Krumm, A. J. P. Garner, and M. P. Müller, "Quantum Darwinism and the spreading of classical information in non-classical theories." arXiv:2012.06559, 2020.

[11] W. H. Zurek, "Relative States and the Environment: Einselection, Envariance, Quantum Darwinism, and the Existential Interpretation." Los Alamos preprint LAUR 07-4568, arXiv:0707.2832, 2007.

[12] W. H. Zurek, "Quantum Darwinism," *Nature Physics* **5**, 181–188 (2009).

[13] L. Masanes, M. P. Müller, D. Pérez-García, and R. Augusiak, "Entanglement and the three-dimensionality of the Bloch ball," *Journal of Mathematical Physics* **55**, 122203 (2014).

[14] G. de la Torre, L. Masanes, A. J. Short, and M. P. Müller, "Deriving quantum theory from its local structure and reversibility," *Physical Review Letters* **109**, 090403 (2012).

[15] M. Krumm and M. P. Müller, "Quantum computation is the unique reversible circuit model for which bits are balls," *npj Quantum Information* **5**, 7 (2019).

[16] J. Barrett, L. Hardy, and A. Kent, "No Signaling and Quantum Key Distribution," *Physical Review Letters* **95**, 010503 (2005).

[17] A. Acin, S. Massar, and S. Pironio, "Efficient quantum key distribution secure against no-signalling eavesdroppers," *New Journal of Physics* **8**, 126 (2006).

[18] A. Acín, N. Gisin, and L. Masanes, "From Bell's Theorem to Secure Quantum Key Distribution," *Physical Review Letters* **97**, 120405 (2006).

[19] U. Vazirani and T. Vidick, "Fully Device-Independent Quantum Key Distribution," *Physical Review Letters* **113**, 140501 (2014).

[20] J. S. Bell, "On the Einstein Podolsky Rosen paradox," *Physics Physique Fizika* **1**, 195–200 (1964).

[21] A. J. P. Garner, M. Krumm, and M. P. Müller, "Semi-device-independent information processing with spatiotemporal degrees of freedom," *Physical Review Research* **2**, 013112 (2020).

[22] O. Oreshkov, F. Costa, and C. Brukner, "Quantum correlations with no causal order," *Nature Commun.* **3**, 1092 (2012), `arXiv:1105.4464 [quant-ph]`.

[23] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, "Witnessing causal nonseparability," *New Journal of Physics* **17**, 102001 (2015).

[24] L. Hardy, "Probability Theories with Dynamic Causal Structure: A New Framework for Quantum Gravity." arXiv:gr-qc/0509120, 2005.

[25] L. Hardy, "Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure," *J. Phys. A: Math. Theor.* **40**, 3081–3099 (2007).

[26] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, "Quantum computations without definite causal structure," *Phys. Rev. A* **88**, 022318 (2013).

[27] M. Zych, *Quantum Systems under Gravitational Time Dilation*. Springer Theses. Springer, 2017.

[28] M. Zych, F. Costa, I. Pikovski, and Č. Brukner, "Bell's theorem for temporal order," *Nature Communications* **10**, 3772 (2019).

[29] D. Jia and N. Sakharwade, "Tensor products of process matrices with indefinite causal structure," *Phys. Rev. A* **97**, 032110 (2018).

[30] P. A. Guérin, M. Krumm, C. Budroni, and Č. Brukner, "Composition rules for quantum processes: a no-go theorem," *New Journal of Physics* **21**, 012001 (2019).

[31] E. Castro-Ruiz, F. Giacomini, A. Belenchia, and C. Brukner, "Quantum clocks and the temporal localisability of events in the presence of gravitating quantum systems," *Nature Commun.* **11**, 2672 (2020), `arXiv:1908.10165 [quant-ph]`.

[32] V. Baumann, M. Krumm, P. A. Guérin, and Č. Brukner, "Page-Wootters formulation of indefinite causal order." arXiv:2105.02304, 2021.

[33] M. Krumm, P. A. Höhn, and M. P. Müller, "Quantum reference frame transformations as symmetries and the paradox of the third particle." arXiv:2011.01951, 2020.

[34] F. Giacomini, E. Castro-Ruiz, and Č. Brukner, "Quantum mechanics and the covariance of physical laws in quantum reference frames," *Nature Communications* **10**, 494 (2019).

[35] A. Vanrietvelde, P. A. Hoehn, F. Giacomini, and E. Castro-Ruiz, "A change of perspective: switching quantum reference frames via a perspective-neutral framework," *Quantum* **4**, 225 (2020).

[36] A.-C. de la Hamette and T. D. Galley, "Quantum reference frames for general symmetry groups," *Quantum* **4**, 367 (2020).

[37] C. Rovelli, *Quantum Gravity*.
Cambridge University Press, Cambridge, 2004.

[38] D. Oriti, ed., *Approaches to quantum gravity*.
Cambridge University Press, Cambridge, 2009.

[39] C. Kiefer, *Quantum Gravity*.
Oxford University Press, Oxford, 3rd ed., 2012.

[40] M. Weilenmann and R. Colbeck, "Analysing causal structures in generalised probabilistic theories," *Quantum* **4**, 236 (2020).

[41] J. Henson, R. Lal, and M. F. Pusey, "Theory-independent limits on correlations from generalized Bayesian networks," *New Journal of Physics* **16**, 113043 (2014).

[42] P. A. Dirac, *Lectures on Quantum Mechanics*.
Yeshiva University Press, 1964.

[43] T. Thiemann, *Modern canonical quantum general relativity*.
Cambridge University Press, 2008.

[44] J. Tambornino, "Relational observables in gravity: a review," *SIGMA* **8**, 017 (2012).

[45] B. Dittrich, "Partial and complete observables for Hamiltonian constrained systems," *General Relativity and Gravitation* **39**, 1891–1927 (2007).

[46] B. Dittrich, "Partial and complete observables for canonical general relativity," *Classical and Quantum Gravity* **23**, 6155 (2006).

[47] Ä. Baumeler and S. Wolf, "Perfect signaling among three parties violating predefined causal order," in *2014 IEEE International Symposium on Information Theory (ISIT)*, pp. 526–530.
IEEE, Piscataway, NJ, June, 2014.

[48] Ä. Baumeler, A. Feix, and S. Wolf, "Maximal incompatibility of locally classical behavior and global causal order in multi-party scenarios," *Phys. Rev. A* **90**, 042106 (2014).

[49] Ä. Baumeler and S. Wolf, "The space of logically consistent classical processes without causal order," *New J. Phys.* **18**, 013036 (2016).

[50] P. A. Guérin and Č. Brukner, "Observer-dependent locality of quantum events," *New Journal of Physics* **20**, 103031 (2018).

[51] J. Wechs, H. Dourdent, A. A. Abbott, and C. Branciard, "Quantum circuits with classical versus quantum control of causal order." arXiv:2101.08796, 2021.

[52] M. P. Mueller and L. Masanes, "Three-dimensionality of space and the quantum bit: an information-theoretic approach," *New Journal of Physics* **15**, 053040 (2013).

[53] B. Dakić and Č. Brukner, "The classical limit of a physical theory and the dimensionality of space." in "Quantum Theory: Informational Foundations and Foils", eds. G. Chiribella and R. W. Spekkens, Springer, 2016.

[54] E. H. Lieb and D. W. Robinson, "The finite group velocity of quantum spin systems," *Commun. Math. Phys* **28**, 251–257 (1972).

[55] M. Cheneau, P. Barmettler, D. Poletti, M. Endres, P. Schauß, T. Fukuhara, C. Gross, I. Bloch, C. Kollath, and S. Kuhr, "Light-cone-like spreading of correlations in a quantum many-body system," *Nature* **481**, 484–487 (2012).

Appendix **A**

List of publications

The following list contains the publications that I worked on and that are included in this thesis. * denotes shared first authorship.

1. V. Baumann*, **M. Krumm***, P. A. Guérin, Č. Brukner, *Page-Wootters formulation of indefinite causal order*, preprint available as arXiv:2105.02304 (2021). Submitted to *Physical Review Research* for review.

   Contribution: Veronika Baumann and I share first authorship. I made essential contributions to all parts of the research. I provided the leading input for developing the protocol to implement arbitrary coherently controlled causal order, and for the result that the causal past of an agent has to be affine-linear in the operations of the other agents.

2. R. D. Baldijão*, **M. Krumm***, A. J. P. Garner, M. P. Müller, *Quantum Darwinism and the spreading of classical information in non-classical theories*, preprint available as arXiv:2012.06559 (2020). Submitted to *Quantum* for review.

   Contribution: Roberto D. Baldijão and I share first authorship of the paper. I made essential contributions to all parts of the research. I provided the leading input for the sufficient conditions for GPT Darwinism. I massively simplified Roberto D. Baldijão's argument for GPT Darwinism in Spekken's toy model, leading to the fanout gate and proof that it achieves GPT Darwinism in Spekken's toy model as presented in the thesis.

3. **M. Krumm**, P. A. Höhn, M. P. Müller, *Quantum reference frame transformations as symmetries and the paradox of the third particle*, preprint available as arXiv:2011.01951 (2020). Submitted to *Quantum* for review.

   Contribution: I made contributions to all aspects of the research. Markus P. Müller initiated the project, I joined in the beginning, and Philipp Höhn joined after a few months. The paper developed as a continuous, dynamical interaction between Philipp Höhn, Markus P. Müller and me, providing conceptual considerations, mathematical lemmas, proofs, proof-reading and corrections for each other.

4. A. J. P. Garner, **M. Krumm**, M. P. Müller, *Semi-device-independent information processing with spatiotemporal degrees of freedom*, Phys. Rev. Research **2**, 013112 (2020). Preprint available as arXiv:1907.09274.

   Contribution: The authors are in alphabetical order. I made essential contributions to all parts of the research. I provided the leading input for the local hidden variable models.

5. P. A. Guérin, **M. Krumm**, C. Budroni, Č. Brukner, *Composition rules for quantum processes: a no-go theorem*, New J. Phys. **21**, 012001 (2019). Preprint available as arXiv:1806.10374.

   Contribution: I made essential contributions to all parts of the research. I provided the leading input for extending convex bilinearity to full bilinearity, and to extending the proof to all dimensions.

6. **M. Krumm**, M. P. Müller, *Quantum computation is the unique reversible circuit model for which bits are balls*, npj Quantum Information **5**, 7 (2019). Preprint available as arXiv:1804.05736.

   Contribution: Markus suggested the original idea for a project to construct interactions of higher dimensional Bloch balls. Instead, I discovered the no-go theorem and developed the first full proof.

I also worked on the following publication that is NOT included in this thesis.

7. **M. Krumm**, M. P. Müller, *Computational irreducibility and compatibilism: towards a formalization*, preprint available as arXiv:2101.12033 (2021).