



Silva, R., Guryanova, Y., Short, T., Skrzypczyk, P., Brunner, N., & Popescu, S. (2017). Connecting processes with indefinite causal order and multi-time quantum states. *New Journal of Physics*, 19(10), Article 103022. <https://doi.org/10.1088/1367-2630/aa84fe>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1088/1367-2630/aa84fe](https://doi.org/10.1088/1367-2630/aa84fe)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via IOP at <http://iopscience.iop.org/article/10.1088/1367-2630/aa84fe/meta>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

PAPER • OPEN ACCESS

Connecting processes with indefinite causal order and multi-time quantum states

To cite this article: Ralph Silva *et al* 2017 *New J. Phys.* **19** 103022

View the [article online](#) for updates and enhancements.

Related content

- [Causal and causally separable processes](#)
Ognyan Oreshkov and Christina Giarmatzi
- [Indefinite causal structures for continuous-variable systems](#)
Flaminia Giacomini, Esteban Castro-Ruiz and aslav Brukner
- [Operational quantum theory without predefined time](#)
Ognyan Oreshkov and Nicolas J Cerf



PAPER

Connecting processes with indefinite causal order and multi-time quantum states

OPEN ACCESS

RECEIVED
3 March 2017REVISED
26 July 2017ACCEPTED FOR PUBLICATION
8 August 2017PUBLISHED
20 October 2017

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Ralph Silva¹, Yelena Guryanova², Anthony J Short³, Paul Skrzypczyk³, Nicolas Brunner¹ and Sandu Popescu³¹ Group of Applied Physics, University of Geneva, CH-1211 Geneva 4, Switzerland² Institute for Quantum Optics and Quantum Information (IQOQI), Boltzmannngasse 3, A-1090, Vienna, Austria³ H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United KingdomE-mail: ralphsilva@gmail.com**Keywords:** process matrices, multi-time states, post-selection, indefinite causal order, two-time states

Abstract

Recently, the possible existence of quantum processes with indefinite causal order has been extensively discussed, in particular using the formalism of process matrices. Here we give a new perspective on this question, by establishing a direct connection to the theory of multi-time quantum states. Specifically, we show that process matrices are equivalent to a particular class of pre- and post-selected quantum states. This offers a new conceptual point of view to the nature of process matrices. Our results also provide an explicit recipe to experimentally implement any process matrix in a probabilistic way, and allow us to generalize some of the previously known properties of process matrices. Finally we raise the issue of the difference between the notions of indefinite temporal order and indefinite causal order, and show that one can have indefinite causal order even with definite temporal order.

1. Introduction

When describing physical phenomena it is commonly assumed that there exists an underlying causal order. Loosely speaking, later events can be influenced by previous ones, but not the other way around. Recently, however, the idea that physical theories necessarily require a causal order has been challenged, in particular in the context of quantum theory. Indeed, one may imagine that the notion of causal order can be subject to fundamental quantum principles, such as the superposition or the uncertainty principle, resulting in indefinite causal structures. Reference [1] introduced the notion of a ‘superposition of quantum evolutions’, and showed how this could be achieved for arbitrary evolutions via post-selection. Reference [2] considered a situation in which the causal order of two evolutions is controlled by a quantum bit, which can become entangled with the causal structure. By post-selecting the control bit, one can then create a superposition of the two evolutions, each of which has a different causal order [24]. The idea of quantum correlations with indefinite causal order has recently triggered an intense research effort and subsequently several frameworks for characterizing quantum processes with undefined or dynamical causal structures have been developed, see e.g. [2–6].

In [7], Oreshkov *et al* introduced the framework of ‘process matrices’ where causal structure can be partially relaxed. Here, operations in local laboratories are described by quantum theory (making use of a local causal structure), however, no global causal structure is assumed. Crucially, this framework captures situations that cannot be explained by any definite causal structure, as witnessed by the violation of so called ‘causal inequalities’ (analogous to Bell inequalities) [7–12]. Possible implications for quantum information processing have also been discussed [13, 14].

Although the process matrix formalism contains features reminiscent of quantum mechanics, it is not derived from quantum theory. The crucial question of its relation to quantum theory was left open in the original work. It could have been the case that process matrices simply represented a reformulation of some specific situations in quantum mechanics, or that they contained new physics outside of quantum theory.

Here we offer an answer to this question. We show that process matrices correspond to a subclass of ‘two-time states’, which were introduced by Aharonov *et al* [15–17]. This means that the world described by process

matrices is equivalent to a particular case of a quantum world with *fundamental* post-selection, i.e. a quantum world with independent initial and final boundary conditions, which are both guaranteed to occur (as opposed to ordinary post-selection in which the final state occurs only probabilistically).

In [10, 26] an alternative connection between quantum mechanics and process matrices was found, via a powerful generalisation of the process matrix formalism. In particular, as a corollary, it was shown that any process matrix can be simulated in quantum mechanics using post-selection. As any two-time state can be probabilistically implemented experimentally in standard quantum theory via post-selection, we also recover this result in our formalism.

A number of further insights also follow from our results. First, and most importantly, they show that there is a distinction between indefinite temporal order and indefinite causal order, and that one can have indefinite causal order with definite temporal order. Furthermore, they allow us to generalise some of the previously known properties of process matrices, which might be of interest in future work.

Finally, we show that the subset of pre- and post-selected states corresponding to process matrices has very special properties, leading to probabilities that are linear functions of the states and of the measurements. This raises a new and important question: why only this subset? What would be problematic if the probabilities were nonlinear functions, as is the case for most pre- and post-selected states?

2. Pre- and post-selection: experimental versus fundamental

The standard procedure for collecting statistics in quantum mechanics is to fix an initial state for a system and accumulate experimental results based on this chosen state, or perhaps on an evolution of it later in time. As pointed out by Aharonov *et al* in [15], in addition to fixing an initial state for a system, one may also specify an independent final state. One way to realise such a specification is with post-selection, which can be understood by considering the following scenario.

At the initial time t_1 , Alice prepares a quantum system in the state $|\psi\rangle$. In the time interval between the initial and a final time t_2 she performs some arbitrary experiments and records their results. At the final time t_2 , she measures an observable O , one of whose non-degenerate eigenstates is $|\phi\rangle$, the desired final state (which is arbitrary and independent of $|\psi\rangle$). Alice considers her experiment to be successful if the measurement of O yields the eigenvalue corresponding to the eigenstate $|\phi\rangle$; otherwise she discards the experiment.

This way, if Alice repeats her experiment on an ensemble of particles, all prepared in the initial state $|\psi\rangle$, she ends up with a sub-ensemble, which we call a pre- and post-selected ensemble, characterised by the initial state $|\psi\rangle$ and the final state $|\phi\rangle$. In this sub-ensemble the statistics of the results of the intermediate measurements is, in general, different from the statistics over the entire ensemble.

Importantly, the procedure described above is purely quantum mechanical, albeit not the one that is usually considered (the standard paradigm considers only pre-selected ensembles). In this procedure, one cannot guarantee *a priori* that the final measurement will yield the eigenvalue corresponding to $|\phi\rangle$. Thus, at no intermediate point can one know if the post-selection will be successful or not (Alice may even decide not to measure O). It follows that there is also no intermediate time at which Alice can know that she is in the desired pre- and post-selected sub-ensemble. Only after the final time, t_2 , will she know which events are successful and must be kept, and which have to be discarded. It is only then, looking back at the records of her intermediate measurement results, that she is able to find out the statistics corresponding to a specific pre- and post-selected ensemble.

Contrary to this experimental realisation of a pre- and post-selected ensemble, the quantum mechanical formalism also allows one to fix a guaranteed final state. Unlike the case of experimental post-selection, here Alice would already see the statistics corresponding to the pre- and post-selected ensemble at the intermediate times $t_1 < t < t_2$. We call this a situation with ‘fundamental’ post-selection, as opposed to the ordinary, measurement based, probabilistic post-selection.

There is no evidence that fundamental post-selection exists in nature, however, some authors have suggested that it could be present in some exotic situations, for example having a final state of the Universe [18–20], or a final state at the singularity of a black-hole [21].

3. The two-time state formalism

Regardless of whether a two-time state arises from experimental or fundamental post-selection, the statistics generated by a given pre- and post-selected ensemble are the same. First introduced for pure states [15, 16], the formalism was recently extended to general mixtures of two-time states [17]. In this section we will review this formalism following the convenient notations of [17].

Suppose, once more, that Alice starts at t_1 with her system in the state $|\psi\rangle$, measures the operator O and selects the cases in which this final measurement yields the eigenvalue corresponding to $|\phi\rangle$ at time t_2 . For

simplicity we assume that the Hamiltonian between t_1 and t_2 is zero and that in the intermediate time Alice performs a detailed measurement, described by the set of Kraus operators $\{\hat{E}_a = \sum_{k,l} \beta_{a,kl} |k\rangle\langle l|\}$. By ‘detailed measurement’ we mean a measurement where each outcome a corresponds to a single Kraus operator, such that the normalisation condition that the operators obey is $\sum_a \hat{E}_a^\dagger \hat{E}_a = \mathbb{I}$. The probability to obtain the outcome a given the pre- and post-selection is then given by

$$P(a) = \frac{|\langle \phi | \hat{E}_a | \psi \rangle|^2}{\sum_{a'} |\langle \phi | \hat{E}_{a'} | \psi \rangle|^2}. \quad (1)$$

In order to use the two-time formalism to its full advantage, we transfer to the two-time language. Formally, we take the state space of Alice to be the tensor Hilbert space $\mathcal{H}_{\mathcal{A}_2} \otimes \mathcal{H}_{\mathcal{A}_1}$, where $\mathcal{H}_{\mathcal{A}_1}$ is the Hilbert space of the pre-selected states, which are denoted by ket vectors (with raised labels) and that evolve forward in time, and $\mathcal{H}_{\mathcal{A}_2}$ is the Hilbert space of the post-selected states, which are denoted by bra vectors (with lowered labels) and that evolve backward in time. The structure is made explicit by defining a two-time state and a two-time version of the Kraus operator in the following way

$$\Psi = {}_{\mathcal{A}_2} \langle \phi | \otimes | \psi \rangle^{\mathcal{A}_1} \in \mathcal{H}_{\mathcal{A}_2} \otimes \mathcal{H}_{\mathcal{A}_1}, \quad (2)$$

$$E_a = \sum_{kl} \beta_{a,kl} |k\rangle^{\mathcal{A}_2} \otimes {}_{\mathcal{A}_1} \langle l| \in \mathcal{H}_{\mathcal{A}_2} \otimes \mathcal{H}_{\mathcal{A}_1}, \quad (3)$$

where we differentiate between the usual Kraus operator and its two-time version by the presence or absence of a hat.

This notation places states and measurement operators on an equal footing such that they are dual to one another, which allows us to re-write the amplitude $\langle \phi | \hat{E}_a | \psi \rangle$ appearing in (1) as $\Psi \cdot E_a$. The operation (\cdot) applies to vectors belonging to different Hilbert spaces and is a combined composition/contraction: it contracts any dual vector pairs (i.e. bra and ket pairs) in Hilbert spaces with the same labels to generate a scalar⁴, e.g. ${}_{\mathcal{A}_2} \langle \phi | \cdot | i \rangle^{\mathcal{A}_2} = | i \rangle^{\mathcal{A}_2} \cdot {}_{\mathcal{A}_2} \langle \phi | = \langle \phi | i \rangle$ ⁵ and it performs the tensor product between any unpaired vectors. In the present example all the vectors are paired, so \cdot effectively results in a scalar product; the tensoring will be useful in later examples.

The probability to obtain the outcome a , as seen in (1), now takes on the alternative form

$$P(a) = \frac{|\Psi \cdot E_a|^2}{\sum_{a'} |\Psi \cdot E_{a'}|^2}. \quad (4)$$

The advantage of viewing pre- and post-selections as two-time states in the tensor product Hilbert space $\mathcal{H}_{\mathcal{A}_2} \otimes \mathcal{H}_{\mathcal{A}_1}$ is that one can take superpositions of simple ‘direct product’ two-time states and obtain general pure two-time states in which the pre-selection is entangled with the post-selection. An arbitrary pure two-time state has the form

$$\Psi = \sum_{ij} \alpha_{ij} {}_{\mathcal{A}_2} \langle i | \otimes | j \rangle^{\mathcal{A}_1} \in \mathcal{H}_{\mathcal{A}_2} \otimes \mathcal{H}_{\mathcal{A}_1}. \quad (5)$$

This general two-time state can be understood as a fundamentally post-selected state, or one that can be implemented via experimental post-selection using entangled ancillas [22], the procedure for which is detailed in figure 1.

Generalizing further, one may also consider mixtures of pure two-time states [17]. Consider that the experimenter follows a similar procedure as above (see figure 1). However, instead of preparing a pure state for \mathcal{A}_1 and the ancilla C , the experimenter prepares any possible mixed state, i.e. the ensemble $\{\sum_{ij} \alpha_{r,ij} |j\rangle^{\mathcal{A}_1} \otimes |i\rangle_C\}_r$ with associated probabilities p_r .

In this case, the statistics of Alice’s operation obey

$$P(a) = \frac{\sum_r p_r |\Psi_r \cdot E_a|^2}{\sum_{a'} \sum_r p_r |\Psi_r \cdot E_{a'}|^2}, \quad (6)$$

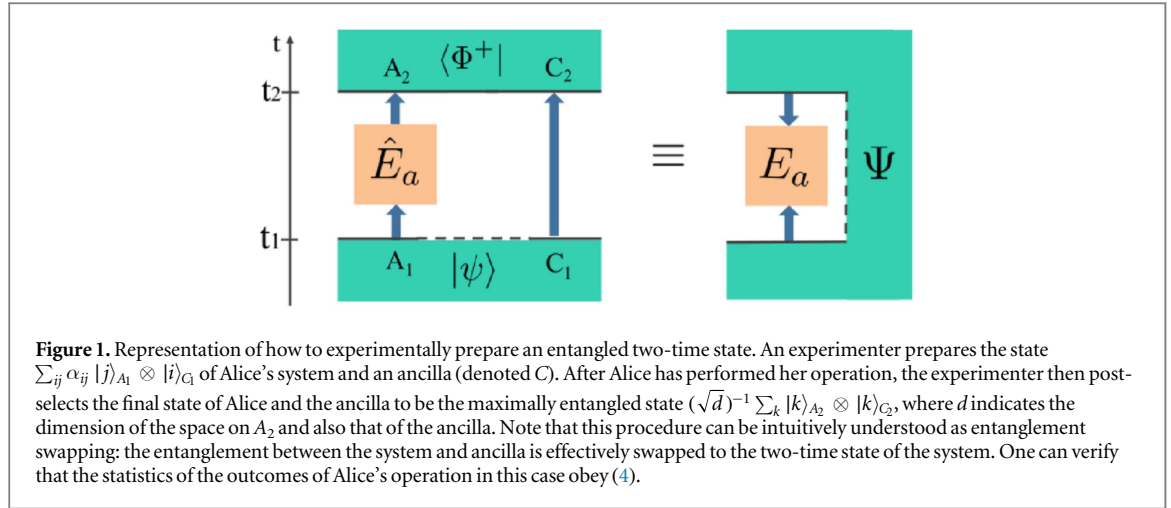
where $\Psi_r = \sum_{ij} \alpha_{r,ij} {}_{\mathcal{A}_2} \langle i | \otimes | j \rangle^{\mathcal{A}_1}$.

Analogous to the case of ordinary quantum mechanics where the density operator captures the information of a mixed state at a single time, one can also construct a similar object for multi-time states. The *density vector* of a pure multi-time state Ψ is given by $\Psi \otimes \Psi^\dagger$, where the Hilbert spaces pertaining to Ψ^\dagger are differentiated by dagger labels, such that every ket (bra) in Ψ is transformed into a daggered bra (ket) in Ψ^\dagger ⁶.

⁴ Note that we follow the convention from general relativity—a contraction can only occur between a raised and a lowered label.

⁵ Since in this framework every Hilbert space (bra or ket) has a distinct label, the order in which the Hilbert spaces are written is arbitrary.

⁶ Note that, given our convention that ket vectors have raised labels and bra vectors lowered labels, the dagger operation also swaps the position of the label.



For the pure state in (5),

$$\begin{aligned} \Psi \otimes \Psi^\dagger &\in \mathcal{H}_{A_2} \otimes \mathcal{H}^{A_1} \otimes \mathcal{H}_{A_1^\dagger} \otimes \mathcal{H}^{A_2^\dagger}, \\ \Psi \otimes \Psi^\dagger &= \sum_{ijmn} \alpha_{ij} \alpha_{mn}^* |i\rangle_{A_2} \langle j| \otimes |j\rangle^{A_1} \otimes \langle n|_{A_1} \otimes |m\rangle^{A_2^\dagger}. \end{aligned} \tag{7}$$

We will use the shorthand $\mathcal{H}^A := \mathcal{H}^A \otimes \mathcal{H}_{A^\dagger}$ and $\mathcal{H}_A := \mathcal{H}_A \otimes \mathcal{H}^{A^\dagger}$ from here on, such that $\Psi \otimes \Psi^\dagger \in \mathcal{H}^{A_1} \otimes \mathcal{H}_{A_2}$.

The density vector of the ensemble is the convex combination of the pure density vectors,

$$\eta = \sum_r p_r \Psi_r \otimes \Psi_r^\dagger \in \mathcal{H}^{A_1} \otimes \mathcal{H}_{A_2}. \tag{8}$$

Similar to the construction of a density vector for states, one can also construct a ‘Kraus density vector’ $J_a = E_a \otimes E_a^\dagger$ for any Kraus operator \hat{E}_a . If the measurement is not detailed with respect to the outcome a (in other words multiple Kraus operators correspond to a single outcome, a so-called ‘coarse-grained’ measurement), then the density vector corresponding to the outcome is the sum of all the Kraus density vectors corresponding to that outcome. If \hat{E}_a^μ denotes the Kraus operator, where a is the outcome and μ is an index running over all the operators corresponding to that outcome,

$$J_a = \sum_\mu E_a^\mu \otimes E_a^{\mu\dagger} \in \mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}. \tag{9}$$

It is important to note that J_a contains all the information about the dynamics induced by the measurement, and not only information about the outcome probabilities. That is, it is not equivalent to the POVM element $\sum_\mu \hat{E}_a^{\mu\dagger} \hat{E}_a^\mu$, which contains no information about the post-measurement state.

The normalisation condition in the standard formalism, $\sum_{a,\mu} \hat{E}_a^{\mu\dagger} \hat{E}_a^\mu = \mathbb{I}$, associated to the fact that the Kraus operators $\{\hat{E}_a^\mu\}_{a,\mu}$ form a completely-positive trace-preserving channel, expressed in our notation is⁷

$$\mathbb{I}_{A_2} \cdot J_{A_1}^{A_2} = \mathbb{I}_{A_1}, \tag{10}$$

where $J_{A_1}^{A_2} = \sum_a J_a$ and we have introduced the ‘identity’ vector $\mathbb{I}_A := \mathbb{I}_A^{A^\dagger} \in \mathcal{H}_A$ given by

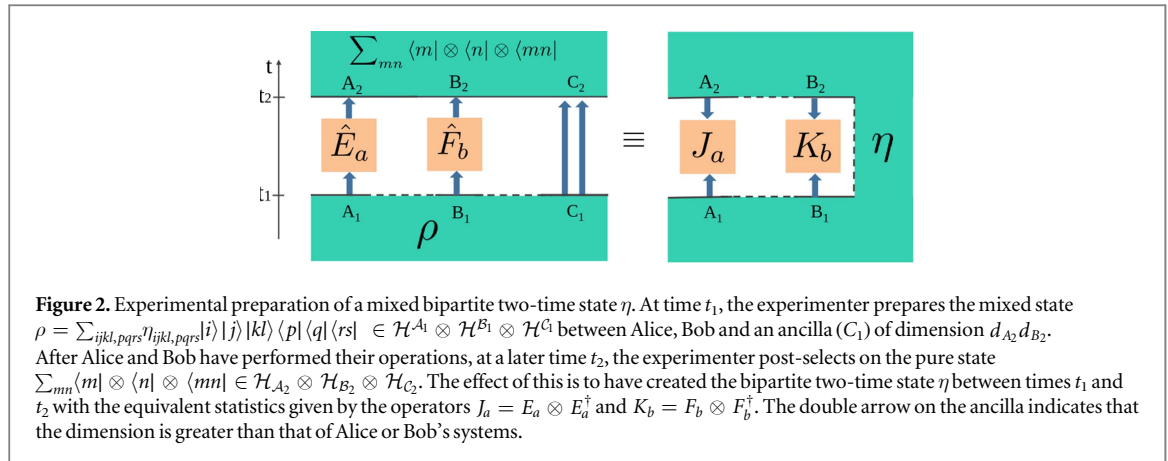
$$\mathbb{I}_A^{A'} = \sum_i |i\rangle^{A'} \otimes \langle i|_A \in \mathcal{H}^{A'} \otimes \mathcal{H}_A. \tag{11}$$

Considering the form of (10), we can think of J that satisfy this condition as *future identity preserving*.

The generalization to mixed states and non-detailed measurements preserves the duality between states and measurements, that is now between the density vectors corresponding to two-time states and the density vectors corresponding to a measurement outcome. The statistics of Alice's operation given by (6) can now be concisely written as

$$P(a) = \frac{\eta \cdot J_a}{\sum_{a'} \eta \cdot J_{a'}}. \tag{12}$$

⁷This representation is the analogue of the condition that if $\Lambda(\cdot) = \sum_{a,\mu} \hat{E}_a^\mu(\cdot)\hat{E}_a^{\mu\dagger}$ is a trace-preserving quantum channel, then the conjugate channel $\Lambda^\dagger(\cdot) = \sum_{a,\mu} \hat{E}_a^{\mu\dagger}(\cdot)\hat{E}_a^\mu$ is *unital*, i.e. satisfies $\Lambda^\dagger(\mathbb{I}) = \mathbb{I}$.



The statistics are thus fully captured by the density vector of the two-time ensemble (η), and the density vectors corresponding to each outcome of the measurement (J_a). Note that the denominator, which ensures that the probabilities sum to one, generally depends upon the choice of measurement. In such cases it is impossible to remove it by merely normalising η ⁸. Note that because of this, the probabilities are in general nonlinear with respect to the state η , as well as the measurement J_a .

Finally, we note that we defined density vectors η as being mixtures of pure density vectors. We can easily show that this condition is equivalent to asking that the state η is ‘positive’, in the sense that for any vector $v \otimes v^\dagger$ we have $\eta \cdot (v \otimes v^\dagger) \geq 0$ (which is the two-time version of the usual definition of a positive operator: A is positive if, for any $|\psi\rangle$, $\langle \psi|A|\psi\rangle \geq 0$). Hence every density vector η is positive and every positive η is a density vector. Furthermore, note that the J , as defined in (9) are positive in the same sense. In particular, this ensures that η produces positive probabilities via (12).

We will be particularly interested in this paper in bipartite two-time states (shared between Alice and Bob), which can be characterised by $\eta \in \mathcal{H}^{A_1} \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}^{B_1} \otimes \mathcal{H}_{B_2}$. (See figure 2 for the preparation of such a state via post-selection.) Given that Alice measures J_a and Bob measures K_b , the joint probability to obtain outcomes a and b is given by

$$P(a, b) = \frac{\eta \cdot (J_a \otimes K_b)}{\sum_{a',b'} \eta \cdot (J_{a'} \otimes K_{b'})}. \tag{13}$$

4. The process matrix formalism

One of the key motivating factors behind the introduction of process matrices in [7] was the ability of the formalism to capture ‘indefinite causal orders’. We will describe the formalism for two parties; the extension to more parties is straightforward and can be found in [11].

The set-up is summarised as follows: Alice and Bob reside in spatially separated laboratories, which are sealed off from the outside world. The doors of Alice’s (Bob’s) lab may open once to let a system in, and once to let a system out. Within their respective labs Alice and Bob may perform local, quantum mechanical measurements (with the aid of local ancillas, should they require them). Finally, Alice and Bob’s labs may be connected in some way. For example it could be the case that on receiving a system, processing it, and then releasing it, Alice passes a quantum system to Bob, who opens his lab doors to let it in (or vice versa). However, they may be connected in completely different ways, corresponding to indefinite causal order, which may not be realisable in quantum mechanics.

The novel element in this formalism is the so-called ‘process matrix’ W which specifies how the two labs are connected, and loosely speaking, plays the role of a joint state for Alice and Bob. More precisely, we associate two Hilbert spaces to Alice, one to the system that enters her laboratory, and one to the system that leaves it, and similarly for Bob. The process matrix is a mathematical object playing the role of a generalised ‘density operator’ acting on these four Hilbert spaces.

On the other hand, measurements in this formalism are usual quantum measurements. Suppose Alice performs a measurement where \hat{E}_a^μ denotes the (possibly non-detailed) Kraus operators corresponding to the outcome a ; the index μ runs over the various operators corresponding to this individual outcome. In the standard quantum formalism, if the state of Alice’s system is ρ , the unnormalised state after the measurement,

⁸ This also means that two density vectors which are identical up to normalization represent the same physical state.

given that the outcome a was obtained, is $\rho_a = \mathcal{M}_a(\rho) = \sum_{\mu} \hat{E}_a^{\mu} \rho \hat{E}_a^{\mu \dagger}$, with $\mathcal{M}_a: \mathcal{L}(\mathcal{H}^{A_1}) \rightarrow \mathcal{L}(\mathcal{H}^{A_2})$ denoting the linear map *from* linear operators on the input Hilbert space *to* linear operators on the output space, thus describing the measurement. We similarly denote Bob's Kraus operators by \hat{F}_b^{ν} and the action of this map as $\mathcal{N}_b(\rho) = \sum_{\nu} \hat{F}_b^{\nu} \rho \hat{F}_b^{\nu \dagger}$.

In the process-matrix formalism the measurements of Alice are represented (via a Choi-Jamiołkowski transformation) by [7, 9]

$$\begin{aligned} M_a^{A_1 A_2} &= [\mathcal{I} \otimes \mathcal{M}_a(|\Phi^+\rangle \langle \Phi^+|)]^{\Gamma} \\ &= \sum_{\mu} [(\mathbb{I} \otimes \hat{E}_a^{\mu}) |\Phi^+\rangle \langle \Phi^+| (\mathbb{I} \otimes \hat{E}_a^{\mu \dagger})]^{\Gamma}, \end{aligned} \quad (14)$$

where $|\Phi^+\rangle = \sum_i |i\rangle |i\rangle \in \mathcal{H}^{A_1} \otimes \mathcal{H}^{A_1}$ is the unnormalized maximally entangled state on two copies of Alice's input Hilbert space, and $\mathcal{M}_a: \mathcal{L}(\mathcal{H}^{A_1}) \rightarrow \mathcal{L}(\mathcal{H}^{A_2})$. This state is written in an arbitrary but fixed orthonormal basis $\{|i\rangle\}$ and Γ is the (full) *transpose* operation taken with respect to this basis. Finally, $\mathcal{I}: \mathcal{L}(\mathcal{H}^{A_1}) \rightarrow \mathcal{L}(\mathcal{H}^{A_1})$, is the identity map. The measurements of Bob are represented analogously, and denoted $N_b^{B_1 B_2}$.

Having defined the measurements in this framework the next step is to define the object on which they act. This object is the process matrix $W^{A_1 A_2 B_1 B_2} \in \mathcal{L}(\mathcal{H}^{A_1} \otimes \mathcal{H}^{A_2} \otimes \mathcal{H}^{B_1} \otimes \mathcal{H}^{B_2})$, and is an operator acting on the input and output Hilbert spaces of Alice and Bob. The probability rule for the measurement outputs is then given in analogy to the Born rule of quantum mechanics,

$$P_W(a, b) = \text{tr}[W^{A_1 A_2 B_1 B_2} (M_a^{A_1 A_2} \otimes N_b^{B_1 B_2})]. \quad (15)$$

For W to be a valid process matrix, the probabilities $P_W(a, b)$ must be positive and normalised, $\sum_{ab} P_W(a, b) = 1$, in addition to W itself being a positive operator. These are the only requirements to be a valid process matrix, and the set of valid process matrices is any operator satisfying these constraints.

5. Connecting process matrices with multi-time states

We are now in position to present our main result, namely to connect process matrices and multi-time states. In particular, we will show that to every process matrix we can associate a two-time state that will produce the same probabilities for all measurements. This shows that every process matrix can be realised within quantum theory, if both pre- and post-selection are allowed.

We will present the mapping in two ways; first in terms of matrix elements and second via an isomorphism between Hilbert spaces. Starting with the former, let us denote an arbitrary process matrix as

$$W^{A_1 A_2 B_1 B_2} = \sum_{\substack{ijkl \\ pqrs}} w_{ijkl, pqrs} |ijkl\rangle \langle pqrs|, \quad (16)$$

where this decomposition is in the same basis as the $|\Phi^+\rangle$ used to define $M_a^{A_1 A_2}$ and $N_b^{B_1 B_2}$. Then, to this process matrix we associate the bipartite two-time state

$$\eta_W = \sum_{\substack{ijkl \\ pqrs}} w_{ijkl, pqrs} \mathcal{A}_2 \langle j| \otimes |i\rangle^{A_1} \otimes \mathcal{A}_1 \langle p| \otimes |q\rangle^{A_2} \otimes \mathcal{B}_2 \langle l| \otimes |k\rangle^{B_1} \otimes \mathcal{B}_1 \langle r| \otimes |s\rangle^{B_2}. \quad (17)$$

Henceforth, we denote a two-time state which has been mapped from a process matrix W as η_W . One can explicitly check that W and η_W lead to the same probabilities for all measurements. We show in appendix A1 that

$$\begin{aligned} P_W(a, b) &= \text{tr}[W^{A_1 A_2 B_1 B_2} (M_a^{A_1 A_2} \otimes N_b^{B_1 B_2})] \\ &= \eta_W \cdot (J_a \otimes K_b), \end{aligned} \quad (18)$$

where $J_a = \sum_{\mu} E_a^{\mu} \otimes E_a^{\mu \dagger}$ are the Kraus density vectors of Alice's measurement, and $K_b = \sum_{\nu} F_b^{\nu} \otimes F_b^{\nu \dagger}$ of Bob's. Since all process matrices satisfy $\sum_{ab} P_W(a, b) = 1$, the above shows that every η_W satisfies

$$\sum_{ab} \eta_W \cdot (J_a \otimes K_b) = 1, \quad (19)$$

and hence

$$\begin{aligned} P_W(a, b) &= \frac{\eta_W \cdot (J_a \otimes K_b)}{\sum_{a'b'} \eta_W \cdot (J_{a'} \otimes K_{b'})} \\ &= P_{\eta_W}(a, b), \end{aligned} \quad (20)$$

which completes the claim, and shows that W and η_W lead to the same probabilities.

From this reasoning we also deduce that every η_W has the special property that *the probability rule becomes linear*.

In particular the probability rule for two-time states in (13) is a nonlinear function with respect to the density vector η and Kraus density vectors J_a and K_b . However, for those η_W which arise through (17), the probability reduces to $P_{\eta_W}(a, b) = \eta_W \cdot (J_a \otimes K_b)$, which is now a linear function of the two-time state, and measurements. We will refer to any state that satisfies the property

$$P_\eta(a, b) = \eta \cdot (J_a \otimes K_b), \quad (21)$$

as a *linear two-time state*. Note that this is equivalent to $\sum_{ab} \eta \cdot (J_a \otimes K_b) = \eta \cdot (J \otimes K) = 1$ for all channels J, K . This also highlights the fact that process matrices only map into a subset of two-time states.

We have shown above that any process matrix maps into a linear two-time state. Conversely, every linear two-time state corresponds to some process matrix W . In particular, if one writes η as in (17) and then considers the corresponding matrix W given by (16) then the positivity of W follows from the positivity of η , and the normalisation of probabilities for any measurement follows from inserting the linearity condition (21) in (18), and noting that $P_\eta(a, b)$ is normalised. Thus the matrix W is a process matrix.

We end by giving a second way in which one can express the connection between process matrices and two-time states via a mapping between operators that act on the Hilbert spaces in the process matrix formalism, and vectors in the Hilbert spaces of the two-time state formalism. Recall that $W^{A_1 A_2 B_1 B_2} \in \mathcal{L}(\mathcal{H}^{A_1} \otimes \mathcal{H}^{A_2} \otimes \mathcal{H}^{B_1} \otimes \mathcal{H}^{B_2})$ is an operator acting on a four-party Hilbert space. We introduce the following mappings

$$\begin{aligned} |i\rangle\langle p| \in \mathcal{L}(\mathcal{H}^{A_1}) &\rightarrow {}_{A_1} \langle p| \otimes |i\rangle^{A_1} \in \mathcal{H}^{A_1}, \\ |j\rangle\langle q| \in \mathcal{L}(\mathcal{H}^{A_2}) &\rightarrow {}_{A_2} \langle j| \otimes |q\rangle^{A_2} \in \mathcal{H}_{A_2}, \\ |k\rangle\langle r| \in \mathcal{L}(\mathcal{H}^{B_1}) &\rightarrow {}_{B_1} \langle r| \otimes |k\rangle^{B_1} \in \mathcal{H}^{B_1}, \\ |l\rangle\langle s| \in \mathcal{L}(\mathcal{H}^{B_2}) &\rightarrow {}_{B_2} \langle l| \otimes |s\rangle^{B_2} \in \mathcal{H}_{B_2}. \end{aligned} \quad (22)$$

Applying this set of transformations to any process matrix W as in (16) leads to the two-time state η_W as in (17). Note that the second and fourth relations, which map the spaces $\mathcal{L}(\mathcal{H}^{A_2})$ and $\mathcal{L}(\mathcal{H}^{B_2})$ from the process matrix formalism to the bra-vector spaces and their conjugates, $\mathcal{H}_{A_2} \equiv \mathcal{H}_{A_2} \otimes \mathcal{H}^{A_2^\dagger}$ and $\mathcal{H}_{B_2} \equiv \mathcal{H}_{B_2} \otimes \mathcal{H}^{B_2^\dagger}$ respectively, are basis-dependent mappings, and must be done in the same basis in which the process matrix is represented (i.e. in this case in the same basis as used in (16)).

6. The set of two-time states corresponding to process matrices

In [7, 9] it is shown that the condition that process matrices give normalised probabilities (that are also positive) for all possible measurements is in fact equivalent to a finite set of conditions, which provide an explicit and compact characterisation of the set of valid process matrices. Here we give a translation of these conditions in the two-time state formalism.

In the bipartite case the set of process matrices is specified by the following five necessary and sufficient conditions:

$$W \geq 0, \quad (23a)$$

$$\text{tr} W = d_{A_2} d_{B_2}, \quad (23b)$$

$${}_{B_1 B_2} W = {}_{A_2 B_1 B_2} W, \quad (23c)$$

$${}_{A_1 A_2} W = {}_{A_1 A_2 B_2} W, \quad (23d)$$

$$W = {}_{A_2} W + {}_{B_2} W - {}_{A_2 B_2} W, \quad (23e)$$

where d_{A_2} (d_{B_2}) is the dimension of Alice's (Bob's) output Hilbert space; the notation in the last three equations is defined using the 'trace-and-replace' operation

$${}_X W := \frac{\mathbb{I}^X}{d_X} \otimes \text{tr}_X W. \quad (24)$$

Returning to two-time states, we first need to introduce two operations

$$T_A^{A'} := \frac{1}{d_{A'}} \mathbb{I}^{A'} \otimes \mathbb{I}_A, \quad (25)$$

$$I_A^{A'} := \mathbb{I}_A^{A'} \otimes \mathbb{I}_{A'}^{A'}, \quad (26)$$

defined in terms of the identity vector (11). The first corresponds to a ‘throw-away-and-replace’ operation, replacing a pre-selected system A by the maximally mixed (pre-selected) state on A' , and is analogous to the ‘trace-and-replace’ operation from (24) for process matrices. The second corresponds to the ‘do-nothing’ operation, taking any state of A to the same state of A' .

Note that in both of these operations we introduce new primed spaces. Depending on the context, these primed spaces should either be thought of as corresponding to a time t'_1 just *after* time t_1 , or to a time t'_2 just *before* time t_2 . For example, $T_{A_1}^{A'_1} \cdot \eta_{A_2}^{A_1} \equiv (\eta')_{A_2}^{A'_1} \in \mathcal{H}^{A'_1} \otimes \mathcal{H}_{A_2}$ is a two-time state between the times t'_1 and t_2 .

With these operations, a translation of the conditions in (23) is⁹

$$\eta_W \geq 0, \quad (28a)$$

$$(\mathbb{I}_{A_1} \otimes \mathbb{I}^{A_2} \otimes \mathbb{I}_{B_1} \otimes \mathbb{I}^{B_2}) \cdot \eta_W = d_{A_2} d_{B_2}, \quad (28b)$$

$$(I_{A_2}^{A_2} \otimes T_{B_1}^{B_2}) \cdot \eta_W = (T_{A_2}^{A_2} \otimes T_{B_1}^{B_2}) \cdot \eta_W, \quad (28c)$$

$$(T_{A_1}^{A_2} \otimes I_{B_2}^{B_2}) \cdot \eta_W = (T_{A_1}^{A_2} \otimes T_{B_2}^{B_2}) \cdot \eta_W, \quad (28d)$$

$$(I_{A_2}^{A_2} \otimes I_{B_2}^{B_2}) \cdot \eta_W = (I_{A_2}^{A_2} \otimes T_{B_2}^{B_2}) \cdot \eta_W + (T_{A_2}^{A_2} \otimes I_{B_2}^{B_2}) \cdot \eta_W - (T_{A_2}^{A_2} \otimes T_{B_2}^{B_2}) \cdot \eta_W. \quad (28e)$$

7. Generalised properties of process matrices

In this section we show that the properties (28) of process matrices, that are expressed in terms of the ‘throw-away-and-replace’ and the ‘do-nothing’ channels ($T_A^{A'}$ and $I_A^{A'}$), can be generalized to arbitrary channels. This demonstrates the usefulness of connecting the process matrix formalism to the formalism of two-time states, and might be useful in future research on process matrices, as it provides further structural information about them.

In appendix B we prove the following result concerning *linear* bipartite two-time states (i.e. about η_W states):

Theorem. *Given any η_W , i.e. any bipartite two time state that satisfies¹⁰*

$$(J_{A_1}^{A_2} \otimes K_{B_1}^{B_2}) \cdot \eta_W = 1, \quad (29)$$

for all $J_{A_1}^{A_2}$ and $K_{B_1}^{B_2}$ corresponding to completely positive trace preserving maps (i.e. of the form (9) and such that $\mathbb{I}_{A_2} \cdot J_{A_1}^{A_2} = \mathbb{I}_{A_1}$ and $\mathbb{I}_{B_2} \cdot K_{B_1}^{B_2} = \mathbb{I}_{B_1}$) then the following properties hold:

$$(C_{A_2}^{A_2} \otimes K_{B_1}^{B_2}) \cdot \eta_W = (D_{A_2}^{A_2} \otimes K_{B_1}^{B_2}) \cdot \eta_W, \quad (30a)$$

$$(J_{A_1}^{A_2} \otimes G_{B_2}^{B_2}) \cdot \eta_W = (J_{A_1}^{A_2} \otimes H_{B_2}^{B_2}) \cdot \eta_W, \quad (30b)$$

for all $C_{A_2}^{A_2}$, $D_{A_2}^{A_2}$, $G_{B_2}^{B_2}$ and $H_{B_2}^{B_2}$ corresponding to completely positive trace preserving maps. From either of the above two conditions, it also follows that¹¹

$$(C_{A_2}^{A_2} \otimes G_{B_2}^{B_2}) \cdot \eta_W = (C_{A_2}^{A_2} \otimes H_{B_2}^{B_2}) \cdot \eta_W + (D_{A_2}^{A_2} \otimes G_{B_2}^{B_2}) \cdot \eta_W - (D_{A_2}^{A_2} \otimes H_{B_2}^{B_2}) \cdot \eta_W. \quad (30c)$$

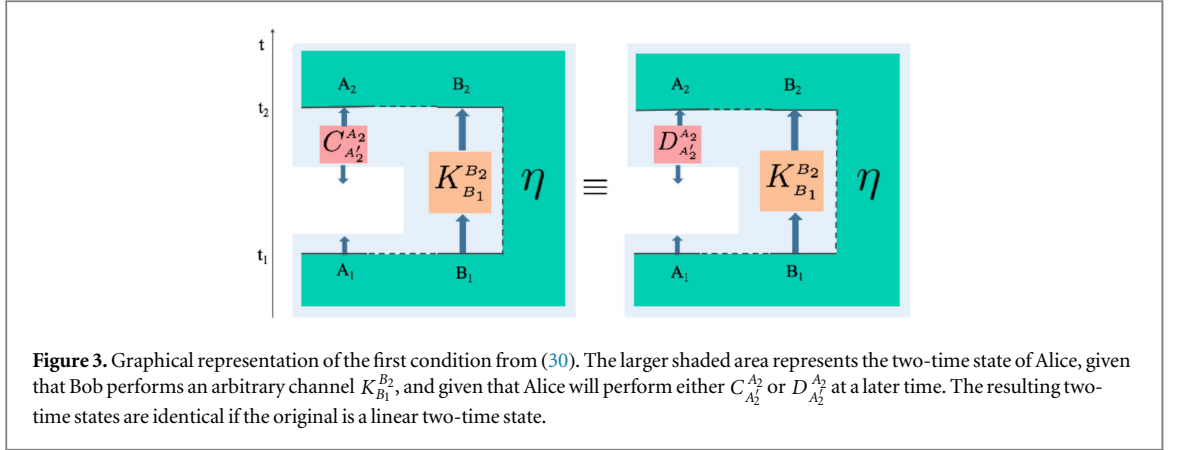
⁹ Strictly speaking, the third condition from (23) translated into two-time notation reads

$$(I_{A_2}^{A_2} \otimes T_{B_1}^{B_1} \otimes T_{B_2}^{B_2}) \cdot \eta_W = (T_{A_2}^{A_2} \otimes T_{B_1}^{B_1} \otimes T_{B_2}^{B_2}) \cdot \eta_W, \quad (27)$$

which is the same as the third condition in (28), with maximally mixed states on Bob’s input and output spaces tensored in on both sides of the equation, which does not change the condition. A similar argument holds for the fourth conditions from (23) and (28).

¹⁰ Since the overall normalisation of a two-time state is unphysical, given a two-time state η that satisfies $(J_{A_1}^{A_2} \otimes K_{B_1}^{B_2}) \cdot \eta = \lambda$ for all $J_{A_1}^{A_2}$ and $K_{B_1}^{B_2}$, we can always re-scale $\eta \rightarrow \frac{1}{\lambda} \eta$ such that this condition holds.

¹¹ Note that the final condition of (30) can be expressed more symmetrically as $[(C_{A_2}^{A_2} - D_{A_2}^{A_2}) \otimes (G_{B_2}^{B_2} - H_{B_2}^{B_2})] \cdot \eta_W = 0$, but we adopt the above form to emphasise its similarity with (23e).



It is interesting that in the case of general channels, the condition (30c) now follows from (30a) and (30b), in contrast to the properties (23), where (23e) does not follow from (23c) and (23d). See figure 3 for a graphical representation of the first condition.

We can translate the above conditions into the process matrix formalism, and present a new set of conditions that are satisfied by all process matrices. To do so, let us introduce arbitrary Choi-Jamiołkowski operators corresponding to completely positive and trace-preserving (CPTP) maps $Q^{A_2 A_2}$, $R^{A_2 A_2}$ and $M^{A_1 A_2}$ for Alice, and similarly $S^{B_2 B_2}$, $Z^{B_2 B_2}$ and $N^{B_1 B_2}$ for Bob. Then, the conditions (30) can be re-expressed as

$$\text{tr}_{A_2 B_1 B_2}[(\mathbb{I}^{A_1} \otimes Q^{A_2 A_2} \otimes N^{B_1 B_2})(W \otimes \mathbb{I}^{A_2'})] = \text{tr}_{A_2 B_1 B_2}[(\mathbb{I}^{A_1} \otimes R^{A_2 A_2} \otimes N^{B_1 B_2})(W \otimes \mathbb{I}^{A_2'})], \quad (31a)$$

$$\text{tr}_{A_1 A_2 B_2}[(M^{A_1 A_2} \otimes \mathbb{I}^{B_1} \otimes S^{B_2 B_2})(W \otimes \mathbb{I}^{B_2'})] = \text{tr}_{A_1 A_2 B_2}[(M^{A_1 A_2} \otimes \mathbb{I}^{B_1} \otimes Z^{B_2 B_2})(W \otimes \mathbb{I}^{B_2'})], \quad (31b)$$

$$\begin{aligned} \text{tr}_{A_2 B_2}[(\mathbb{I}^{A_1 B_1} \otimes Q^{A_2 A_2} \otimes S^{B_2 B_2})(W \otimes \mathbb{I}^{A_2' B_2'})] &= \text{tr}_{A_2 B_2}[(\mathbb{I}^{A_1 B_1} \otimes Q^{A_2 A_2} \otimes Z^{B_2 B_2})(W \otimes \mathbb{I}^{A_2' B_2'})] \\ + \text{tr}_{A_2 B_2}[(\mathbb{I}^{A_1 B_1} \otimes R^{A_2 A_2} \otimes S^{B_2 B_2})(W \otimes \mathbb{I}^{A_2' B_2'})] &- \text{tr}_{A_2 B_2}[(\mathbb{I}^{A_1 B_1} \otimes R^{A_2 A_2} \otimes Z^{B_2 B_2})(W \otimes \mathbb{I}^{A_2' B_2'})]. \end{aligned} \quad (31c)$$

These conditions are not directly implied by the conditions (23), and may prove of independent interest for proving new results regarding process matrices.

8. Indefinite causal structure versus indefinite temporal structure.

In the usual causal scenario an event A can be a cause of an event B only if A takes place before B, and B can be a cause of A only if B takes place first. The most general causal structure is a simple mixture of these cases, by probabilistically sometimes letting A be before B and sometimes B before A. In quantum mechanics one can envisage a more interesting situation in which one can arrange for a *superposition* of these two time orderings, resulting in an ‘indefinite’ causal order. The process matrix formalism is inspired by this idea.

Note that in the above the causal structure is constrained by the temporal structure, and indefinite causal order is induced by indefinite temporal order. On the other hand, given the correspondence between two-time states and process matrices presented here, we have implicitly shown that in quantum mechanics with post-selection (and presumably in every post-selected probabilistic theory) temporal order and causal order are independent notions. In particular one can have indefinite causal order even with definite temporal order.

In the process matrix formalism the actual time at which Alice’s and Bob’s actions take place is left undefined. The only ordering we have is a local time order inside Alice and Bob’s labs, and nothing is said about the relative timing between Alice and Bob. The possibility is left open for Alice to pass her output state to Bob as his input state, and for Bob to pass his output state to Alice as her input state, in some sort of potential superposition, though how exactly this is encoded in the formalism is less clear.

On the other hand, in the two-state formalism the temporal order is perfectly well-defined: pre-selection takes place at time t_1 , post-selection at time t_2 , and the actions of both Alice and Bob take place between these times. Since their initial times are synchronised, in the usual causal scenario, there would be no way for Alice to send her final state to Bob, or vice versa.

However, in cases with fundamental post-selection, this is no longer true, as the post-selection can realize communication backward in time via deterministic teleportation (this is related to the notion of closed timelike curves via post-selection [23, 25]).

In the case of experimental post-selection, the same behaviour is simulated when the post-selection succeeds (although note that this will only be discovered after the experiment is over.)

In both cases, it is the entanglement present in the joint two-time state that enables, for example, Alice's final state to be teleported backward in time to Bobs initial time. Thus we can realise a causal order from Alice to Bob, or vice versa, or even an indefinite causal order (when the joint two-time state corresponds to a process matrix with indefinite causal order), despite having definite temporal order¹².

9. Discussion

In our paper we addressed the question of the relation between the process matrix formalism and quantum theory: does the process matrix formalism imply new phenomena, not present in quantum theory or not? We showed that correlations predicted by the process matrix formalism are identical to those obtainable in a subset of pre- and post-selected quantum states. In this sense, the process matrix formalism is contained within quantum theory. The subtle question is that of the probability with which we can prepare situations described by process matrices. In ordinary quantum theory pre- and post-selected states cannot be prepared with certainty, unless they are trivial (i.e. no post-selection): as their name suggests, sometimes we don't succeed in obtaining them, so, at the end of the experiment we need to select the cases when we were successful and reject the other cases. Hence, if Nature would turn out to be such that all process matrices can be prepared with certainty, then Nature may not be described by quantum theory, but something new. (Here we say 'may not' rather than 'would not', since we have not proven that trivial post-selections are not enough, or that quantum mechanics does not allow simulating W matrices by some other means than by pre- and post-selection.) On the other hand, if one enlarges the scope of quantum theory to quantum theory + fundamental post-selection, in which nature provides the 'post-selected' state with certainty (such as by giving a final state of the Universe, in addition to and independent of the initial state), then the situations described by the process matrix formalism are completely contained within it.

As noted above, the correlations described by the process matrix formalism are the same as those arising from a *subset* of pre- and post-selected states. This raises a new and important question: why only this subset? Why not other pre- and post-selected states? As we have seen, the limitation to this set of states stems from the requirement that the probabilities are linear functions of the states and of the measurements. What would be problematic if the probabilities were nonlinear functions? After all, in the context of pre- and post-selection the probabilities are nonlinear in general. We will address this question in forthcoming work.

We generalised the defining conditions of the process matrices which are given by relations that are obeyed when the systems are subjected to simple 'trace and replace' channels to arbitrary channels. This may be useful in proving further results, or to allow us to gain a deeper understanding of process matrices.

We also discussed the issues of indefinite temporal and causal order and argued that they are actually two independent concepts.

Finally, we note a technical issue which stems from a deeper conceptual issue and shows an advantage of viewing process matrix situations within the two-time state formalism. The usual formalism used to study process matrices, in which they are represented as operators on standard ket vector spaces and which involves a Choi-Jamiołkowski representation, is basis dependent. This leads to the rather complicated relation (14) between the Kraus operators that describe Alice and Bobs measurements and the corresponding operators $M_a^{A_1 A_2}$ and $N_b^{B_1 B_2}$ that are used for calculating probabilities. On the other hand, the two-time formalism, in which a ket vector and a bra vector space are used to represent the forward and backward in time propagating states respectively (and vice versa for the daggered spaces), is basis independent. Indeed, the state $\sum_i |i\rangle^{A_1} |i\rangle^{A_2}$ is basis dependent, while the two-time equivalent $\sum_i |i\rangle^{A_1} \otimes \langle i|^{A_2}$ is basis independent. Correspondingly, the relation (9) between the Kraus operators E_a and the operator J_a used for calculating probabilities is straightforward.

Acknowledgments

YG acknowledges funding from the Austrian Science Fund (FWF) through the START project Y879-N27. SP and PS acknowledge support from the ERC through the AdG NLST. PS acknowledges support from the Royal

¹² Note that in fact any fixed timings for the pre- and post-selections of Alice and Bob, not necessarily the synchronised timings considered above, allows for indefinite causal order with definite temporal order.

Society. RS and NB acknowledge financial support from the Swiss National Science Foundation (starting grants DIAQ and QSIT). AJA acknowledges support from the FQXi via the SVCF.

Appendix A. Proof of the connection between process matrices and two-time states

Here we give the proof of (18), which connects $P_W(a, b)$ and η_W .

$$\begin{aligned}
P_W(a, b) &= \text{tr}[W^{A_1 A_2 B_1 B_2} (M_a^{A_1 A_2} \otimes N_b^{B_1 B_2})] \\
&= \text{tr} \sum_{\mu\nu} W^{A_1 A_2 B_1 B_2} ((\mathbb{I} \otimes \hat{E}_a^\mu) |\Phi^+\rangle \langle \Phi^+| (\mathbb{I} \otimes \hat{E}_a^{\mu\dagger})^\Gamma \otimes [(\mathbb{I} \otimes \hat{F}_b^\nu) |\Phi^+\rangle \langle \Phi^+| (\mathbb{I} \otimes \hat{F}_b^{\nu\dagger})]^\Gamma) \\
&= \text{tr} \sum_{\mu\nu} [W^{A_1 A_2 B_1 B_2}]^\Gamma ((\mathbb{I} \otimes \hat{E}_a^\mu) |\Phi^+\rangle \langle \Phi^+| (\mathbb{I} \otimes \hat{E}_a^{\mu\dagger}) \otimes (\mathbb{I} \otimes \hat{F}_b^\nu) |\Phi^+\rangle \langle \Phi^+| (\mathbb{I} \otimes \hat{F}_b^{\nu\dagger})) \\
&= \text{tr} \sum_{ijkl} \sum_{\mu\nu} \sum_{pqrs} w_{ijkl,pqrs} |pqrs\rangle \langle ijkl| (|t\rangle \langle u| \otimes \hat{E}_a^\mu |t\rangle \langle u| \hat{E}_a^{\mu\dagger} \otimes |v\rangle \langle w| \otimes \hat{F}_b^\nu |v\rangle \langle w| \hat{F}_b^{\nu\dagger}) \\
&= \sum_{ijkl} \sum_{\mu\nu} w_{ijkl,pqrs} \langle j| \hat{E}_a^\mu |i\rangle \langle p| \hat{E}_a^{\mu\dagger} |q\rangle \langle l| \hat{F}_b^\nu |k\rangle \langle r| \hat{F}_b^{\nu\dagger} |s\rangle \\
&= \left(\sum_{ijkl} w_{ijkl,pqrs} A_2 \langle j| \otimes |i\rangle^{A_1} \otimes A_1 \langle p| \otimes |q\rangle^{A_2} \otimes B_2 \langle l| \otimes |k\rangle^{B_1} \otimes B_1 \langle r| \otimes |s\rangle^{B_2} \right) \cdot \left(\sum_{\mu} E_a^\mu \otimes E_a^{\mu\dagger} \otimes \sum_{\nu} F_b^\nu \otimes F_b^{\nu\dagger} \right) \\
&= \eta_W \bullet (J_a \otimes K_b).
\end{aligned} \tag{A1}$$

Appendix B. Proof of the generalized properties of process matrices

In this appendix we give the proof of the generalized properties of process matrices given in the main text (30). We begin by restating the theorem:

Theorem. Given any η_W , i.e. any bipartite two time state that satisfies

$$(J_{A_1}^{A_2} \otimes K_{B_1}^{B_2}) \bullet \eta_W = 1, \tag{B1}$$

for all $J_{A_1}^{A_2}$ and $K_{B_1}^{B_2}$ corresponding to completely positive trace preserving maps (i.e. of the form (9) and such that $\mathbb{I}_{A_2} \bullet J_{A_1}^{A_2} = \mathbb{I}_{A_1}$ and $\mathbb{I}_{B_2} \bullet K_{B_1}^{B_2} = \mathbb{I}_{B_1}$) then the following properties hold:

$$(C_{A_2}^{A_2} \otimes K_{B_1}^{B_2}) \bullet \eta_W = (D_{A_2}^{A_2} \otimes K_{B_1}^{B_2}) \bullet \eta_W, \tag{B2a}$$

$$(J_{A_1}^{A_2} \otimes G_{B_2}^{B_2}) \bullet \eta_W = (J_{A_1}^{A_2} \otimes H_{B_2}^{B_2}) \bullet \eta_W, \tag{B2b}$$

for all $C_{A_2}^{A_2}$, $D_{A_2}^{A_2}$, $G_{B_2}^{B_2}$ and $H_{B_2}^{B_2}$ corresponding to completely positive trace preserving maps. From either of the above two conditions, it also follows that

$$(C_{A_2}^{A_2} \otimes G_{B_2}^{B_2}) \bullet \eta_W = (C_{A_2}^{A_2} \otimes H_{B_2}^{B_2}) \bullet \eta_W + (D_{A_2}^{A_2} \otimes G_{B_2}^{B_2}) \bullet \eta_W - (D_{A_2}^{A_2} \otimes H_{B_2}^{B_2}) \bullet \eta_W. \tag{B2c}$$

Proof. Consider the following CPTP map for Alice

$$\tilde{J}_{A_1}^{A_2} = T_{A_1}^{A_2} + \epsilon (C_{A_2}^{A_2} - D_{A_2}^{A_2}) \bullet X_{A_1}^{A_2'}, \tag{B3}$$

where $T_{A_1}^{A_2}$ is the throw-away-and-replace operation (see (25)), $C_{A_2}^{A_2}$ and $D_{A_2}^{A_2}$ correspond to two arbitrary CPTP maps, $X_{A_1}^{A_2'} \in \mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}'$ is an arbitrary vector¹³, and $\epsilon > 0$ is a positive constant, taken to be sufficiently small such that $\tilde{J}_{A_1}^{A_2}$ is positive. Similarly, for Bob,

$$\tilde{K}_{B_1}^{B_2} = T_{B_1}^{B_2} + \delta (G_{B_2}^{B_2} - H_{B_2}^{B_2}) \bullet Y_{B_1}^{B_2'}, \tag{B4}$$

where $G_{B_2}^{B_2}$ and $H_{B_2}^{B_2}$ are two arbitrary CPTP maps, $Y_{B_1}^{B_2'} \in \mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}'$ is an arbitrary vector, and $\delta > 0$ is a sufficiently small positive constant so that $\tilde{K}_{B_1}^{B_2}$ is positive.

¹³ The only requirement on $X_{A_1}^{A_2'}$ is that $X_{A_1}^{A_2'} \bullet \eta_{A_2}^{A_1} \in \mathbb{R}$ for all $\eta_{A_2}^{A_1} \in \mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$, which can be seen as a ‘hermiticity’ requirement.

Now, for linear two-time states it follows that

$$(\tilde{J}_{A_1}^{A_2} \otimes K_{B_1}^{B_2}) \cdot \eta_{A_2 B_2}^{A_1 B_1} = (T_{A_1}^{A_2} \otimes K_{B_1}^{B_2}) \cdot \eta_{A_2 B_2}^{A_1 B_1}, \quad (\text{B5})$$

for all CPTP maps $K_{B_1}^{B_2}$. Using (B3) to expand out $\tilde{J}_{A_1}^{A_2}$, this is seen to be equivalent to

$$[((C_{A_2}^{A_2'} - D_{A_2}^{A_2'}) \cdot X_{A_1}^{A_2'}) \otimes K_{B_1}^{B_2}] \cdot \eta_{A_2 B_2}^{A_1 B_1} = 0, \quad (\text{B6})$$

where we have used the fact that $\epsilon > 0$ to cancel it. The only way that this can hold for an arbitrary vector $X_{A_1}^{A_2'}$ is if

$$(C_{A_2}^{A_2'} \otimes K_{B_1}^{B_2}) \cdot \eta_{A_2 B_2}^{A_1 B_1} = (D_{A_2}^{A_2'} \otimes K_{B_1}^{B_2}) \cdot \eta_{A_2 B_2}^{A_1 B_1}, \quad (\text{B7})$$

which is the first condition of the theorem.

An identical argument can be made, by starting from $(J_{A_1}^{A_2} \otimes \tilde{K}_{B_1}^{B_2}) \cdot \eta_{A_2 B_2}^{A_1 B_1} = (J_{A_1}^{A_2} \otimes T_{B_1}^{B_2}) \cdot \eta_{A_2 B_2}^{A_1 B_1}$, where $J_{A_1}^{A_2}$ is an arbitrary CPTP map, which leads to the second result of the theorem,

$$(J_{A_1}^{A_2} \otimes G_{B_2}^{B_2'}) \cdot \eta_{A_2 B_2}^{A_1 B_1} = (J_{A_1}^{A_2} \otimes H_{B_2}^{B_2'}) \cdot \eta_{A_2 B_2}^{A_1 B_1}. \quad (\text{B8})$$

Finally, using (B8), by setting $J_{A_1}^{A_2} = \tilde{J}_{A_1}^{A_2}$ from (B3) and applying (B8) to cancel the terms involving $T_{A_1}^{A_2}$, it follows that

$$[((C_{A_2}^{A_2'} - D_{A_2}^{A_2'}) \cdot X_{A_1}^{A_2'}) \otimes G_{B_2}^{B_2'}] \cdot \eta_{A_2 B_2}^{A_1 B_1} = [((C_{A_2}^{A_2'} - D_{A_2}^{A_2'}) \cdot X_{A_1}^{A_2'}) \otimes H_{B_2}^{B_2'}] \cdot \eta_{A_2 B_2}^{A_1 B_1}. \quad (\text{B9})$$

Since this must hold for arbitrary $X_{A_1}^{A_2'}$, it follows that

$$(C_{A_2}^{A_2'} \otimes G_{B_2}^{B_2'}) \cdot \eta_{A_2 B_2}^{A_1 B_1} = (C_{A_2}^{A_2'} \otimes H_{B_2}^{B_2'}) \cdot \eta_{A_2 B_2}^{A_1 B_1} + (D_{A_2}^{A_2'} \otimes G_{B_2}^{B_2'}) \cdot \eta_{A_2 B_2}^{A_1 B_1} - (D_{A_2}^{A_2'} \otimes H_{B_2}^{B_2'}) \cdot \eta_{A_2 B_2}^{A_1 B_1}, \quad (\text{B10})$$

which is the final condition of the theorem and completes the proof.

References

- [1] Aharonov Y, Anandan J, Popescu S and Vaidman L 1990 *Phys. Rev. Lett.* **64** 2965
- [2] Chiribella G, D'Ariano G M, Perinotti P and Valiron B 2013 *Phys. Rev. A* **88** 022318
- [3] Hardy L 2007 *J. Phys. A: Math. Theor.* **40** 3081
- [4] Leifer M S and Spekkens R W 2013 *Phys. Rev. A* **88** 052130
- [5] Baumeler A and Wolf S 2014 *2014 IEEE Int. Symp. on Information Theory* pp 526–30
- [6] Allen J-M A, Barrett J, Horsman D C, Lee C M and Spekkens R W 2016 *Phys. Rev. X* **7** 031021
- [7] Oreshkov O, Costa F and Brukner Č 2012 *Nat. Commun.* **3** 1092
- [8] Branciard C, Araújo M, Feix A, Costa F and Brukner Č 2016 *New J. Phys.* **18** 013008
- [9] Araújo M, Branciard C, Costa F, Feix A, Giarmatzi C and Brukner Č 2015 *New J. Phys.* **17** 102001
- [10] Oreshkov O and Giarmatzi C 2016 *New J. Phys.* **18** 093020
- [11] Abbott A A, Giarmatzi C, Costa F and Branciard C 2016 *Phys. Rev. A* **94** 032131
- [12] Baumeler A and Wolf S 2016 *New J. Phys.* **18** 035014
- [13] Chiribella G 2012 *Phys. Rev. A* **86** 040301
- [14] Araújo M, Costa F and Brukner Č 2014 *Phys. Rev. Lett.* **113** 250402
- [15] Aharonov Y, Bergmann P G and Lebowitz J L 1964 *Phys. Rev.* **134** B1410
- [16] Aharonov Y and Vaidman L 1991 *J. Phys. A: Math. Gen.* **24** 2315
- [17] Silva R, Guryanova Y, Brunner N, Linden N, Short A J and Popescu S 2014 *Phys. Rev. A* **89** 012121
- [18] Aharonov Y 1990 personal communication
Aharonov Y and Rohrlich D 2005 *Quantum Paradoxes: Quantum Theory for the Perplexed* (New York: Wiley)
Aharonov Y and Gruss E Y 2005 arXiv:0507269 [quant-ph]
- [19] Hawking S W, Hertog T and Reall H S 2000 *Phys. Rev. D* **62** 043501
- [20] Gell-Mann M and Hartle J 1990 *Complexity, Entropy and the Physics of Information (SFI Studies in the Sciences of Complexity vol VIII)* ed A W Zurek (Reading, MA: Wesley)
- [21] Horowitz G T and Maldacena J 2004 *J. High Energy Phys.* **2004** 008
- [22] Aharonov Y, Popescu S, Tollaksen J and Vaidman L 2009 *Phys. Rev. A* **79** 052110
- [23] Lloyd S, Maccone L, Garcia-Patron R, Giovannetti V and Shikano Y 2011 *Phys. Rev. D* **84** 025007
- [24] Procopio L M et al 2015 *Nat. Commun.* **6** 7913
- [25] Bennett C H 2005 *Proc. QuPoN (Wien)*
- [26] Oreshkov O and Cerf N J 2016 Operational quantum theory without predefined time *New J. Phys.* **18** 073037