PAPER • OPEN ACCESS

Entanglement, non-Markovianity, and causal non-separability

To cite this article: Simon Milz *et al* 2018 *New J. Phys.* **20** 033033

View the article online for updates and enhancements.

New Journal of Physics

The open access journal at the forefront of physics

CrossMark

PAPER

OPEN ACCESS

Entanglement, non-Markovianity, and causal non-separability

Simon Milz[1], Felix A Pollock[1], Thao P Le[1,2], Giulio Chiribella[3,4,5,6] and Kavan Modi[1]

1 School of Physics and Astronomy, Monash University, Clayton, Victoria 3800, Australia
2 Dept. of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom
3 Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, United Kingdom
4 Canadian Institute for Advanced Research, CIFAR Program in Quantum Information Science, Toronto, ON M5G 1Z8, Canada
5 Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong
6 HKU Shenzhen Institute of Research and Innovation, Kejizhong 2nd Road, Shenzhen, China
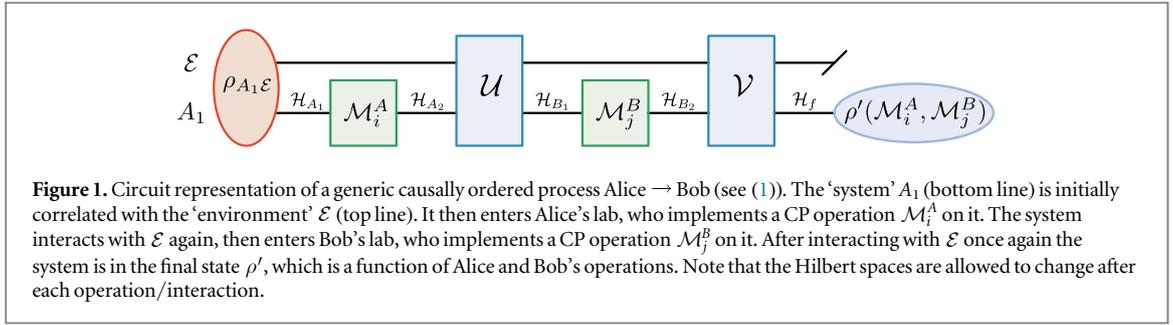
E-mail: simon.milz@monash.edu

Abstract

Quantum mechanics, in principle, allows for processes with indefinite causal order. However, most of these causal anomalies have not yet been detected experimentally. We show that every such process can be simulated experimentally by means of non-Markovian dynamics with a measurement on additional degrees of freedom. In detail, we provide an explicit construction to implement arbitrary a causal processes. Furthermore, we give necessary and sufficient conditions for open system dynamics with measurement to yield processes that respect causality locally, and find that tripartite entanglement and nonlocal unitary transformations are crucial requirements for the simulation of causally indefinite processes. These results show a direct connection between three counter-intuitive concepts: entanglement, non-Markovianity, and causal non-separability.

1. Introduction

Temporal order is one of the fundamental pillars that both our everyday understanding of the world, as well as our physical theories are built on. Events, no matter how complicated the underlying dynamical theory, seem to happen in a causal succession and there is a clear arrow of time that defines in which direction they can influence each other. However, the impression of causal order might only be locally true. Future experiments may challenge the idea that causal order is fundamental, and may reduce it to a property that exists locally but is violated globally. For example, an experiment could consist of two parties (Alice and Bob) conducting measurements in their separated laboratories. The temporal order of events would be heralded by the joint probability distributions of their measurement outcomes. While Alice and Bob experience a well-defined temporal order in their respective laboratories, it is fathomable that a third party (Charlie) that receives measurement data from both Alice and Bob is unable to assign a relative causal order to them based on the received data.

An example of a causally unordered process is the quantum switch, theoretically introduced in [1] and experimentally realised in [2, 3]. Besides the quantum switch, no other exotic causal structure has been implemented experimentally so far. Nonetheless, the mathematical description of such structures is well developed [1, 4] and is subject to active research (see, e.g., [5–10]). The main mathematical object to represent general processes is the *process matrix*, introduced in [4] for two parties and later extended to multiple parties in [8]. In [4], the authors showed that this framework allows for *causally non-separable* process matrices—i.e., process matrices that cannot be written as a probabilistic mixture of causal ones. These causally non-separable process matrices go beyond what can be described by quantum mechanics that respects causal order and also encapsulate processes that can violate *causal inequalities*, i.e., processes that do not allow for an underlying causal model.

By definition, no process that is compatible with a global causal order exhibits correlations that are obtained from a causally non-separable process matrix. However, processes without causal order can be simulated non-

Figure 1. Circuit representation of a generic causally ordered process Alice → Bob (see (1)). The 'system' A_1 (bottom line) is initially correlated with the 'environment' $\mathcal{E}$ (top line). It then enters Alice's lab, who implements a CP operation $\mathcal{M}_i^A$ on it. The system interacts with $\mathcal{E}$ again, then enters Bob's lab, who implements a CP operation $\mathcal{M}_j^B$ on it. After interacting with $\mathcal{E}$ once again the system is in the final state ρ', which is a function of Alice and Bob's operations. Note that the Hilbert spaces are allowed to change after each operation/interaction.

deterministically, i.e., by conditioning the collection of data on an additional measurement outcome. For example, Charlie might measure an additional system that he possesses, which has interacted with Alice and Bob. He could choose to only record the data he receives from Alice and Bob when the measurement of his system yields a particular outcome. Even if the causal ordering of Alice's and Bob's laboratory is fixed, the data that Charlie records could lead him to believe that there is no temporal ordering between Alice and Bob. More generally, it has been shown that *any* process matrix—causally ordered or not—can be implemented experimentally by a quantum circuit (i.e., a causally ordered process) with an additional measurement [11–14] [7].

In this article, we consider the task of implementing processes with indefinite causal order. We answer two natural questions: Given a process, what is its circuit implementation (with measurement)? What resources are necessary to simulate a causally non-separable process? We concretely relate process matrices to *quantum combs* [11, 12, 19] and *process tensors* [20, 21], which are a generalisation of completely positive (CP) maps used to describe general causally-ordered processes and non-Markovian quantum phenomena. Building on this relation, we provide a general implementation scheme for arbitrary processes. This scheme requires a genuinely tripartite entangled initial state. Moreover, we provide necessary and sufficient conditions for a general circuit with measurement to yield a valid process, and give an explicit example of causally non-separable process matrices that can be simulated with a probability that exceeds 50%. Finally, we show that—independent of the implementation scheme—the simulation of causally non-separable process matrices requires both genuine tripartite entanglement in the initial state, as well as nonlocal unitary dynamics, i.e., it requires the underlying causal process to be non-Markovian. These results provide a constructive way to experimentally simulate arbitrary process matrices and establish a clear connection between entanglement, non-Markovianity and causal non-separability.

2. Causally ordered processes

In order to render the structure of causally unordered processes more transparent, we reiterate existing results about *causally ordered* processes. These are processes where the temporal order of the operations performed by Alice and Bob, respectively, is well-defined. Throughout this article, we will mainly focus on the two-party case. Hereinafter, we consider the following scenario: Alice (Bob) has a quantum instrument $\mathcal{J}_A$ ($\mathcal{J}_B$) in her (his) laboratory, i.e., a set of CP trace non-increasing maps $\{\mathcal{M}_i^X\}$, $\mathcal{M}_i^X : \mathcal{L}(\mathcal{H}_{X_1}) \rightarrow \mathcal{L}(\mathcal{H}_{X_2})$ with $X \in \{A, B\}$. $\mathcal{L}(\mathcal{H}_{X_y})$ represents the space of bounded operators on $\mathcal{H}_{X_y}$ and the input Hilbert space $\mathcal{H}_{X_1}$ is not necessarily the same as the output Hilbert space $\mathcal{H}_{X_2}$. The labels i correspond to the outcomes of the instrument, and the entire procedure, $\sum_{i=1} \mathcal{M}_i^X$, is a CP and trace-preserving (CPTP) map. For example, suppose Alice, upon receiving a state $\rho \in \mathcal{L}(\mathcal{H}_{A_1})$, performs a measurement in the computational basis. Upon observing outcome m she prepares a state $\rho_m \in \mathcal{L}(\mathcal{H}_{A_2})$ and sends it forward. This choice of instrument corresponds to the CPTP map $\sum_m \mathcal{M}_m^A[\rho] = \sum_m \langle m|\rho|m \rangle \rho_m$, where for each m we have the CP map $\mathcal{M}_m^A[\rho] = \langle m|\rho|m \rangle \rho_m$.

A generic causally ordered process, where Alice goes before Bob, is of the following form: Alice receives a quantum system $\mathcal{S}$ in state $\rho \in \mathcal{L}(\mathcal{H}_{A_1})$—possibly correlated with an environment $\mathcal{E}$—and performs (non-deterministically) a CP operation $\mathcal{M}_i^A$ on it. Alice's output state is sent to Bob via a quantum communication channel. He also performs (non-deterministically) a CP operation $\mathcal{M}_j^B$. In the end, Bob's output state is sent through another quantum communication channel to yield the final state $\rho' \in \mathcal{L}(\mathcal{H}_f)$. This scenario is depicted in terms of a circuit representation in figure 1. The two quantum communication channels are—in general—correlated and can be represented by two system-environment unitary maps $\mathcal{U}$ and $\mathcal{V}$ (where, e.g., $\mathcal{U}[\rho_{\mathcal{SE}}] = U \rho_{\mathcal{SE}} U^\dagger$, with $UU^\dagger = \mathbf{1}_{\mathcal{SE}}$), and consequently, the final state, which depends on the performed CP operations, can be written as

---

[7] In a slightly different context, schemes involving conditioning of data are also actively investigated both theoretically as well as experimentally with respect to the simulation of closed timelike curves (see, e.g., [15–18]).
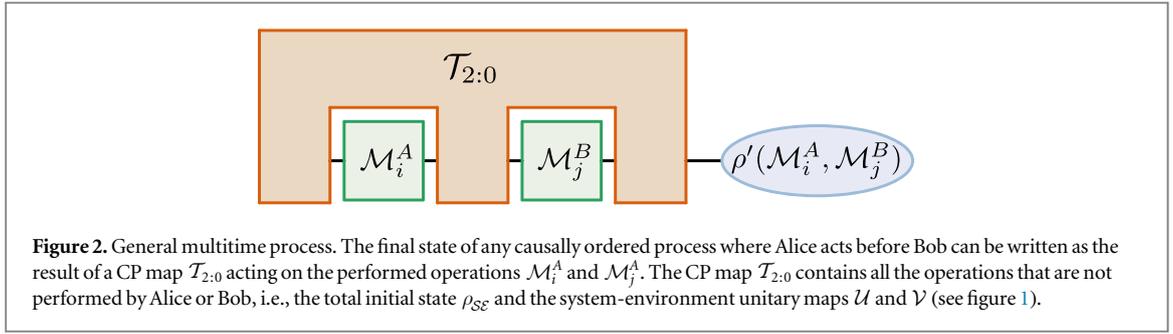
Figure 2. General multitime process. The final state of any causally ordered process where Alice acts before Bob can be written as the result of a CP map $\mathcal{T}_{2:0}$ acting on the performed operations $\mathcal{M}_i^A$ and $\mathcal{M}_j^A$. The CP map $\mathcal{T}_{2:0}$ contains all the operations that are not performed by Alice or Bob, i.e., the total initial state $\rho_{S\mathcal{E}}$ and the system-environment unitary maps $\mathcal{U}$ and $\mathcal{V}$ (see figure 1).
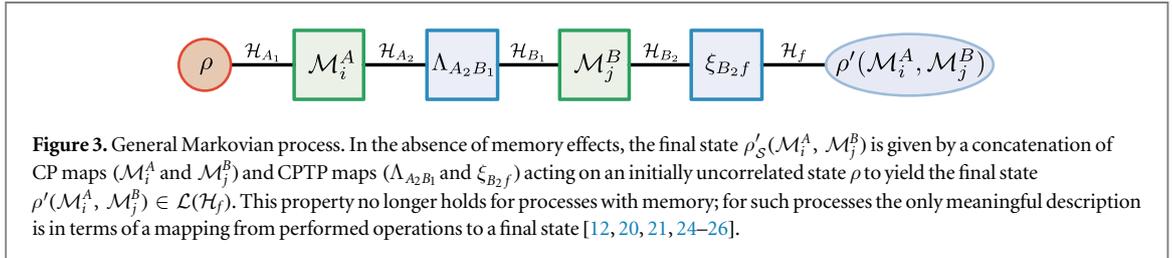

Figure 3. General Markovian process. In the absence of memory effects, the final state $\rho_S'(\mathcal{M}_i^A, \mathcal{M}_j^B)$ is given by a concatenation of CP maps ($\mathcal{M}_i^A$ and $\mathcal{M}_j^B$) and CPTP maps ($\Lambda_{A_2B_1}$ and ξ_{B_2f}) acting on an initially uncorrelated state ρ to yield the final state $\rho'(\mathcal{M}_i^A, \mathcal{M}_j^B) \in \mathcal{L}(\mathcal{H}_f)$. This property no longer holds for processes with memory; for such processes the only meaningful description is in terms of a mapping from performed operations to a final state [12, 20, 21, 24–26].

$$\rho'(\mathcal{M}_i^A, \mathcal{M}_j^B) = \mathrm{tr}_{\mathcal{E}}\{(\mathcal{V} \circ \mathcal{M}_j^B \circ \mathcal{U} \circ \mathcal{M}_i^A)[\rho_{S\mathcal{E}}]\}, \tag{1}$$

where we have omitted the respective identity operators on the environment. We emphasise the dependence of the output state on Alice's and Bob's quantum operations, while we omit the dependence on the unitary maps $\mathcal{U}$ and $\mathcal{V}$. This is because we regard Alice's and Bob's operations as variables that Alice and Bob can choose freely, while the rest of the circuit is fixed.

Since $\mathcal{S}$ is initially correlated with $\mathcal{E}$, and $\mathcal{U}$ and $\mathcal{V}$ act on the same $\mathcal{E}$, this noisy process can be temporally correlated. Such processes are also known as a *non-Markovian processes*. Mathematically, they can be described with the framework of quantum combs [11, 12, 19], or equivalently, with the framework of process tensors [20], where every process of the form (1) can be rewritten as a linear mapping $\mathcal{T}_{2:0}$ from the performed operations $\mathcal{M}_i^A$ and $\mathcal{M}_j^B$ to the final state $\rho'(\mathcal{M}_i^A, \mathcal{M}_j^B)$. Explicitly, we have

$$\rho'(\mathcal{M}_i^A, \mathcal{M}_j^B) = \mathcal{T}_{2:0}[\mathcal{M}_i^A, \mathcal{M}_j^B], \tag{2}$$

where the linear map $\mathcal{T}_{2:0}$ is called a quantum supermap [12], or a process tensor [20]. The notation $\mathcal{T}_{2:0}$ refers to the fact that the process has two open slots and produces an output with no open slot. Similar frameworks were proposed by Gutoski and Watrous [22] and Hardy [23].

A graphical representation for multitime processes is provided in figure 2. The linear map $\mathcal{T}_{2:0}$ is the generalisation of a quantum channel to multiple time steps [24]. Such a generalisation straightforwardly accounts for temporal correlations, i.e., non-Markovian open quantum dynamics. When a process is Markovian, it reduces to a sequence of CPTP maps (see figure 3). We will show below that Non-Markovianity plays an important role when simulating causally inseparable process. In the following section, we provide the mathematical restrictions that have to be imposed on the linear map $\mathcal{T}_{2:0}$ in order for it to describe a valid causally ordered process.
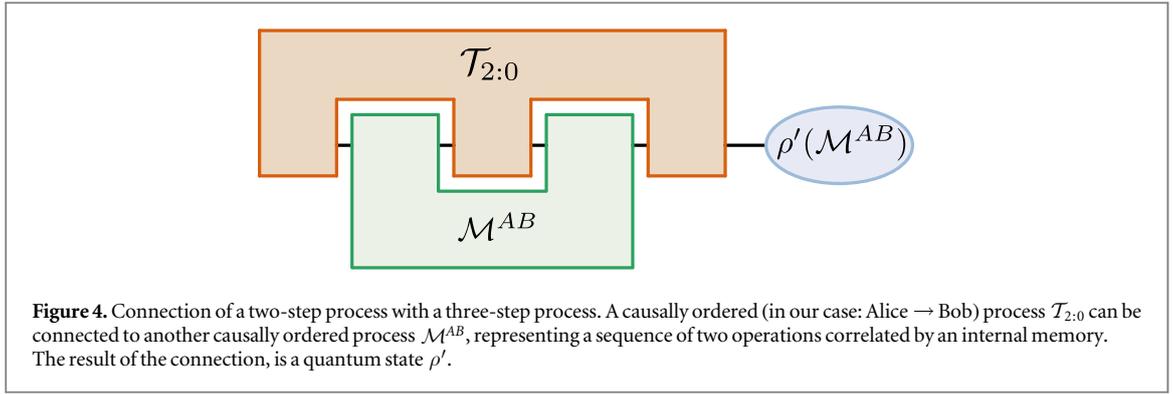
2.1. Properties of causally ordered processes

Mathematically, the map $\mathcal{T}_{2:0}$ is a mapping from pairs of CP maps to a final state $\rho' \in \mathcal{L}(\mathcal{H}_f)$. This mathematical structure can be made more manifest by employing the *Choi-Jamiołkowski isomorphism* [27, 28]:

Let $\mathcal{L}(\mathcal{H}_{X_1})$ denote the set of linear operators on the Hilbert space $\mathcal{H}_{X_1}$. Every CP map $\mathcal{M}^X : \mathcal{L}(\mathcal{H}_{X_1}) \to \mathcal{L}(\mathcal{H}_{X_2})$ can be mapped isomorphically onto a positive matrix $M^{X_2X_1} \in \mathcal{L}(\mathcal{H}_{X_2} \otimes \mathcal{H}_{X_1})$. For an arbitrary CP map $\mathcal{M}_k^X : \mathcal{L}(\mathcal{H}_{X_1}) \to \mathcal{L}(\mathcal{H}_{X_2})$ the *Choi state* of the map is given by

$$M_k^{X_2X_1} := d_{X_1}(\mathcal{M}_k^X \otimes \mathcal{I})[|\phi_{X_1}^+\rangle\langle\phi_{X_1}^+|] := (\mathcal{M}_k^X \otimes \mathcal{I})[\phi_{X_1}^+], \tag{3}$$

where $|\phi_{X_1}^+\rangle = \frac{1}{\sqrt{d_{X_1}}}\sum_{n=1}^{d_{X_1}}|nn\rangle \in \mathcal{H}_{X_1} \otimes \mathcal{H}_{X_1}$ is a normalised maximally entangled state and $\mathcal{I}$ is the identity map in the appropriate dimension. For the map to also be trace preserving, it has to satisfy the additional constraint $\mathrm{tr}_{X_2}(M^{X_2X_1}) = \mathbf{1}_{X_1}$, i.e., the trace over the Hilbert space of the output of the map has to yield the identity matrix on the input space. The action of a CP map $\mathcal{M}_k^X$ on a state $\rho \in \mathcal{L}(\mathcal{H}_{X_1})$ can be written in terms of its Choi state:

Figure 4. Connection of a two-step process with a three-step process. A causally ordered (in our case: Alice → Bob) process $\mathcal{T}_{2:0}$ can be connected to another causally ordered process $\mathcal{M}^{AB}$, representing a sequence of two operations correlated by an internal memory. The result of the connection, is a quantum state ρ'.

$$\mathcal{M}_k^X(\rho) = \mathrm{tr}_{X_1}[(\mathbf{1}_{X_2} \otimes \rho^{\mathrm{T}})M_k^{X_2 X_1}], \qquad (4)$$

where $\cdot^{\mathrm{T}}$ denotes the transpose with respect to the computational basis.

Similar mathematical relations hold for multitime processes. Specifically, a process $\mathcal{T}_{2:0}$ that maps a pair of CP maps into a state, can be represented by a Choi state $\Upsilon_{2:0} \in \mathcal{L}(\mathcal{H}_f \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}_{A_1} \otimes \mathcal{H}_{B_2} \otimes \mathcal{H}_{B_1})$. The action of the process $\mathcal{T}_{2:0}$ on the Choi states of the two input CP maps can—in clear analogy to (4)—be written as [12]

$$\rho'(\mathcal{M}_i^A, \mathcal{M}_j^B) = \mathrm{tr}_{\mathcal{S}_2 \mathcal{S}_1}[(\mathbf{1}_f \otimes (M_i^{A_2 A_1})^{\mathrm{T}} \otimes (M_j^{B_2 B_1})^{\mathrm{T}})\Upsilon_{2:0}], \qquad (5)$$

where we have employed the convention $\mathcal{S}_y = A_y B_y$. The corresponding probability to measure the outcomes i and j (given the instruments $\mathcal{J}_A$ and $\mathcal{J}_B$) in the same run can be computed via [12]

$$p(i, j|\mathcal{J}_A, \mathcal{J}_B) = \mathrm{tr}_{\mathcal{S}_1 \mathcal{S}_2}\{[(M_i^{A_2 A_1})^{\mathrm{T}} \otimes (M_j^{B_2 B_1})^{\mathrm{T}}]\mathrm{tr}_f(\Upsilon_{2:0})\}. \qquad (6)$$

This expression allows us to derive restrictions on $\Upsilon_{2:0}$ to represent a causally ordered process. For example, the statement 'Alice goes before Bob', means that no operation that Bob performs in his laboratory can influence the measurement statistics of Alice's experiment. In terms of the outcome probabilities, this means that $p(i|\mathcal{J}_A) = \sum_j p(i, j|\mathcal{J}_A, \mathcal{J}_B)$ is independent of the choice of instrument $\mathcal{J}_B$ for all possible choices of instruments in Alice's laboratory. It is straightforward to prove [11, 12] that this requirement implies

$$\mathrm{tr}_f(\Upsilon_{2:0}) = \mathbf{1}_{B_2} \otimes \Upsilon_{1:0}^{B_1 A_2 A_1}, \qquad (7)$$

where $\Upsilon_{1:0}^{B_1 A_2 A_1} \in \mathcal{L}(\mathcal{H}_{B_1} \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}_{A_1})$ is the Choi operator of a multitime process with a single open slot. Employing the same reasoning again yields the final restriction

$$\mathrm{tr}_{B_1}(\Upsilon_{1:0}^{B_1 A_2 A_1}) = \mathbf{1}_{A_2} \otimes \rho, \qquad (8)$$

where ρ is the initial system state, i.e., $\rho = \mathrm{tr}_{\mathcal{E}}(\rho_{\mathcal{SE}})$ [12, 25].

The unitary circuit of (1) automatically leads to Choi operators that fulfil these requirements. Conversely, every Choi operator satisfying equations (7) and (8) can be represented by a unitary circuit like the one depicted in figure 1 [12, 20]. Analogously, a process that is causally ordered Bob → Alice would have to fulfil the same trace requirements, but with the roles of B_2 (B_1) and A_2 (A_1) interchanged.

A causally ordered (Alice → Bob) process $\mathcal{T}_{2:0}$ cannot only be meaningfully applied to independent CP operations $\mathcal{M}_i^A$ and $\mathcal{M}_j^B$, but also to temporally correlated operations $\mathcal{M}^{AB}$ (see figure 4). For example, Alice could send the result of her measurement to Bob, and he conditions his choice of instrument on said outcome; or Alice could send Bob the ancilla that she used to implement her instrument, and he uses the same ancilla to perform his operation. The resulting *temporally correlated* operation $\mathcal{M}^{AB}$ is itself a causally ordered process (Alice → Bob), and as such, its Choi state $M^{B_2 B_1 A_2 A_1}$ has to fulfil

$$M^{B_2 B_1 A_2 A_1} \geqslant 0, \quad \mathrm{tr}_{B_2}(M^{B_2 B_1 A_2 A_1}) = \mathbf{1}_{B_1} \otimes M^{A_2 A_1} \quad \text{and} \quad \mathrm{tr}_{A_2}(M^{A_2 A_1}) = \mathbf{1}_{A_1}. \qquad (9)$$

With this, we can equivalently restate the requirements on $\Upsilon_{2:0}$ in terms of a probability preservation condition; a matrix $\Upsilon_{2:0} \in \mathcal{L}(\mathcal{H}_f \otimes \mathcal{H}_{B_2} \otimes \mathcal{H}_{B_1} \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}_{A_1})$ is a valid causally ordered two-step process tensor, if it maps every causally ordered $M^{B_2 B_1 A_2 A_1}$ to a unit trace quantum state, i.e.:

$$\Upsilon_{2:0} \geqslant 0 \quad \text{and} \quad \mathrm{tr}[\mathrm{tr}_f(\Upsilon_{2:0})(M^{B_2 B_1 A_2 A_1})^{\mathrm{T}}] = 1 \quad \forall \; M^{B_2 B_1 A_2 A_1} \text{ that fulfil (9)}. \qquad (10)$$

We will see in the following section that a relaxation of condition (10) directly leads to processes with indefinite causal order.

3. Locally causal processes—the process matrix

A causally ordered process imposes a global temporal order (e.g., Alice goes before Bob), as well as a local temporal order, i.e., the outputs of Alice's (Bob's) instrument do not influence its respective inputs. Oreshkov *et al* [4] introduced the *process matrix* framework in order to describe quantum processes without reference to a global causal order. For two parties—Alice and Bob—the framework is a generalisation of the framework described in section 2 (the generalisation to more than two parties is straightforward [8]): Alice and Bob both have quantum instruments $\mathcal{J}_A$ and $\mathcal{J}_B$. They can choose their instrument at will; in contrast to the above case, their respective laboratories are separated, such that they cannot create correlated operations. All they can do is to individually receive a quantum state, apply an operation to it, and send out the result. Consequently, the Choi states[8] of operations that Alice and Bob can perform are of the form $M_i^{B_2 B_1} \otimes M_j^{A_2 A_1}$ only.

For the choice of instruments $\mathcal{J}_A$ and $\mathcal{J}_B$ of Alice and Bob, the probability of outcomes i and j is then given by

$$\mathbb{P}(i, j \mid \mathcal{J}_A, \mathcal{J}_B) = \mathrm{tr}\{W^{B_2 B_1 A_2 A_1}[(M_i^{A_2 A_1})^\mathrm{T} \otimes (M_j^{B_2 B_1})^\mathrm{T}]\}, \tag{11}$$

where the positive-semidefinite matrix $W^{B_2 B_1 A_2 A_1} \in \mathcal{L}(\mathcal{H}^{B_2} \otimes \mathcal{H}^{B_1} \otimes \mathcal{H}^{A_2} \otimes \mathcal{H}^{A_1})$ is called the *process matrix*. We denote the probabilities obtained from a process matrix by $\mathbb{P}$ to clearly distinguish them from probabilities obtained from causal process (possibly with conditioning—see section 4.1). The only restriction on $W^{B_2 B_1 A_2 A_1}$ is that it has the correct normalisation for independent CPTP operations, i.e.,

$$W^{B_2 B_1 A_2 A_1} \geqslant 0 \quad \text{and} \quad \mathrm{tr}\{W^{B_2 B_1 A_2 A_1}[(M^{A_2 A_1})^\mathrm{T} \otimes (M^{B_2 B_1})^\mathrm{T}]\} = 1, \tag{12}$$

for all Choi state $M^{A_2 A_1}$ and $M^{B_2 B_1}$ of CPTP maps. Positivity of $W^{B_2 B_1 A_2 A_1}$ ensures that probabilities are positive[9], while the trace condition enforces local causality (i.e., in Alice's and Bob's separate laboratories).

It is important to note the similarity between the conditions (12) for process matrices and the conditions (10) imposed on quantum combs. Process matrices yield unit probability on the affine span of the set of product CPTP maps $M^{B_2 B_1} \otimes M^{A_2 A_1}$; this set coincides with the set of no-signalling operations [1, 30], which is strictly smaller than the set of temporally correlated causally ordered operations $M^{B_2 B_1 A_2 A_1}$. Consequently, the set of admissible process matrices $W^{B_2 B_1 A_2 A_1}$ is strictly larger than the set of temporally ordered processes $\mathrm{tr}_f(\Upsilon)$ (here, and in what follows, we will drop the subscripts of the process tensor and its Choi state); process matrices can describe temporal correlations that do not agree with a global causal order. We list the explicit restrictions that local causality imposes on process matrices in the appendix.

If $W^{B_2 B_1 A_2 A_1}$ corresponded to a causally ordered process, it would be of the form $W^{B_2 B_1 A_2 A_1} = \mathbb{1}_{B_2} \otimes \Upsilon^{B_1 A_2 A_1}$ (Alice goes before Bob) or $W^{B_2 B_1 A_2 A_1} = \mathbb{1}_{A_2} \otimes \Upsilon^{A_1 B_2 B_1}$ (Bob goes before Alice), where $\Upsilon^{B_1 A_2 A_1}$ and $\Upsilon^{A_1 B_2 B_1}$ are correctly causally ordered one-step process tensors. It is also conceivable that the causal structure is not known with certainty, or depends on an exterior statistical parameter (like, e.g., the flipping of a coin). The corresponding process matrix could then be expressed by a convex combination of causally ordered ones:

$$W_{\mathrm{caus.sep.}}^{B_2 B_1 A_2 A_1} = q(\mathbb{1}_{B_2} \otimes \Upsilon^{B_1 A_2 A_1}) + (1 - q)(\mathbb{1}_{A_2} \otimes \Upsilon^{A_1 B_2 B_1}), \quad q \in [0, 1]. \tag{13}$$

Any process matrix that can be written in the form (13) is called *causally separable*. Processes with an underlying causal order, or processes where Alice and Bob cannot influence each other (e.g., Alice and Bob could share an entangled state but are spacelike separated) can be described by a process matrix that is of the form (13). The set of process matrices defined by (12) contains process matrices that are *not* causally separable [4].

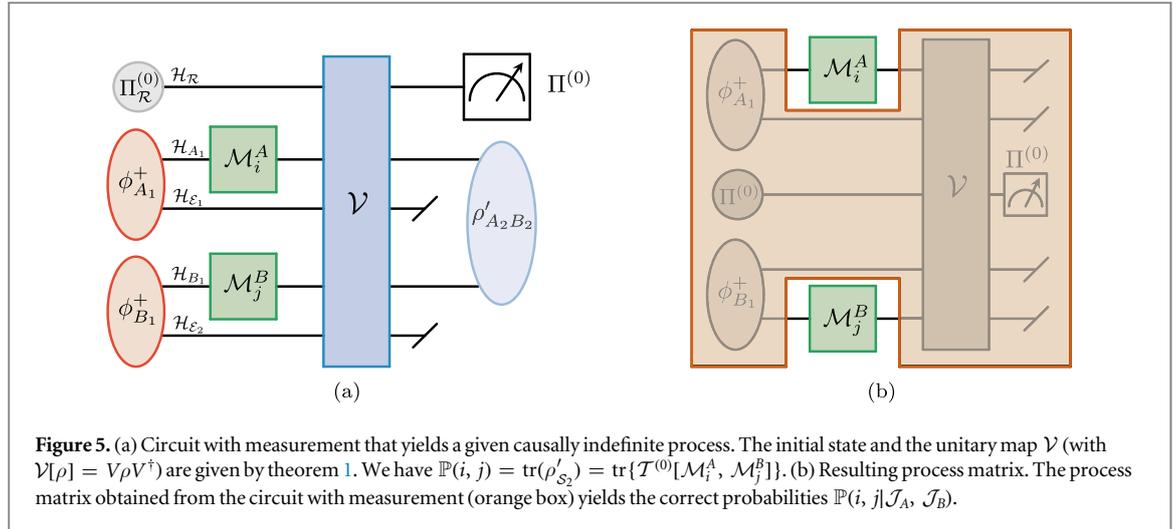## 4. Conditional simulation of causally indefinite processes

By definition, causally non-separable processes cannot be realised deterministically as quantum circuits or as probabilistic mixtures of quantum circuits. However, it has been shown that every process matrix can be simulated probabilistically by a quantum circuit with postselection [11–14]. Explicitly, this means that the joint probability distributions $\mathbb{P}(i, j \mid \mathcal{J}_A, \mathcal{J}_B)$, corresponding to a given process matrix $W^{B_2 B_1 A_2 A_1}$, can be simulated by conditioning the probabilities $p(i, j, \mu \mid \mathcal{J}_A, \mathcal{J}_B, \mathcal{J}_C)$, with respect to a successful outcome μ_{succ}, i.e., with respect to a successful outcome of Charlie's measurement[10].

In this section, we provide a direct, constructive proof (in the spirit of the one provided in [11]) that every process matrix can be simulated by a circuit with postselection. With respect to existing results of this type, our

---

[8] Our definition (3) of the Choi state, as well as the ordering of Hilbert spaces differs slightly from [4]. These particular choices have no influence on the results of the paper.

[9] More precisely, positivity of $W^{B_2 B_1 A_2 A_1}$ is *sufficient* for positive probabilities, but not necessary [29]. Demanding $W^{B_2 B_1 A_2 A_1} \geqslant 0$ can be justified under the additional assumption that Alice and Bob can share a maximally entangled state on top of the temporal correlations that are given by $W^{B_2 B_1 A_2 A_1}$ [4, 6].

[10] It is important to note that conditioning on the other measurement outcomes $\mu \neq \mu_{\mathrm{succ}}$ can also lead to valid process matrices (see, e.g. example 1).

Figure 5. (a) Circuit with measurement that yields a given causally indefinite process. The initial state and the unitary map $\mathcal{V}$ (with $\mathcal{V}[\rho] = V\rho V^\dagger$) are given by theorem 1. We have $\mathbb{P}(i, j) = \mathrm{tr}(\rho'_{\mathcal{S}_2}) = \mathrm{tr}\{\mathcal{T}^{(0)}[\mathcal{M}_i^A, \mathcal{M}_j^B]\}$. (b) Resulting process matrix. The process matrix obtained from the circuit with measurement (orange box) yields the correct probabilities $\mathbb{P}(i, j|\mathcal{J}_A, \mathcal{J}_B)$.

construction is useful because it yields a higher probability of success. Subsequently, we will analyse how, and under what circumstances a *valid* process matrix emerges in general from a conditioned circuit. This analysis will then enable an investigation of the necessary resources to simulate a causally inseparable process, that will be carried out in the next section.

4.1. Conditional simulation of arbitrary causally non-separable processes

A circuit like the one depicted in figure 1 yields a joint probability distribution $p(i, j|\mathcal{J}_A, \mathcal{J}_B)$ to obtain the outcomes i and j in Alice's and Bob's laboratory, respectively. Allowing for an additional (given by a fixed instrument $\mathcal{J}_C$) measurement on the environment leads to a joint probability distribution $p(i, j, \mu|\mathcal{J}_A, \mathcal{J}_B, \mathcal{J}_C)$. In what follows, we will omit the explicit dependence on $\mathcal{J}_C$, for compact notation, and to emphasise that we consider Charlie's instrument to be fixed and part of the underlying process. By *conditioning* $p(i, j, \mu|\mathcal{J}_A, \mathcal{J}_B)$ on a measurement outcome on the environment, i.e., by only recording data when the measurement on the environment yields the correct outcome, it is possible to simulate any process—causally non-separable or not. We have the following theorem:

Theorem 1. *Any process matrix $W^{B_2 B_1 A_2 A_1}$ can be simulated by a circuit with an initial state $\rho_{\mathcal{S}_1 \mathcal{E}} = \phi_{A_1}^+ \otimes \phi_{B_1}^+ \otimes \Pi_{\mathcal{R}}^{(0)}$, where and $\Pi_{\mathcal{R}}^{(0)} = |0\rangle\langle 0|$ is a pure state of the environment $\mathcal{R}$. After the instruments $\mathcal{J}_A$ and $\mathcal{J}_B$ are applied, the systems and the environment evolve through the unitary V that satisfies*

$$\mathrm{tr}_{\mathcal{R}}[\Pi_{\mathcal{R}}^{(0)} V] = \sqrt{\lambda_{\max}}^{-1} \sqrt{(W^{B_2 B_1 A_2 A_1})^{\mathrm{T}}}, \tag{14}$$

where $\lambda_{\max}$ is the largest eigenvalue of $W^{B_2 B_1 A_2 A_1}$. The desired process is simulated by measuring the environment in the computational basis and conditioning on the outcome 0. The probability of success is

$$p_{\mathrm{succ}} = \frac{1}{d_{A_1 B_1} \lambda_{\max}}. \tag{15}$$

Before we prove the theorem, we want to emphasise that it is constructive; for any given process matrix, it allows one to find an explicit circuit plus conditioning procedure that will yield the same statistics as the process matrix. This circuit is depicted in figure 5(a), where we delineate between different spaces: the initial system space includes one half of $\phi_{A_1}^+$ and $\phi_{B_1}^+$, i.e., $\mathcal{S}_1 = A_1 B_1$; while, the other half, along with $\mathcal{R}$, makes up $\mathcal{E} = \mathcal{R}\mathcal{E}_1\mathcal{E}_2$.

Proof. Inserting (3) into (11) we obtain $\mathbb{P}(i, j|\mathcal{J}_A, \mathcal{J}_B) = d_{A_1 B_1}\mathrm{tr}\{W^{\mathrm{T}}(\mathcal{M}_i^A \otimes \mathcal{M}_j^B)[\phi_{A_1}^+ \otimes \phi_{B_1}^+]\}$, where we have set $W := W^{B_2 B_1 A_2 A_1}$. Since W^{T} is positive we can think of it as an element of a POVM; by Neumark's theorem [31–33], there is a unitary V, and a projector $\Pi_{\mathcal{R}}^{(0)} := |0\rangle\langle 0|$ such that $\sqrt{\alpha W^{\mathrm{T}}} = \mathrm{tr}_{\mathcal{R}}[\Pi_{\mathcal{R}}^{(0)} V]$, where $\alpha > 0$ is chosen such that $\mathbb{1} - \alpha W^{\mathrm{T}} \geqslant 0$. Putting it together we get

$$\mathbb{P}(i, j|\mathcal{J}_A, \mathcal{J}_B) = \frac{d_{A_1 B_1}}{\alpha}\mathrm{tr}\{\Pi_{\mathcal{R}}^{(0)} V[(\mathcal{M}_i^A \otimes \mathcal{M}_j^B)[\phi_{A_1}^+ \otimes \phi_{B_1}^+] \otimes \Pi_{\mathcal{R}}^{(0)}]V^\dagger\}. \tag{16}$$

The right-hand side of (16) describes a circuit with a measurement on $\mathcal{R}$ in the computational basis that yields 0. The probability to measure 0 on the environment is

$$p(0) = \sum_{i,j} p(i, j, 0) = \frac{\alpha}{d_{A_1 B_1}} \sum_{i,j} \mathbb{P}(i, j) = \frac{\alpha}{d_{A_1 B_1}}, \tag{17}$$

where $p(i, j, 0)$ is the probability to measure i, j and 0. The maximum success probability p_{succ} is hence obtained for $\alpha = \lambda_{\max}^{-1}$. □

This result is robust; due to the linearity of the scheme, a deviation of the experimentally prepared initial state from the desired one of order ϵ leads to a deviation from the desired process matrix (original success probability) of the same order. It is convenient to rewrite equation (16) as

$$\mathbb{P}(i, j | \mathcal{J}_A, \mathcal{J}_B) = \frac{1}{p(0)} \text{tr}(\mathcal{N}^{(0)} \circ \mathcal{V} \circ \mathcal{M}_j^B \circ \mathcal{M}_i^A [\rho_{\mathcal{S}_i \mathcal{E}}]) := \text{tr}(\mathcal{T}^{(0)} [\mathcal{M}_i^A, \mathcal{M}_j^B]), \tag{18}$$

where $\mathcal{N}^{(0)}[\rho_{\mathcal{R}}] := \Pi_{\mathcal{R}}^{(0)} \rho_{\mathcal{R}} \Pi_{\mathcal{R}}^{(0)}$ is the projection on $\mathcal{R}$, which defines the *conditioned process tensor* $\mathcal{T}^{(0)}$. Comparison of (18) and (5) shows that $W^{B_2 B_1 A_2 A_1} = \text{tr}_f \Upsilon^{(0)}$, where $\Upsilon^{(0)}$ is the Choi state of $\mathcal{T}^{(0)}$; every process matrix can be simulated by a conditioned process tensor. For the above scenario, we have

$$\mathbb{P}(i, j | \mathcal{J}_A, \mathcal{J}_B) = \frac{1}{p(0)} p(i, j, 0 | \mathcal{J}_A, \mathcal{J}_B). \tag{19}$$

The simplest implementation of theorem 1 is the implementation where the ancilla is a qubit. Defining $X := W^T / \lambda_{\max}$ and $X_\sharp := \mathbf{1} - X$, a possible unitary V which implements the desired process matrix W, as one of two possible process matrices $\{W, W_\sharp\}$, can be written as

$$V = \sqrt{X} \otimes |0\rangle\langle 0| - \sqrt{X_\sharp} \otimes |0\rangle\langle 1| + \sqrt{X_\sharp} \otimes |1\rangle\langle 0| + \sqrt{X} \otimes |1\rangle\langle 1|. \tag{20}$$

This choice of V is indeed well-defined and unitary, as $[\sqrt{X}, \sqrt{X_\sharp}] = 0$, and $X_\sharp$ is a positive operator. Conditioning on the outcome 0 yields W, whereas conditioning on 1 yields the process matrix $W_\sharp$. The above construction is not restricted to a two-party scenario, but can straightforwardly be generalised to process matrices that apply to an arbitrary number of parties.

We conclude this section by illustrating theorem 1 for an explicit example.

Example 1. In [4] the following process matrix was introduced as an example for a causally non-separable process that can violate a causal inequality:

$$W^{B_2 B_1 A_2 A_1} = \frac{1}{4}\left(\mathbf{1}^{B_2 B_1 A_2 A_1} + \frac{1}{\sqrt{2}}(\sigma_z^{B_1} \otimes \sigma_z^{A_2} + \sigma_z^{B_2} \otimes \sigma_x^{B_1} \otimes \sigma_z^{A_1})\right), \tag{21}$$

where σ_a^X are Pauli matrices on the Hilbert space $\mathcal{H}_X = \mathbb{C}^2$, and we have omitted the respective identity matrices. For this process matrix, we can choose α in (16) to be equal to 2. Consequently:

$$\sqrt{X} = \frac{1}{2}\left(\mathbf{1}^{B_2 B_1 A_2 A_1} + \frac{1}{\sqrt{2}}(\sigma_z^{B_2} \otimes \sigma_x^{B_1} \otimes \sigma_z^{A_1} + \sigma_z^{B_1} \otimes \sigma_z^{A_1})\right),$$

$$\sqrt{X_\sharp} = \frac{1}{2}\left(\mathbf{1}^{B_2 B_1 A_2 A_1} - \frac{1}{\sqrt{2}}(\sigma_z^{B_2} \otimes \sigma_x^{B_1} \otimes \sigma_z^{A_1} + \sigma_z^{B_1} \otimes \sigma_z^{A_1})\right). \tag{22}$$

The corresponding unitary V can be constructed using (20). The probability $p(0)$ of success for the outcome 0 on the environment for this choice of unitary is equal to $1/2$. The process matrix $W_\sharp$ that one would obtain by conditioning on the outcome 1 is given by $W_\sharp = p(1)X_\sharp = \frac{1}{2}\mathbf{1} - W$. Both W as well as $W_\sharp$ are causally non-separable; conditioning his data recording on either of the outcomes 0 or 1 on the environment, Charlie could not ascribe a causal order to Alice and Bob's actions based on the joint probability distributions he obtains. As expected, the average of W and $W_\sharp$ is causally ordered, though.

By construction, the initial state $\rho_{\mathcal{SE}}$ exhibits genuine tripartite entanglement, i.e., it is entangled across all possible bipartitions $\{A_1 : B_1 \mathcal{E}, B_1 : A_1 \mathcal{E}, \mathcal{E} : A_1 B_1\}$. On top of that, it is easy to check, that the constructed V is nonlocal, i.e., it cannot be written as $V^{A_2} \otimes Z^{B_2 \mathcal{E}}, V^{B_2} \otimes Z^{A_2 \mathcal{E}}$ or $V^{B_2 A_2} \otimes Z^{\mathcal{E}}$. For this example, it is even tripartite entangling[11]. We will see in section 5 that both of these properties—initial tripartite entanglement and a nonlocal unitary—are necessary requirements for the simulation of causally non-separable process matrices.

4.2. Conditional circuits and valid processes
In the previous section, we have shown that every causally non-separable process can be obtained via a circuit with measurement. On the other hand, not every circuit with measurement yields a valid process. The following theorem fixes the set of circuits that lead to a valid process matrix when conditioned on an outcome μ on the environment:

---

[11] In general, nonlocality of a unitary operation is necessary for it to be entangling, but not sufficient.

Theorem 2. *A circuit with measurement on the environment yielding outcome μ leads to a valid process matrix if the success probability $p(\mu)$ does not depend on the choices of instruments $\mathcal{J}_A$ and $\mathcal{J}_B$.*

In a slightly different context, this was also discussed in [14], where the authors pointed out that valid process matrices can be simulated by two-time states that have the property that the probability rule becomes linear, i.e., probabilities do not depend on the choice of instruments. Here, we provide a direct proof of this statement. Theorem 2 shows that the conditioned simulation of a process matrix is well-defined; since the probability for success does not depend on Alice's and Bob's instruments, the reconstructed process matrix is independent of how Alice and Bob choose to run their respective experiments. Consequently, conditioning is a clear-cut experimental prescription. However, this also means that the respective circuits are highly fine-tuned; an arbitrary circuit with measurement would almost always lead to success probabilities that depend on the choices of instruments, or, put another way, would lead to a reconstructed process that violates local causality.

Proof. The probability to measure i, j and μ in a run of a general two-step circuit with measurement on the environment is given by:

$$p(i, j, \mu | \mathcal{J}_A, \mathcal{J}_B) = \text{tr}\{\text{tr}_f(\widetilde{\Upsilon}^{(\mu)})[(M_i^{A_2 A_1})^{\text{T}} \otimes (M_j^{B_2 B_1})^{\text{T}}]\}, \tag{23}$$

where $\widetilde{\Upsilon}^{(\mu)} := p(\mu)\Upsilon^{(\mu)}$ is the Choi state of an unnormalised conditioned process tensor, which does not depend on the respective instruments in Alice's and Bob's laboratories.

If the success probability $p(\mu) = \sum_{i,j} p(i, j, \mu | \mathcal{J}_A, \mathcal{J}_B)$ to measure μ is independent of $\mathcal{J}_A$ and $\mathcal{J}_B$, we can define the positive matrix $W^\mu = \text{tr}_f(\widetilde{\Upsilon}^{(\mu)})$. This matrix is also independent of $\mathcal{J}_A$ and $\mathcal{J}_B$ and it is straightforward to see that it is correctly normalised on products of CPTP maps, i.e.,

$$\text{tr}\{W^\mu[(M^{A_2 A_1})^{\text{T}} \otimes (M^{B_2 B_1})^{\text{T}}]\} = \frac{1}{p(\mu)} \sum_{i,j} p(i, j, \mu | \mathcal{J}_A, \mathcal{J}_B) = 1. \tag{24}$$

To prove the converse statement, let the circuit with measurement be such that it yields a valid process matrix W^μ. This means that Alice and Bob—choosing their respective instruments independently—can reconstruct a valid process matrix by only recording data when a measurement on the environment yields the outcome μ. Consequently, the process matrix they reconstruct is the Choi state defined in (23), normalised by the *overall* probability $\tilde{p}(\mu)$ to measure μ, i.e., $W^\mu = \frac{1}{\tilde{p}(\mu)}\text{tr}_f(\widetilde{\Upsilon}^{(\mu)})$. With this, we can show that the probability to measure μ for given instruments $\mathcal{J}_A$ and $\mathcal{J}_B$, given in (18), is independent of the choice of instruments:

$$p(\mu | \mathcal{J}_A, \mathcal{J}_B) = \sum_{i,j} p(i, j, \mu | \mathcal{J}_A, \mathcal{J}_B) = \tilde{p}(\mu)\text{tr}\{W^\mu[(M^{A_2 A_1})^{\text{T}} \otimes (M^{B_2 B_1})^{\text{T}}]\} = \tilde{p}(\mu), \tag{25}$$

where we have used that W^μ is a valid process matrix, i.e., it satisfies (12). □

We have already seen an example of theorem 2 in section 4.1; the probability to measure 0 on the environment in example 1 was independent of the respective instruments and equal to $p(0) = 1/2$. While the above proof only applies to the case of two parties (Alice and Bob), it can straightforwardly be extended to the case of multiple parties.

If a circuit with measurement satisfies theorem 2, we can explicitly compute the resulting process matrix as well as the probability of successful conditioning. A circuit like the one depicted in figure 6 is defined by a triple $\{\rho_{\mathcal{S}_1 \mathcal{E}}, V, \Pi^{(\mu)}\}$ of initial system-environment state, intermediate system-environment unitary and a projection on the environment. We have the following corollary:

Corollary 1. *For a triple $\{\rho_{\mathcal{S}_1 \mathcal{E}}, V, \Pi^{(\mu)}\}$ that yields a valid process matrix W^μ, we have*

$$(W^\mu)^{\text{T}} = \frac{1}{p(\mu)}\text{tr}_{\mathcal{E}}[[\mathbf{1}_{\mathcal{S}_1} \otimes V^\dagger(\mathbf{1}_{\mathcal{S}_2} \otimes \Pi^{(\mu)}) V](\mathbf{1}_{\mathcal{S}_2} \otimes \rho_{\mathcal{S}_1 \mathcal{E}}^{\text{T}_{\mathcal{S}_1}})] \tag{26}$$

and $$p(\mu) = \frac{1}{d_{A_2 B_2}}\text{tr}_{\mathcal{S}_2 \mathcal{E}}\{[V^\dagger(\mathbf{1}_{\mathcal{S}_2} \otimes \Pi^{(\mu)}) V](\mathbf{1}_{\mathcal{S}_2} \otimes \rho_{\mathcal{E}})\}, \tag{27}$$

where $\rho_{\mathcal{E}} = \text{tr}_{\mathcal{S}_1}(\rho_{\mathcal{S}_1 \mathcal{E}})$ and $\cdot^{\text{T}_{\mathcal{S}_1}}$ is the partial transpose with respect to $\mathcal{S}_1$.

The above case of orthogonal projections on the environment is very general; it includes—by Neumark's theorem—all possible POVMs. A more general construction, where $\mathcal{M}^A$ and $\mathcal{M}^B$ act sequentially, follows in a similar manner, and is discussed in section 5.2.
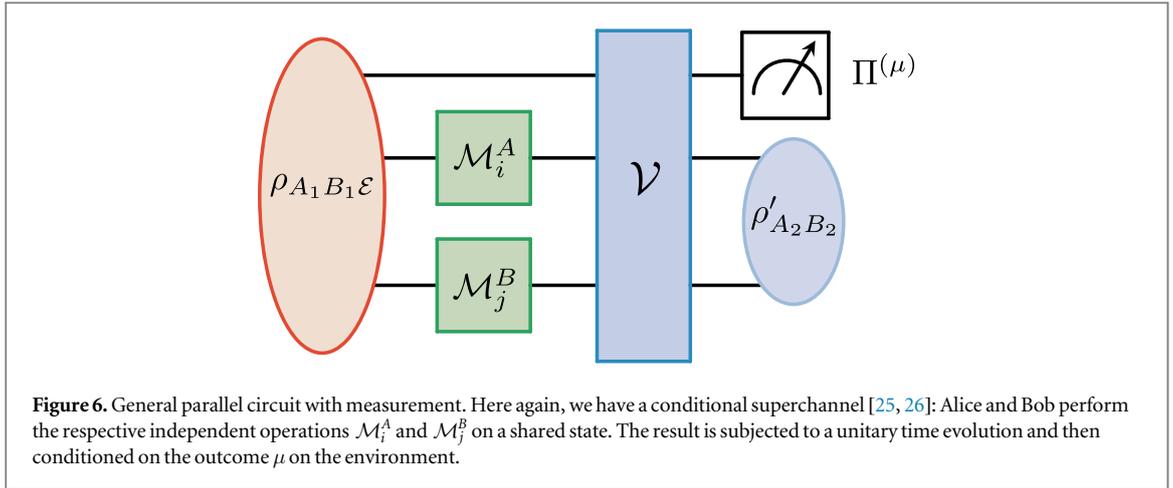
Figure 6. General parallel circuit with measurement. Here again, we have a conditional superchannel [25, 26]: Alice and Bob perform the respective independent operations $\mathcal{M}_i^A$ and $\mathcal{M}_j^B$ on a shared state. The result is subjected to a unitary time evolution and then conditioned on the outcome μ on the environment.

Proof. Using (18), we can write the probability to obtain outcomes i, j and μ in a single run of the experiment as

$$p(i, j, \mu | \mathcal{J}_A, \mathcal{J}_B) = \mathrm{tr}[(\mathbf{1}_{\mathcal{S}_2} \otimes \Pi^{(\mu)}) V (\mathcal{M}_i^A \otimes \mathcal{M}_j^B \otimes \mathcal{I}_{\mathcal{E}})[\rho_{\mathcal{S}_1 \mathcal{E}}] V^\dagger]$$

$$= \mathrm{tr}\{\mathrm{tr}_{\mathcal{E}}[(\mathbf{1}_{\mathcal{S}_1} \otimes V^\dagger (\mathbf{1}_{\mathcal{S}_2} \otimes \Pi^{(\mu)}) V)(\mathbf{1}_{\mathcal{S}_2} \otimes \rho_{\mathcal{S}_1 \mathcal{E}}^{T_{\mathcal{S}_1}})](M_i^{A_2 A_1} \otimes M_j^{B_2 B_1})\}. \tag{28}$$

The resulting process matrix W^μ can be directly read off of (28). By theorem 2, the probability $p(\mu)$ to measure μ on the environment is independent of the instrument. Using $\sum_i M_i^{A_2 A_1} = \frac{1}{d_{A_2}} \mathbf{1}_{A_1 A_2}$, $\sum_i M_i^{B_2 B_1} = \frac{1}{d_{B_2}} \mathbf{1}_{B_1 B_2}$, and $p(\mu) = \sum_{i,j} p(i, j, \mu | \mathcal{J}_A, \mathcal{J}_B)$ yields (27). $\qquad \square$

While the process matrix obtained by conditioning on a particular measurement outcome can be signalling or even causally non-separable, the average process matrix is compatible with a definite causal order, as it must be. From the construction of (28) we have $\sum_\mu p(\mu) W^\mu = \mathrm{tr}_f[\Upsilon] = \mathbf{1}_{\mathcal{S}_2} \otimes \rho_{\mathcal{S}_1}$, where we have used $\sum_\mu \Pi^{(\mu)} = \mathbf{1}_{\mathcal{E}}$ and $\rho_{\mathcal{S}_1} = \mathrm{tr}(\rho_{\mathcal{S}_1 \mathcal{E}})$. This is in agreement with the results from example 1, where we had $p(0) W + p(1) W_\sharp = \frac{1}{2} W + \frac{1}{2} \widetilde{W} = \frac{1}{4} \mathbf{1}_{\mathcal{S}_1 \mathcal{S}_2}$ and $\rho_{\mathcal{S}_1} = \frac{1}{4} \mathbf{1}_{\mathcal{S}_1}$.

The proof of corollary 1 suggests that a circuit with measurement has to be highly fine-tuned to yield a valid process matrix. Since the valid process matrices belong to a lower-dimensional vector space *random* choice of $\{V, \rho_{\mathcal{S}_1 \mathcal{E}}, \Pi^{(\mu)}\}$ will almost always lead to a matrix W^μ that violates local causality. For the case of a properly fine-tuned circuit with measurement, the form (26) allows us to investigate what resources are necessary to simulate process matrices with indefinite causal order. We will carry out this investigation in section 5.

4.3. Probability of success

The probability of success for simulating a process matrix depends—amongst others—on its causal structure and the protocol that is employed for its implementation [18]. Using the scheme provided in section 4.1, we can show the following notable property:
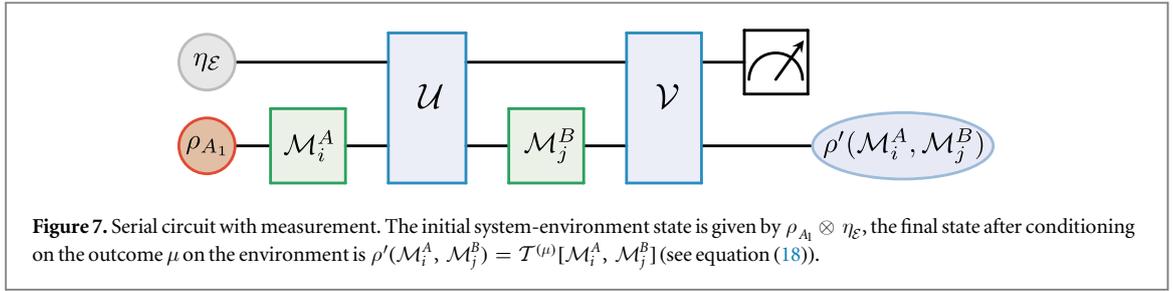
Remark 1. With the protocol of theorem 1, the success probability for the implementation of a process matrix that violates a causal inequality can exceed $1/2$.

We illustrate this fact with an explicit example.

Proof. It has been shown in [34] that the process matrix of example 1 can be mixed with a certain amount of *white noise* and still be causally non-separable. In detail, the process matrix

$$W' = \frac{\gamma}{4(\gamma + 1)} \mathbf{1} + \frac{1}{1 + \gamma} W_{\mathrm{OCB}} \quad \text{is causally non-separable for } \gamma \in [0, \sqrt{2} - 1), \tag{29}$$

where W_{OCB} is the process matrix defined in (21). In order to be able to implement W' with success probability $p = (d_{A_1} d_{A_2})^{-1} \alpha$ according to the procedure provided in section 4.1, the relation $\mathbf{1} - \alpha W' \geqslant 0$ has to hold. The minimal eigenvalue of $\mathbf{1} - \alpha W'$ is equal to $4 + 4\gamma - 2\alpha - \alpha\gamma$. For $\gamma \to \sqrt{2} - 1$, the maximal allowed α tends to $4\sqrt{2}/(1 + \sqrt{2})$. Consequently, there are causally non-separable process matrices that can be implemented with a probability arbitrarily close to $p = \sqrt{2}/(1 + \sqrt{2}) \approx 0.59$. As W' also violates a causal inequality for $\gamma < \sqrt{2} - 1$ [34], this means that there are process matrices that violate causal inequalities and can be implemented with a success probability of more than 50%. $\qquad \square$

Figure 7. Serial circuit with measurement. The initial system-environment state is given by $\rho_{A_1} \otimes \eta_{\mathcal{E}}$, the final state after conditioning on the outcome μ on the environment is $\rho'(\mathcal{M}_i^A, \mathcal{M}_j^B) = \mathcal{T}^{(\mu)}[\mathcal{M}_i^A, \mathcal{M}_j^B]$ (see equation (18)).

It is important to contrast this result with the scheme for the simulation of process matrices proposed in [13]; independent of the process matrix that is to be simulated, this scheme always yields a success probability of 1/16. We now show that both entanglement, nonlocal operations, and non-Markovian features are needed to simulate process matrices that are causally non-separable.

5. Resources for causally non-separable process matrices

The constructive procedure presented in section 4.1 to simulate any process matrix W via conditioning requires both genuine tripartite entanglement, as well as a nonlocal unitary V. In this section, we will show our main result: the simulation of a causally non-separable process matrix via conditioning *always* requires both a genuinely entangled initial state, as well as a nonlocal unitary—no matter the strategy. In a first step, we will prove this statement for the special case of the circuit depicted in figure 6, which we will—for obvious reasons— call the *parallel case*. This circuit is a special case of a circuit including two unitary evolutions (depicted in figure 7), and it is natural to ask if the requirement of initial entanglement can be lifted, if two intermediary evolutions are available. We show in section 5.2 that this is not the case, and initial entanglement and nonlocal unitaries are indeed crucial for the simulation of a causally non-separable process matrix.

5.1. The parallel case
The parallel case is described by a triple $\{\rho_{\mathcal{S}_1\mathcal{E}}, V, \Pi^{(\mu)}\}$, i.e., an initial total state, an intermediary unitary dynamics and a conditioning on the environment. Possible resources for the simulation of a causally non-separable process matrix are the initial state $\rho_{\mathcal{S}_1\mathcal{E}}$, as well as the unitary V. We have the following theorem:

Theorem 3. *For the conditional simulation of a causally non-separable process matrix, it is necessary that both the initial state $\rho_{\mathcal{S}_1\mathcal{E}}$ is genuinely tripartite entangled as well as the unitary matrix V is nonlocal, i.e., it cannot be written as a product operation in any possible bipartition.*

Proof. To prove the first part of the theorem, let V be an arbitrary unitary matrix, $\Pi^{(\mu)} \in \mathcal{L}(\mathcal{H}_{\mathcal{E}})$ an arbitrary orthogonal projection on the environment, and let the initial system-environment state be of the form $\rho_{\mathcal{S}_1\mathcal{E}} = \rho_{\mathcal{S}_1} \otimes \xi_{\mathcal{E}}$. We define $\Gamma^{(\mu)} := V^{\dagger}(\mathbf{1}_{\mathcal{S}_2} \otimes \Pi^{(\mu)})V$. Recall that $\mathcal{S}_2$ denotes the output Hilbert spaces of Alice and Bob, while V acts on these output spaces as well as the environment $\mathcal{E}$ (see figure 6). Consequently, $\Gamma^{(\mu)} \in \mathcal{L}(\mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2} \otimes \mathcal{H}_{\mathcal{E}})$; the resulting process matrix W^{μ} follows from (26) and is given by

$$W^{\mu} = \frac{1}{p(\mu)}\{[\mathrm{tr}_{\mathcal{E}}(\Gamma^{(\mu)}(\mathbf{1}_{\mathcal{S}_2} \otimes \xi_{\mathcal{E}}))]^{\Gamma} \otimes \rho_{\mathcal{S}_1}\} := \Theta_{\mathcal{S}_2} \otimes \rho_{\mathcal{S}_1}. \tag{30}$$

Local causality forbids terms of the form A_2B_2, B_2 or A_2 to appear in the process matrix (see appendix). If $\Theta_{\mathcal{S}_2}$ is not proportional to $\mathbf{1}_{\mathcal{S}_2}$, one of these terms is bound to appear in W^{μ}. With a meaningfully chosen conditioning, the process matrix is then of the form $\mathbf{1}_{\mathcal{S}_2} \otimes \rho_{\mathcal{S}_1}$, which is causal (non-signalling).

A similar argument holds for the case $\rho_{\mathcal{S}_1\mathcal{E}} = \rho_{A_1} \otimes \xi_{B_1\mathcal{E}}$. For this case, we have

$$W^{\mu} = \frac{1}{p(\mu)}\{[\mathrm{tr}_{\mathcal{E}}(\Gamma^{(\mu)}(\mathbf{1}_{\mathcal{S}_2} \otimes (\xi_{B_1\mathcal{E}})^{\mathrm{T}_{B_1}}))]^{\Gamma} \otimes \rho_{A_1}\} := \omega_{B_2B_1A_2} \otimes \rho_{A_1}. \tag{31}$$

Local causality forbids terms of the form B_2, A_2B_2, B_2B_1 and $A_2B_2B_1$. This forces $\omega_{B_2B_1A_2}$ to be of the form $\omega_{B_2B_1A_2} = \mathbf{1}_{B_2} \otimes \zeta_{B_1A_2}$ (where $\zeta_{B_1A_2} \in \mathcal{L}(\mathcal{H}_{B_1} \otimes \mathcal{H}_{A_2})$), which leads to a causally separable process matrix (Alice goes before Bob). The same argument applies for an initial state of the form $\rho_{\mathcal{S}_1\mathcal{E}} = \rho_{B_1} \otimes \xi_{A_1\mathcal{E}}$. Consequently, any initial state $\rho_{\mathcal{S}_1\mathcal{E}}$ of the form

$$\rho_{\mathcal{S}_1\mathcal{E}} = p\,\rho_{A_1B_1} \otimes \xi_{\mathcal{E}} + q\,\rho_{A_1} \otimes \xi_{B_1\mathcal{E}} + (1 - p - q)\,\rho_{B_1} \otimes \xi_{A_1\mathcal{E}}, \qquad \forall\ p, q, (q + p) \in [0, 1] \tag{32}$$

does not lead to a causally non-separable process matrix.

To prove the second part of the theorem, let $\rho_{\mathcal{S}_1\mathcal{E}}$ be an arbitrary state and $V = V^{A_2 B_2} \otimes U^{\mathcal{E}}$ a unitary of product form. The resulting process matrix is given by

$$W^\mu = \frac{1}{p(\mu)}\{\mathbf{1}_{\mathcal{S}_2} \otimes [\text{tr}_{\mathcal{E}}((\mathbf{1}_{\mathcal{S}_1} \otimes \widetilde{\Pi}^{(\mu)})\rho_{\mathcal{S}_1\mathcal{E}}^{\text{T}_{\mathcal{S}_1}})]^{\text{T}}\}, \tag{33}$$

where $\widetilde{\Pi}^{(\mu)} = U^{\mathcal{E}\dagger}\Pi^{(\mu)}U^{\mathcal{E}}$. This is obviously a causal process matrix (non-signalling).

If V is of the form $V = V^{A_2} \otimes U^{B_2\mathcal{E}}$, we obtain the following process matrix:

$$W^\mu = \frac{1}{p(\mu)}\{\mathbf{1}_{A_2} \otimes [\text{tr}_{\mathcal{E}}(\Xi^{(\mu)}(\mathbf{1}_{B_2} \otimes \rho_{\mathcal{S}_1\mathcal{E}}^{\text{T}_{\mathcal{S}_1}}))]^{\text{T}}\} := \mathbf{1}_{A_2} \otimes \eta_{A_1 B_1 B_2} \tag{34}$$

where $\Xi^{(\mu)} = U^{B_2\mathcal{E}\dagger}(\mathbf{1}_{B_2} \otimes \Pi^{(\mu)})U^{B_2\mathcal{E}}$. Again, this process matrix is causally ordered (it allows signalling from Bob to Alice only). A similar argument holds for total unitaries of the form $U = U^{B_2} \otimes V^{A_2\mathcal{E}}$. Consequently only non-product unitaries lead to causally non-separable process matrices. □

In agreement with the results in section 4.1 genuine tripartite entanglement does not mean that Alice and Bob have to initially share entanglement amongst each other. However, the total state of the environment, Alice, and Bob has to be entangled in any possible bipartition. Genuine tripartite entanglement in the initial state constitutes a quantum memory of the past, that can be used to implement a causally non-separable process. In other words, pre-shared quantum memory is a crucial resource for the simulation of causal non-separability.

Non-product unitaries are signalling [35–37], which makes the above theorem perspicuous; causal non-separability can only be simulated if a resource is available that enables communication between Alice, Bob and the environment. Such a unitary propagates the initial memory in a detectable way. Consequently, it is the *non-Markovianity* of the underlying circuit that enables the simulation of causally indefinite processes. Having these results for the parallel case at hand, we now discuss, if the requirements of initial entanglement and nonlocal unitaries can be relaxed if two intermediate unitary dynamics are available.

5.2. Serial case

In the previous sections, we have analysed the implementation of causally non-separable processes by means of a parallel circuit with additional conditioning. Obviously, if we allow for *any* possible initial state, the parallel circuit is a special case of the *serial* one (depicted in figure 7), i.e., a circuit with two intermediary unitaries. It is hence natural to ask, if a serial circuit with measurement allows us to relax the requirement of initial tripartite entanglement and nonlocality of the system-environment unitaries for the simulation of causally non-separable process matrices. This question is answered by the following theorem:

Theorem 4. *The conditional simulation of a causally non-separable process matrix with a serial circuit requires initial system-environment entanglement and nonlocal intermediate system-environment unitaries.*

Proof. Let $\rho_{\mathcal{S}\mathcal{E}} = \rho_{A_1} \otimes \eta_{\mathcal{E}}$ be the initial system-environment state and let $\mathcal{U}$ and $\mathcal{V}$ be arbitrary system-environment unitary maps. The final system state obtained by conditioning on the outcome μ on the environment, given $\mathcal{M}_i^A$ and $\mathcal{M}_j^B$, is (see figure 7):

$$\rho'(\mathcal{M}_i^A, \mathcal{M}_j^B) = \mathcal{T}^{(\mu)}[\mathcal{M}_i^A, \mathcal{M}_j^B] = \text{tr}_{\mathcal{E}}(\widetilde{\mathcal{N}}^{(\mu)} \circ \mathcal{V} \circ \mathcal{M}_j^B \circ \mathcal{U} \circ \mathcal{M}_i^A[\rho_{A_1} \otimes \eta_{\mathcal{E}}]). \tag{35}$$

Analogous to the definition of the conditioned process tensor (with $\widetilde{\mathcal{N}}^{(\mu)} = \frac{1}{p(\mu)}\mathcal{N}^{(\mu)}$, see (18)), this equation can be rewritten in terms of a (non-deterministic) supermap $\widetilde{S}^{(\mu)}$ acting on the CP map $\mathcal{M}_j^B$ (see [11] and figure 8):

$$\rho'(\mathcal{M}_i^A, \mathcal{M}_j^B) = \frac{1}{p(\mu)}(\widetilde{S}^{(\mu)}[\mathcal{M}_j^B])[\mathcal{M}_i^A[\rho_{A_1}]], \tag{36}$$

where $\widetilde{S}^{(\mu)}$ is a CP map (in the sense of [11]), and $\widetilde{S}^{(\mu)}[\mathcal{M}_j^B]: \mathcal{L}(\mathcal{H}_{A_2}) \rightarrow \mathcal{L}(\mathcal{H}_{B_2'})$. We distinguish between $\mathcal{H}_{B_2'}$ and $\mathcal{H}_{B_2}$ for notational purposes only, i.e., $\mathcal{H}_{B_2'} \cong \mathcal{H}_{B_2}$.

Let $S^{(\mu)}$ denote the analogous CP map to $\widetilde{S}^{(\mu)}$ on the level of Choi states, i.e. $S^{(\mu)}[M_j^{B_2 B_1}] \in \mathcal{L}(\mathcal{H}_{B_2'}) \otimes \mathcal{L}(\mathcal{H}_{A_2})$ is the Choi state of $\widetilde{S}^{(\mu)}[\mathcal{M}_j^B]$. With this, (36) can be written as

$$\rho'(\mathcal{M}_i^A, \mathcal{M}_j^B) = \frac{1}{p(\mu)}\text{tr}_{A_2}[(\mathbf{1}_{B_2'} \otimes \mathcal{M}_i^A[\rho_{A_1}]^{\text{T}})S^{(\mu)}[M^{B_2 B_1}]]. \tag{37}$$

Defining the Choi state of the map $S^{(\mu)}$ as $\$^{(\mu)} \in \mathcal{L}(\mathcal{H}_{B_2'} \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2} \otimes \mathcal{H}_{B_1})$, we can rewrite the action of $S^{(\mu)}$ via (see (4)):
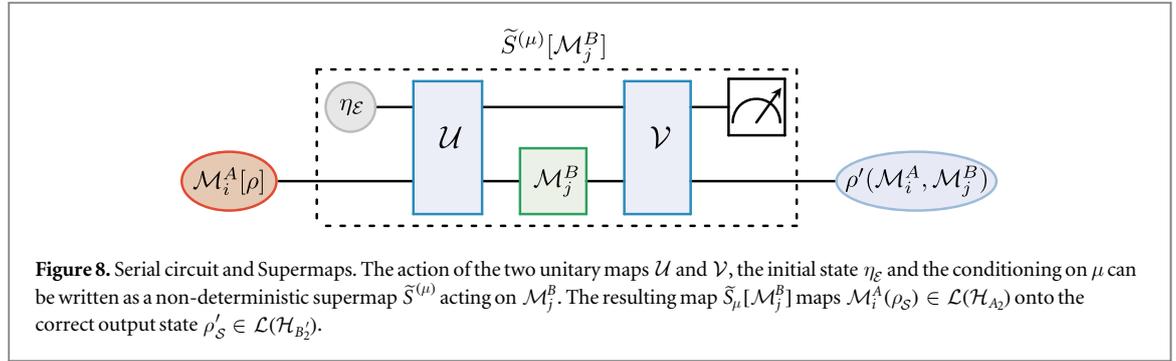
Figure 8. Serial circuit and Supermaps. The action of the two unitary maps $\mathcal{U}$ and $\mathcal{V}$, the initial state $\eta_\mathcal{E}$ and the conditioning on μ can be written as a non-deterministic supermap $\widetilde{S}^{(\mu)}$ acting on $\mathcal{M}_j^B$. The resulting map $\widetilde{S}_\mu[\mathcal{M}_j^B]$ maps $\mathcal{M}_i^A(\rho_\mathcal{S}) \in \mathcal{L}(\mathcal{H}_{A_2})$ onto the correct output state $\rho'_\mathcal{S} \in \mathcal{L}(\mathcal{H}_{B_2'})$.

$$S^{(\mu)}(M_j^{B_2 B_1}) = \mathrm{tr}_{B_2 B_1}[(\mathbb{1}_{B_2' A_2} \otimes (M_j^{B_2 B_1})^\mathrm{T}) \$^{(\mu)}]. \tag{38}$$

With this, (37) reads

$$\rho'(\mathcal{M}_i^A, \mathcal{M}_j^B) = \frac{1}{p(\mu)} \mathrm{tr}_{\mathcal{S}_2 \mathcal{S}_1}\{[\rho_{A_1}^\mathrm{T} \, M_i^{A_2 A_1}]^{\mathrm{T}_{A_2}} [(M_j^{B_2 B_1})^\mathrm{T} \, \$^{(\mu)}]\}. \tag{39}$$

Hence, we obtain

$$p(i, j|\mathcal{J}_A, \mathcal{J}_B, \mu) = \frac{1}{p(\mu)} \mathrm{tr}\{(\mathrm{tr}_{B_2'} \$^{(\mu)} \otimes \rho_{A_1})[(M_i^{A_2 A_1})^\mathrm{T} \otimes (M_j^{B_2 B_1})^\mathrm{T})]\}, \tag{40}$$

which means that for the serial case with initial product state the resulting process matrix is of the form

$$W^\mu = \frac{1}{p(\mu)} \mathrm{tr}_{B_2'} \$^{(\mu)} \otimes \rho_{A_1}. \tag{41}$$

This process matrix has exactly the same form as (31), the process matrix obtained for the parallel case with a separable initial state. It is causally separable (only allowing signalling from Alice to Bob) for the same reasons. Consequently, any separable initial state $\rho_{\mathcal{S}\mathcal{E}}$ will lead to a causally non-separable process matrix.

The necessity of nonlocal unitaries U and V can be proven in a similar way as in theorem 3. Let $V = V^{B_2} \otimes Z^\mathcal{E}$. Rewriting (35) in terms of Choi states, it is straightforward to see that the resulting process matrix is of the form $W^\mu = \mathbb{1}_{B_2} \otimes \upsilon_{B_1 A_2 A_1}$, which is causal (Alice $\rightarrow$ Bob). Analogously, a product unitary $U = U^{A_2} \otimes Q^\mathcal{E}$ leads to a process matrix of the form $W^\mu = \nu_{B_1 A_2} \otimes \varrho_{B_2 A_1}$, where $\nu_{B_1 A_2}$ is the Choi state of a unitary map. Up to normalisation, $\nu_{B_1 A_2}$ is a maximally entangled state, which implies that terms of the form $B_1 A_2$ appear in its decomposition. As terms of the form $B_2 B_1 A_2$ and $B_2 B_1 A_2 A_1$ cannot appear in W^μ, this implies that $\varrho_{B_2 A_1} = \mathbb{1}_{B_2} \otimes \varphi_{A_1}$, which means that the resulting process matrix is causal (Alice $\rightarrow$ Bob). $\qquad\square$

As for the parallel case, the nonlocality of the system-environment unitaries is perspicuous. If the first unitary was of product form, local causality in Alice's laboratory would automatically dictate a global order between the two laboratories. Nonlocality of the final system-environment unitary enables communication between Bob and the environment, which is necessary to 'blur' the causal order between Alice and Bob.

As for the serial case, the theorem shows the importance of genuine pre-existing quantum memory, and system-environment unitaries that transport memory in a detectable way. This implies the following straightforward Corollary:

Corollary 2. *Independent of the strategy, a* Markovian *process is not sufficient for the conditional simulation of a causally non-separable processes.*

For a process—like, e.g., the one depicted in figure 3—that does not allow to store information in the environment and access it at a later time, local causality fixes the global temporal order. The simulation of processes with indefinite causal order via conditioning requires underlying non-Markovian dynamics, i.e., dynamics that display detectable memory effects.

6. Conclusions

Quantum mechanics is compatible with the existence of processes without a definite causal order. To date, however, no such processes has been found in nature or has been realised experimentally, besides the quantum switch [1–3]. In this article, we proposed a way to simulate every causally unordered process through a causally ordered circuit with postselection. With respect to previous results of this type, we have found a simulation

strategy that ensures a higher probability of success, facilitating the experimental observation of causal anomalies.

The simulation of causally unordered processes can be obtained by a simple circuit with measurement on the environment. This simulation works also for process matrices that would be forbidden if purification postulates were imposed [10]. It is important to note that—in contrast to the results of [13, 14]—the conditioning in our scheme happens on the environment, and not on the outputs of Alice and Bob; Charlie can decide whether or not to record data, without having direct access to Alice's or Bob's degrees of freedom. Additionally, beyond the proof of existence, we provided a *constructive* way to obtain a triple $\{\rho_{\mathcal{S}_1\mathcal{E}}, V, \Pi^{(\mu)}\}$ of initial state, unitary evolution and measurement outcome on the environment that yields a given process matrix W.

Even though conditioning seems like a cherry-picking of data to obtain statistics that display causal anomalies, it is not a mathematical post-processing procedure, performed offline, but an experimental procedure; data is collected, whenever the measurement on the environment yields the correct outcome. The whole procedure is well-defined, as the probability for successful conditioning does not depend on the choices of instruments. In a slightly different context, this has also been noted in [14], where the authors showed that valid process matrices can be simulated by two-time states that have the property that the probability rule becomes linear, i.e., probabilities do not depend on the choice of instruments. This understanding of the conditioning process makes causality become an emergent average property. For example, for the conditioning process presented in example 1, both process matrices W and $W_\sharp$ obtained by conditioning on the two possible outcomes 0 and 1 are causally non-separable, but their average $p(0)\,W + p(1)\,W_\sharp$ is—as it should be—causally ordered.

The simulation of causally unordered processes is highly non-unique. A randomly chosen triple $\{\rho_{\mathcal{S}_1\mathcal{E}}, V, \Pi^{(\mu)}\}$ almost always leads to a matrix W^μ that violates local causality. Put differently, there are spatial correlations that cannot be understood as temporal correlations [38]. We have provided a necessary and sufficient condition for a conditioned circuit to yield a valid process matrix. These results also show that, should this kind of conditioning actually happen in nature, it is a highly fine-tuned process.

Finally, we analysed in detail the resources necessary to implement a causally non-separable process matrix via a circuit with conditioning. Our results show that the implementation of causally unordered processes requires both genuine tripartite entanglement in the initial state as well as nonlocal unitary dynamics. The requirement of initial entanglement cannot even be lifted if we allow for more nonlocal communication. Initial entanglement represents a genuine quantum memory of the past, while a nonlocal unitary dynamics allows for a detectable propagation of this quantum memory. In this sense the obtained results—loosely speaking[12]—establish that only genuinely quantum non-Markovian processes allow for the simulation of causally non-separable processes via conditioning. This result, however, only holds for the two-party case; if more parties are involved, causal inequalities can be violated with purely classical processes [39].

Our results provide a complete picture of the resources that go into the simulation of (two party) causally non-separable processes via conditioning. The success probability $p(\mu)$ of the implementation depends on the respective choice of circuit (but not on the choice of instruments). It remains an open question if the maximum success probability for the serial case is—except for trivial cases—always strictly larger than for the parallel case. This is certainly true for process matrices that allow for one-way signalling; they can be simulated deterministically in the serial case, but require conditioning in the parallel one. Determining the relation between signalling and the maximum success probability is an interesting avenue of future research, which we plan to explore in a future work.

---

[12] A generally agreed upon definition of non-Markovianity in the quantum regime is still subject of debate.

Appendix. Allowed terms in the process matrix $W^{B_2 B_1 A_2 A_1}$

Process matrices must respect local causality. This requirement is expressed explicitly in (12). The process matrix $W^{A_2 A_1 B_2 B_1}$ is positive—and hence Hermitian. Consequently, it can be written in the form [4]

$$W^{B_2 B_1 A_2 A_1} = \sum_{\alpha\beta\gamma\epsilon = 0} w_{\alpha\beta\gamma\epsilon}\, \sigma_\alpha^{B_2} \otimes \sigma_\beta^{B_1} \otimes \sigma_\gamma^{A_2} \otimes \sigma_\epsilon^{A_1} \tag{A1}$$

where the matrices $\{\sigma_a^{X_y}\}_{a=0}^{d_{X_y}^2 - 1}$ are generalised Pauli matrices, i.e., they are traceless (except for $\sigma_0^{X_y} = \mathbf{1}_{X_y}$) and $\mathrm{tr}(\sigma_a^{X_y} \sigma_b^{X_y}) = d_{X_y} \delta_{ab}$. The prefactor w_{0000} is equal to $\frac{1}{d_{A_1 B_1}}$ for correct normalisation. Not all positive matrices $W^{A_2 A_1 B_2 B_1}$ of the form (A1) satisfy the requirement (12) for local causality; in order for (12) to hold for all CPTP maps $M^{B_2 B_1}$ and $M^{A_2 A_1}$, $W^{A_2 A_1 B_2 B_1}$ can only contain terms that do not appear in $(M^{A_2 A_1})^{\mathrm{T}} \otimes (M^{B_2 B_1})^{\mathrm{T}}$ (except for $\mathbf{1}_{A_2 A_1 B_2 B_1}$). Otherwise, it would always be possible to find two valid CPTP maps, such that (12) is violated [4].

Using the property $\mathrm{tr}_{X_2}(M^{X_2 X_1}) = \mathbf{1}_{X_1}$ of CPTP maps, we can explicitly write down conditions that define the terms that can appear in the decomposition (A1). In a concise notation, we have

$$\mathrm{tr}(\sigma_\alpha^{X_2} W^{\mathrm{T}}) = 0, \quad \mathrm{tr}[(\sigma_\alpha^{X_2} \otimes \sigma_\beta^{Y_1}) W^{\mathrm{T}}] = 0 \quad \mathrm{tr}[(\sigma_\alpha^{X_2} \otimes \sigma_\beta^{X_1}) W^{\mathrm{T}}] = 0,$$
$$\mathrm{tr}[(\sigma_\alpha^{X_2} \otimes \sigma_\beta^{X_1} \otimes \sigma_\gamma^{Y_2}) W^{\mathrm{T}}] = 0, \quad \text{and} \quad \mathrm{tr}[(\sigma_\alpha^{X_2} \otimes \sigma_\beta^{X_1} \otimes \sigma_\gamma^{Y_2} \otimes \sigma_\epsilon^{Y_1}) W^{\mathrm{T}}] = 0, \tag{A2}$$

where we have omitted the respective identity matrices, $W := W^{B_2 B_1 A_2 A_1}$, α, β, γ, $\epsilon \geqslant 1$, $X, Y \in \{B, A\}$ and $X \neq Y$ when they both appear in the same equation. For simplicity of notation, following the convention of [4], we label terms in the decomposition (A1) of the form $\sigma_\alpha^{A_2} \otimes \mathbf{1}_{A_1 B_2 B_1}$ ($\alpha \geqslant 1$) by A_2, terms of the form $\mathbf{1}_{B_2 B_1} \otimes \sigma_\alpha^{A_2} \otimes \sigma_\beta^{A_1}$ ($\alpha, \beta \geqslant 1$) by $A_2 A_1$, etc. In this notation, for example, the second equation in (A2) states that terms of the form $A_2 A_1$ and $B_2 B_1$ cannot appear in a valid process matrix.

ORCID iDs

Giulio Chiribella ◉ https://orcid.org/0000-0002-1339-0656
Kavan Modi ◉ https://orcid.org/0000-0002-2054-9901

References

[1] Chiribella G, D'Ariano G M, Perinotti P and Valiron B 2013 *Phys. Rev. A* **88** 022318 arXiv:0912.0195
[2] Procopio L M, Moqanaki A, Araújo M, Costa F, Calafell I A, Dowd E G, Hamel D R, Rozema L A, Brukner Č and Walther P 2015 *Nat. Commun.* **6** 8913
[3] Rubino G, Rozema L A, Feix A, Araújo M, Zeuner J M, Procopio L M, Brukner Č and Walther P 2017 *Sci. Adv.* **3** e1602589
[4] Oreshkov O, Costa F and Brukner Č 2012 *Nat. Commun.* **3** 1092
[5] Morimae T 2014 *Phys. Rev. A* **90** 010101
[6] Araújo M, Branciard C, Costa F, Feix A, Giarmatzi C and Brukner Č 2015 *New J. Phys.* **17** 102001
[7] Brukner Č 2015 *New J. Phys.* **17** 083034
[8] Oreshkov O and Giarmatzi C 2016 *New J. Phys.* **18** 093020
[9] Branciard C, Araújo M, Feix A, Costa F and Brukner Č 2016 *New J. Phys.* **18** 013008
[10] Araújo M, Feix A, Navascués M and Brukner Č 2017 *Quantum* **1** 10
[11] Chiribella G, D'Ariano G M and Perinotti P 2008a *Europhys. Lett.* **83** 30004
[12] Chiribella G, D'Ariano G M and Perinotti P 2009 *Phys. Rev. A* **80** 022339
[13] Oreshkov O and Cerf N J 2016 *New J. Phys.* **18** 073037
[14] Silva R, Guryanova Y, Short A J, Skrzypczyk P, Brunner N and Popescu S 2017 *New J. Phys.* **19** 103022
[15] Bennett C H 2005 Simulated time travel, teleportation without communication, and how to conduct a romance with someone who has fallen into a black hole (QUPON)
[16] Lloyd S, Maccone L, Garcia-Patron R, Giovannetti V and Shikano Y 2011a *Phys. Rev. D* **84** 025007
[17] Lloyd S *et al* 2011b *Phys. Rev. Lett.* **106** 040403
[18] Genkina D, Chiribella G and Hardy L 2012 *Phys. Rev. A* **85** 022330
[19] Chiribella G, D'Ariano G M and Perinotti P 2008b *Phys. Rev. Lett.* **101** 060401
[20] Pollock F A, Rodrìguez-Rosario C, Frauenheim T, Paternostro M and Modi K 2018 *Phys. Rev. Lett.* **120** 040405
    Pollock F A, Rodrìguez-Rosario C, Frauenheim T, Paternostro M and Modi K 2018 *Phys. Rev. A* **97** 012127
[21] Costa F and Shrapnel S 2016 *New J. Phys.* **18** 063032
[22] Gutoski G and Watrous J 2007 *Proc. 39th Annual ACM Symp. on Theory of Computing* (New York: ACM) pp 565–74
[23] Hardy L 2012 *Phil. Trans. R. Soc. A* **370** 3385
[24] Milz S, Pollock F A and Modi K 2017 *Open Syst. Inf. Dyn.* **24** 1740016
[25] Modi K 2012 *Sci. Rep.* **2** 581
[26] Ringbauer M, Wood C J, Modi K, Gilchrist A, White A G and Fedrizzi A 2015 *Phys. Rev. Lett.* **114** 090402
[27] Choi M-D 1975 *Linear Algebr. Appl.* **10** 285
[28] Jamiołkowski A 1972 *Rep. Math. Phys.* **3** 275
[29] Barnum H, Fuchs C A, Renes J M and Wilce A 2005 arXiv:0507108
[30] Gutoski G 2009 *Quantum Inf. Comput.* **9** 739
[31] Neumark M 1940 *Izv. Akad. Nauk SSSR Ser. Mat.* **4** 53

[32] Peres A 1995 *Quantum Theory: Concepts and Methods* (Dordrecht: Springer)
[33] Paulsen V 2003 *Completely Bounded Maps and Operator Algebras* (Cambridge: Cambridge University Press)
[34] Feix A, Araújo M and Brukner Č 2016 *New J. Phys.* **18** 083040
[35] Beckman D, Gottesman D, Nielsen M A and Preskill J 2001 *Phys. Rev. A* **64** 052309
[36] Piani M, Horodecki M, Horodecki P and Horodecki R 2006 *Phys. Rev. A* **74** 012305
[37] Bennett C H, Harrow A W, Leung D W and Smolin J A 2003 *IEEE Trans. Inf. Theory* **49** 1895
[38] Costa F, Ringbauer M, Goggin M E, White A G and Fedrizzi A 2017 arXiv:1710.01776
[39] Baumeler Ä, Feix A and Wolf S 2014 *Phys. Rev. A* **90** 042106