On the definition and characterisation of multipartite causal (non)separability

Julian Wechs, Alastair A. Abbott, Cyril Branciard

PAPER • OPEN ACCESS

On the definition and characterisation of multipartite causal (non)separability

To cite this article: Julian Wechs *et al* 2019 *New J. Phys.* **21** 013027

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
1 August 2018

REVISED
21 November 2018

ACCEPTED FOR PUBLICATION
23 November 2018

PUBLISHED
25 January 2019

PAPER

On the definition and characterisation of multipartite causal (non) separability

Julian Wechs, Alastair A Abbott and Cyril Branciard

Univ. Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, F-38000 Grenoble, France

E-mail: julian.wechs@neel.cnrs.fr

Keywords: quantum foundations, quantum information, causality, indefinite causal order, process matrix framework

Abstract

The concept of causal nonseparability has been recently introduced, in opposition to that of causal separability, to qualify physical processes that locally abide by the laws of quantum theory, but cannot be embedded in a well-defined global causal structure. While the definition is unambiguous in the bipartite case, its generalisation to the multipartite case is not so straightforward. Two seemingly different generalisations have been proposed, one for a restricted tripartite scenario and one for the general multipartite case. Here we compare the two, showing that they are in fact inequivalent. We propose our own definition of causal (non)separability for the general case, which—although *a priori* subtly different—turns out to be equivalent to the concept of 'extensible causal (non)separability' introduced before, and which we argue is a more natural definition for general multipartite scenarios. We then derive necessary, as well as sufficient conditions to characterise causally (non)separable processes in practice. These allow one to devise practical tests, by generalising the tool of witnesses of causal nonseparability.

1. Introduction

The notion of a causal order between events is an essential ingredient in our understanding of the world. Our conventional view of causality is that events are ordered according to some global time parameter, with past events influencing future events, but not vice versa. One may however wonder whether this concept is really fundamental, or whether scenarios without such an underlying background causal structure are conceivable. The situation is particularly interesting in quantum theory, where the properties of physical systems are not always well-defined, and where the question arises of whether the causal structure itself can be subject to quantum effects in a similar way. These questions are of great importance for the foundations of physics [1–3], but they are also motivated by a more practical point of view, as new resources for quantum information processing become available when the assumption of a definite causal structure is relaxed [4]. Recent works have demonstrated that, for instance, indefinite causal orders can enable advantages in regard to query complexity [5–8], communication complexity [9, 10] and other information processing tasks [11–13].

A particular model describing causal relations between quantum events is the so-called process matrix formalism [2]. In this framework, quantum events are assumed to take place locally, but the causal order between them is not specified *a priori*. The physical resource relating the local events is described by a *process matrix*, which, broadly speaking, is a generalisation of a multipartite density matrix allowing also for the description of signalling scenarios, such as quantum channels. As it turns out, some scenarios arising within this formalism are indeed incompatible with any definite causal order. The process matrices corresponding to these scenarios are called *causally nonseparable*, while the process matrices describing scenarios compatible with a well-defined causal structure are called *causally separable*.

The process matrix formalism was initially introduced for two local events. In that bipartite case, the notion of causal (non)separability is clearly defined and well understood. In particular, the causal (non)separability of any bipartite process matrix can be determined using *witnesses of causal nonseparability* [14, 15], similar conceptually to entanglement witnesses. In order to comprehensively understand causal indefiniteness from a fundamental perspective, and to explore more deeply the question of how they can be harnessed as a quantum

information processing resource, it is essential to clarify how the absence of a causal order can be described, characterised and certified also in multipartite scenarios. While the formalism of process matrices generalises rather easily to more parties [14, 16, 17], the notion of causal (non)separability becomes less clear. In fact, several different definitions have recently been proposed to generalise the bipartite case [14, 17] which, as it turns out, are not equivalent.

In this work, we clarify the definition of causal (non)separability in multipartite scenarios. After recalling the framework and definitions in the bipartite case, we compare the generalisations of causal (non)separability that have been proposed so far, before proposing and motivating our own definition for the multipartite case (Definition 5). We then provide a characterisation of multipartite causally (non)separable processes via necessary as well as sufficient conditions (Propositions 3, 4 and 5), allowing us to generalise the tool of witnesses of causal nonseparability.

2. Process matrix formalism: the basics

The process matrix formalism is perhaps most easily understood in the bipartite case [2]. To begin with, let us briefly recall its framework for this case, before turning to the generalisation to multipartite scenarios.

2.1. Bipartite process matrices

The formalism of process matrices was introduced in [2] to study correlations between events that locally obey the laws of quantum theory, but which are not *a priori* embedded into any global causal order. In the bipartite scenario, two parties, who we shall call Alice (A) and Bob (B), are each associated with closed laboratories. The parties perform an experiment during which their interactions with the 'outside world' (and hence with each other) are restricted to opening their laboratories only once to let an incoming physical system enter, and once to send out an outgoing system. Alice and Bob may choose local operations to perform within their laboratories, possibly depending on some external (classical) input x or y for A and B, and producing (classical) measurement outcomes a and b, respectively. The correlations established between the parties after repeating the experiment many times are described by the conditional probability distribution $P(a, b|x, y)$.

While no assumption is made about the global causal order between the parties, we assume that the local operations performed inside the laboratories are described by standard quantum theory. We can therefore assign some 'incoming' and 'outgoing' Hilbert spaces to the parties, which we denote $\mathcal{H}^{A_I}$, $\mathcal{H}^{A_O}$ (for Alice) and $\mathcal{H}^{B_I}$, $\mathcal{H}^{B_O}$ (for Bob), of dimensions d_{A_I}, d_{A_O}, d_{B_I} and d_{B_O}, respectively. The spaces of Hermitian linear operators over these Hilbert spaces will simply be denoted by A_I, A_O, B_I and B_O. For convenience we also define $A_{IO} \coloneqq A_I \otimes A_O$, $B_{IO} \coloneqq B_I \otimes B_O$, $d_{A_{IO}} \coloneqq d_{A_I} d_{A_O}$ and $d_{B_{IO}} \coloneqq d_{B_I} d_{B_O}$. In this paper, we will only consider finite-dimensional Hilbert spaces; for a generalisation of the framework to infinite-dimensional systems, see [18].

According to quantum theory, Alice and Bob's local operations can most generally be described as quantum instruments [19]—that is, sets of completely positive (CP) maps that sum up to CP trace-preserving (CPTP) maps. The Choi–Jamiołkowski (CJ) isomorphism [20, 21] allows us to represent these CP maps as positive semidefinite matrices $M_{a|x}^{A_{IO}}$, $M_{b|y}^{B_{IO}}$, and the CPTP maps as positive semidefinite matrices $M_x^{A_{IO}} \coloneqq \sum_a M_{a|x}^{A_{IO}}$, $M_y^{B_{IO}} \coloneqq \sum_b M_{b|y}^{B_{IO}}$ that satisfy $\mathrm{Tr}_{A_O} M_x^{A_{IO}} = \mathbb{1}^{A_I}$ and $\mathrm{Tr}_{B_O} M_y^{B_{IO}} = \mathbb{1}^{B_I}$. Here, Tr_X denotes the partial trace over the system X, and $\mathbb{1}^X$ denotes the identity operator in the space X (in general, superscripts on operators, which may be omitted when clear enough, denote the system(s) they apply to).

As shown in [2], requiring compatibility with quantum mechanics locally and assuming the non-contextuality of the probabilities imply that the probabilities $P(a, b|x, y)$ must be bilinear in the CP maps associated with the operations of A and B—or, equivalently, bilinear in their CJ representations. (Throughout this paper we will often refer to CP maps by their equivalent CJ representation and vice versa.) It follows that the overall process can be described by a Hermitian operator, a 'process matrix' $W \in A_{IO} \otimes B_{IO}$ [2], such that the correlations are obtained via the *generalised Born rule*

$$P(a, b|x, y) = \mathrm{Tr}[M_{a|x}^{A_{IO}} \otimes M_{b|y}^{B_{IO}} \cdot W] \tag{1}$$

(where Tr is now the full trace).

The framework also permits the parties to share, in addition to the process matrix, some (possibly entangled) ancillary quantum state that can be accessed via their local operations. The parties may thus have access also to some extra incoming Hilbert spaces $\mathcal{H}^{A_{I'}}$ and $\mathcal{H}^{B_{I'}}$ of arbitrary (finite) dimension, and be able to perform CP maps $M_{a|x}^{A_{II'O}} \in A_{II'O} \coloneqq A_I \otimes A_{I'} \otimes A_O$ and $M_{b|y}^{B_{II'O}} \in B_{II'O} \coloneqq B_I \otimes B_{I'} \otimes B_O$, respectively (where as before, $A_{I'}$ and $B_{I'}$ are the spaces of Hermitian linear operators over $\mathcal{H}^{A_{I'}}$ and $\mathcal{H}^{B_{I'}}$). This implies that any process matrix $W \in A_{IO} \otimes B_{IO}$ can be extended to a process matrix $W \otimes \rho \in A_{II'O} \otimes B_{II'O}$, for any extra incoming spaces $A_{I'}$, $B_{I'}$ and any $\rho \in A_{I'} \otimes B_{I'}$ [2].

Requiring equation (1) to yield valid (i.e. nonnegative and normalised) probabilities, even when the parties share arbitrary ancillary states, is equivalent to W satisfying the following constraints:

$$W \geqslant 0, \quad W \in \mathcal{L}^{\{A,B\}}, \quad \text{and} \quad \operatorname{Tr} W = d_{A_O} d_{B_O} \tag{2}$$

for some particular linear subspace $\mathcal{L}^{\{A,B\}}$ of $A_{IO} \otimes B_{IO}$; see section 4.1 and appendix A.1 for an explicit characterisation [2, 14]. In the following we will refer to a matrix satisfying the first two constraints above (i.e. without necessarily imposing the normalisation constraint $\operatorname{Tr} W = d_{A_O} d_{B_O}$) as a *valid* process matrix, and whenever we talk about a process matrix W we always implicitly assume it is valid. Hermitian matrices that are not valid process matrices will simply be referred to as 'matrices'.

2.2. Bipartite causal (non)separability

One may now consider the question, whether the situation described by a process matrix can be embedded in a well-defined causal structure, with a fixed causal order between the events happening in each party's laboratory, or not.

A process matrix is said to be 'compatible with (the causal order) $A \prec B$' (sometimes abbreviated to just '$A \prec B$', e.g. in superscripts) if all the correlations it generates are compatible with a causal order where A acts before B, which is to be understood operationally: such a process matrix $W^{A\prec B}$ does not allow for any signalling from B to A. More precisely, whatever the CP and CPTP maps $M^{A_{IO}}_{a|x}$, $M^{B_{IO}}_{y'|y}$ of A and B, the resulting correlations respect the no-signalling condition $P(a|x, y) = P(a|x, y')$, or $\operatorname{Tr}[M^{A_{IO}}_{a|x} \otimes M^{B_{IO}}_{y} \cdot W^{A\prec B}] = \operatorname{Tr}[M^{A_{IO}}_{a|x} \otimes M^{B_{IO}}_{y'} \cdot W^{A\prec B}]$ according to equation (1). This constrains $W^{A\prec B}$ to be in a linear subspace $\mathcal{L}^{A\prec B} \subset \mathcal{L}^{\{A,B\}}$ of $A_{IO} \otimes B_{IO}$; see section 4.1 and appendix A.2 for an explicit characterisation of $\mathcal{L}^{A\prec B}$.

Likewise, process matrices that do not allow signalling from A to B are said to be compatible with the causal order $B \prec A$, and will typically be denoted $W^{B\prec A} \in \mathcal{L}^{B\prec A}$. One can also conceive of situations where the causal order is not fixed to be the same for all experimental runs, but where there is instead a probabilistic mixture of the two possibilities. Such a scenario is described by a convex combination of process matrices compatible with $A \prec B$ and $B \prec A$, respectively. Process matrices of this form remain compatible with an underlying causal framework and are the subject of the following definition, first introduced by Oreshkov, Costa and Brukner [2]:

Definition 1 (Bipartite causal (non)separability [2]). A bipartite process matrix W is said to be *causally separable* if and only if it can be written as a convex combination

$$W = q\, W^{A\prec B} + (1 - q)\, W^{B\prec A}, \tag{3}$$

with $q \in [0, 1]$ and where $W^{A\prec B}$ and $W^{B\prec A}$ are two process matrices compatible with the causal orders $A \prec B$ and $B \prec A$, respectively.

A process matrix that cannot be decomposed as above is said to be *causally nonseparable*.

Causally separable process matrices thus describe the most general bipartite situations where one can identify a definite causal order between the parties, be it fixed for all experimental runs or subject to classical randomness. In contrast, if a process matrix is causally nonseparable, it is incompatible with any causal order between A and B. In the bipartite case, causal (non)separability can be easily and efficiently verified; in particular, any causally nonseparable process can be detected using a *witness of causal nonseparability* [14, 15] (see section 4.4).

2.3. Towards generalising to more parties

The process matrix framework itself generalises rather easily to the multipartite case.

Let us first introduce some generalised notations. We shall consider N parties denoted by A_k for $k \in \{1, \ldots, N\} \coloneqq \mathcal{N}$, with corresponding inputs and outputs denoted by x_k, and a_k, respectively. We define the input and output vectors $\vec{x} \coloneqq (x_1, \ldots, x_N)$ and $\vec{a} \coloneqq (a_1, \ldots, a_N)$. The 'incoming' and 'outgoing' Hilbert spaces for each party are denoted by $\mathcal{H}^{A^k_I}$, $\mathcal{H}^{A^k_O}$ (of dimensions $d_{A^k_I}$, $d_{A^k_O}$, respectively), while the spaces of Hermitian linear operators over these Hilbert spaces are denoted by A^k_I, A^k_O. We also define $A^k_{IO} \coloneqq A^k_I \otimes A^k_O$, and $d_{A^k_{IO}} \coloneqq d_{A^k_I} d_{A^k_O}$.

For a subset $\mathcal{K} \subseteq \mathcal{N}$ of parties, we will denote by $\vec{x}_\mathcal{K}$ and $\vec{a}_\mathcal{K}$ the vectors of inputs and outputs restricted to the parties in $\mathcal{K}$, and use shorthand notations like $A^\mathcal{K}_{IO} \coloneqq \bigotimes_{k\in\mathcal{K}} A^k_{IO}$ ($=\mathbb{R}$ if $\mathcal{K} = \varnothing$), $\mathbb{1}^\mathcal{K} \coloneqq \bigotimes_{k\in\mathcal{K}} \mathbb{1}^{A^k_{IO}} = \mathbb{1}^{A^\mathcal{K}_{IO}}$, and $\operatorname{Tr}_\mathcal{K}$ for the trace over all (incoming and outgoing) systems of the parties in $\mathcal{K}$—i.e. $\operatorname{Tr}_{A^\mathcal{K}_{IO}}$ or $\operatorname{Tr}_{A^\mathcal{K}_{II'O}}$, as appropriate (see below), and with $\operatorname{Tr}_\varnothing$ the identity operation and $\operatorname{Tr}_\mathcal{N}$ the full trace. For notational simplicity, we shall identify the parties' names with their labels, and singletons of parties (e.g. $\{A_k\}$) with the parties themselves (e.g. A_k) or the corresponding label, so that $\mathcal{N} = \{1, \ldots, N\} \equiv \{A_1, \ldots, A_N\}$, $\mathcal{N}\setminus\{A_k\} \equiv \mathcal{N}\setminus k$, $\operatorname{Tr}_{\{A_k\}} \equiv \operatorname{Tr}_k$, etc.

The CP maps corresponding to the parties' operations are then denoted by $M^{A_{IO}^k}_{a_k|x_k}$, the corresponding CPTP maps $M^{A_{IO}^k}_{x_k} := \sum_{a_k} M^{A_{IO}^k}_{a_k|x_k}$, and the overall process is represented by a process matrix $W \in A_{IO}^{\mathcal{N}}$. The resulting correlations are then obtained through a generalised Born rule as before:

$$P(\vec{a}|\vec{x}) = \mathrm{Tr}[M^{A_{IO}^1}_{a_1|x_1} \otimes \cdots \otimes M^{A_{IO}^N}_{a_N|x_N} \cdot W]. \tag{4}$$

As in the bipartite case, the parties may also share some ancillary state ρ in some extra incoming spaces $A_{I'}^1 \otimes \cdots \otimes A_{I'}^N = A_{I'}^{\mathcal{N}}$, and extend their local operations to act on these spaces as well. Requiring again the nonnegativity and normalisation of all obtainable probabilities, including for arbitrary extensions $W \otimes \rho$ of W, imposes validity constraints on W. In the general multipartite case, they read

$$W \geqslant 0, \quad W \in \mathcal{L}^{\mathcal{N}}, \quad \text{and} \quad \mathrm{Tr}\, W = \prod_{k \in \mathcal{N}} d_{A_O^k} \tag{5}$$

for some particular linear subspace $\mathcal{L}^{\mathcal{N}}$ of $A_{IO}^{\mathcal{N}}$; see section 4.1 and appendix A.1 [2, 14]. As for the bipartite case, in this paper a matrix will be called a (valid) process matrix whenever it satisfies the first two constraints above, without necessarily requiring that it is correctly normalised.

The no-signalling constraints can readily be generalised to the N-partite case, allowing the notion of compatibility with a fixed causal order to be extended accordingly. For instance, a process matrix is said to be compatible with the fixed causal order $A_1 \prec A_2 \prec \cdots \prec A_N$ if no party or group of parties can signal to other parties in their causal 'past' (as defined by the specified causal order)—which translates into the constraint that $P(a_1, \ldots, a_k|\vec{x}) = P(a_1, \ldots, a_k|x_1, \ldots, x_k)$ for all $k = 1, \ldots, N-1$. As before, this constrains such a process matrix $W^{A_1 \prec \cdots \prec A_N}$ to be in a linear subspace $\mathcal{L}^{A_1 \prec \cdots \prec A_N} \subset \mathcal{L}^{\mathcal{N}}$ of $A_{IO}^{\mathcal{N}}$; see section 4.1 for an explicit characterisation of $\mathcal{L}^{A_1 \prec \cdots \prec A_N}$ (and appendices A.2–A.4 for further discussions and characterisations of process matrices compatible with other fixed causal orders).

What is not so straightforward, however, is to generalise the concept of causal (non)separability, which turns out to be much more subtle for more than two parties. In particular, additional complexity arises in the multipartite case because the causal order can be *dynamical* as well as probabilistic—that is, the causal order of parties in the future can depend on operations of parties in the past [16, 17, 22]. Simply considering a convex combination of process matrices compatible with different fixed causal orders does not include scenarios with such dynamical causal orders, and is therefore too restrictive to capture all scenarios that should be considered compatible with a well-defined causal order. Perhaps more strikingly, as we shall see the possibility to extend process matrices with ancillary quantum states has nontrivial implications for the definition of causal (non) separability for more than two parties [17]. The main objectives of this paper are precisely to discuss how the concept of causal (non)separability should properly be generalised to the multipartite case, and to characterise causally separable and causally nonseparable process matrices.

3. Defining multipartite causal (non)separability

3.1. Araújo *et al*'s definition

The multipartite case was first considered in a restricted tripartite situation in which one party has no (or, equivalently, a trivial) outgoing system. This particular scenario was studied because of its relevance for a practical protocol where the causal order between two parties A and B, which perform some unitary operations U_A and U_B on a target system initialised in a state $|\psi\rangle^t$, is controlled by another (two-dimensional) quantum system. If this control qubit is initialised in the state $|0\rangle^c$, the operation U_A is applied before U_B, while for a control qubit in the state $|1\rangle^c$, U_B is applied before U_A. If the control qubit is initialised in a superposition state $|+\rangle^c = \frac{1}{\sqrt{2}}(|0\rangle^c + |1\rangle^c)$, the overall transformation on the joint state of the target and control systems is thus

$$|\psi\rangle^t \otimes |+\rangle^c \to \frac{1}{\sqrt{2}}(U_B U_A\, |\psi\rangle^t \otimes |0\rangle^c + U_A U_B\, |\psi\rangle^t \otimes |1\rangle^c), \tag{6}$$

i.e. the unitaries are applied in a 'superposition of orders'. The output state is then sent to a third party C (Charlie) who can measure the control qubit, and possibly also the target system. The protocol just described can straightforwardly be generalised to the case where A and B's operations are general quantum instruments instead of unitaries. This so-called *quantum switch* can be understood as a quantum supermap [23], or higher order transformation, that maps A and B's local operations to the overall global transformation. It cannot be realised by inserting the local operations into a circuit with a well-defined causal order, and therefore constitutes a new resource for quantum computation that goes beyond causally ordered quantum circuits [4]. It has attracted particular interest as a consequence of being readily implementable, and indeed several implementations have been experimentally realised [24–28]. Consequent work has sought to clarify whether such implementations can

really be seen as genuine realisations of indefinite causal orders, and [29] gives arguments clarifying why they can be.

The quantum switch can naturally be described in the process matrix formalism [14, 17] where it indeed corresponds to a tripartite process matrix for parties A, B and C, where Charlie has no outgoing system and therefore cannot signal to the other parties. The situation is thus relatively similar to the bipartite case, since the only relevant causal orders are those where Charlie acts last, i.e. $A \prec B \prec C$ and $B \prec A \prec C$. This observation led Araújo *et al* to propose the following definition (as an initial, '1-step' generalisation of definition 1) for this particular scenario:

Definition 2 (Araújo *et al*'s causal separability [14]). In a tripartite scenario where party C has no outgoing system, a process matrix W is said to be *causally separable* if and only if it can be written as a convex combination

$$W = q\, W^{A\prec B\prec C} + (1-q)\, W^{B\prec A\prec C}, \tag{7}$$

with $q \in [0,\, 1]$ and where $W^{A\prec B\prec C}$ and $W^{B\prec A\prec C}$ are two process matrices compatible with the causal orders $A \prec B \prec C$ and $B \prec A \prec C$, respectively.

It was shown that the process matrix describing the quantum switch is causally nonseparable as per definition 2 [14], and this definition has subsequently been used e.g. in [15, 18, 27].

3.2. Oreshkov and Giarmatzi's definitions

While Araújo *et al*'s definition recalled above applied only to a particular tripartite situation, Oreshkov and Giarmatzi (OG) considered in [17] the general multipartite case—taking into account, in particular, the possibility of dynamical causal orders. They defined in fact two possible generalisations of bipartite causal (non) separability, namely what they called the notions of 'causal (non)separability' and 'extensible causal (non) separability'.

The definition they proposed for causal separability is recursive, in analogy with the definition of multipartite 'causal correlations' [17, 22]—correlations that are compatible with a definite causal order. In [17, 22], these were characterised as those for which it is possible to identify, up to some probability, a party that acts first, and such that, for any behaviour of this first party, the conditional correlations shared by the remaining parties are again causal. Oreshkov and Giarmatzi invoked an analogous 'unravelling argument' for causally separable processes.

More specifically, their definition is based on the concept of a 'conditional (process) matrix', defined for a given matrix W and a given CP map $M_k := M^{A_k^k}_{a_k|x_k}$ applied by a party A_k as

$$W_{|M_k} := \mathrm{Tr}_k[M_k \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W]. \tag{8}$$

In general, even if W is a valid process matrix, $W_{|M_k}$ thus defined may not be a valid process matrix (in which case we shall just talk about a 'conditional matrix'). In fact, as we will see in section 4.1, a process matrix W is compatible with party A_k acting first (i.e. it does not allow signalling from the other parties to A_k) if and only if for any CP map M_k the conditional matrix $W_{|M_k}$, as defined in equation (8), is (up to normalisation[1]) a valid $(N-1)$-partite process matrix for the parties in $\mathcal{N}\setminus k$. In that case, the conditional process matrix $W_{|M_k}$ then represents the process shared by these $N-1$ parties, conditioned on party A_k performing the CP map $M_k = M^{A_k^k}_{a_k|x_k}$ (i.e. conditioned on both receiving the input x_k and obtaining the outcome a_k).

Oreshkov and Giarmatzi then proposed the following (recursive) definition:[2]

Definition 3 (Oreshkov and Giarmatzi's causal separability [17]). For $N = 1$, any process matrix is causally separable. For $N \geqslant 2$, an N-partite process matrix W is said to be *causally separable* if and only if it can be decomposed as

$$W = \sum_{k\in\mathcal{N}} q_k\, W_{(k)}, \tag{9}$$

with $q_k \geqslant 0$, $\sum_k q_k = 1$, and where for each k, $W_{(k)}$ is a process matrix compatible with party A_k acting first, and is such that for any possible CP map $M_k \in A_{IO}^k$ applied by party A_k, the conditional $(N-1)$-partite process matrix $(W_{(k)})_{|M_k} := \mathrm{Tr}_k[M_k \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W_{(k)}]$ is itself causally separable.

[1] For a properly normalised process matrix W compatible with A_k first (i.e. which always gives $P(a_k|\vec{x}) = P(a_k|x_k)$) and a trace-non-increasing CP map $M_k = M_{a_k|x_k}$, one has $\mathrm{Tr}\, W_{|M_k} = P(a_k|x_k) \prod_{j\in\mathcal{N}\setminus k} d_{A_O^j}$, so that $W_{|M_k}$ must be divided by the factor $P(a_k|x_k)$ to also be properly normalised according to equation (5).

[2] More precisely, what we present here as their definition is actually presented in [17] (in a slightly different, but equivalent way) as a characterisation following from a more fundamental recursive definition of causally separable processes (not necessarily quantum mechanical).

As outlined in the previous section, the process matrix framework allows for process matrices to be extended by providing additional ancillary states the the parties. Taking this into account, OG introduced a second definition of causal separability for process matrices that are causally separable even under arbitrary such extensions:

Definition 4 (**Oreshkov and Giarmatzi's extensible causal separability** [17]). An N-partite process matrix W is said to be *extensibly causally separable* if and only if it is causally separable (as per definition 3 above), and it remains so under any extension with incoming systems in an arbitrary joint quantum state—i.e. if and only if for any extension $A_{I'}^{\mathcal{N}}$ of the parties' incoming systems and any ancillary quantum state $\rho \in A_{I'}^{\mathcal{N}}$, $W \otimes \rho$ is causally separable.

It is easy to see that OG's causal separability (CS) and extensible causal separability (ECS) are equivalent in the bipartite case, and, indeed, equivalent to definition 1 given in section 2.2: the process matrix $W \otimes \rho$ obtained by attaching an ancillary state ρ to a causally separable process matrix W of the form of equation (3) remains of the same form, with $W^{A\prec B} \otimes \rho$ ($W^{B\prec A} \otimes \rho$) compatible with A acting before B (B before A), and for both terms $W^{A\prec B} \otimes \rho$ and $W^{B\prec A} \otimes \rho$, whatever operation the first party applies, the resulting conditional process matrix for the other party is single-partite, hence trivially causally separable.

However, OG's CS and ECS are not equivalent in the general multipartite case and thus indeed represent two different possible multipartite generalisations of the same bipartite concept. Of course ECS implies CS, but the converse is not true in general—the result of a phenomenon called 'activation of causal nonseparability' in [17]. An explicit example of a CS process that is not ECS was indeed given in [17], in a tripartite scenario where one party has no incoming system; we will see another example in the following subsection.

3.3. Comparison

We thus now have three potential generalisations of the concept of causal separability to the particular tripartite situation where one party has no outgoing system—namely, the two different definitions of causal separability (Definitions 2 and 3), and that of extensible causal separability (Definition 4). How do they relate to one another? Are the two definitions of causal separability indeed equivalent? These questions are answered by the following result:

Proposition 1. *In a tripartite scenario where party* C *has no outgoing system, Araújo et al's definition of causal separability (Definition 2) is equivalent to Oreshkov and Giarmatzi's definition of extensible causal separability (Definition 4), but nonequivalent to their definition of causal separability (Definition 3).*

The equivalence between definitions 2 and 4 for this particular tripartite scenario is proved explicitly in appendix B.1.1, which we refer to for more details; we simply summarise the argument here as follows. Clearly, any process matrix W of the form of equation (7) is ECS, as any $W \otimes \rho$ is also of that form (and of the form also of equation (9)), and for any $W^{A\prec B\prec C}$ and any M_A, the conditional process $(W^{A\prec B\prec C})_{|M_A}$ is compatible with the order $B \prec C$ (hence it is causally separable; similarly for any $W^{B\prec A\prec C}$ and any M_B). The proof that an ECS process matrix W necessarily has the form of equation (7) is based on a 'teleportation technique' (see lemma B1 in appendix B), already used in [17], that consists in introducing an ancillary system in a maximally entangled state ρ shared by two parties, e.g. A and C. By definition, the global process $W \otimes \rho_{A_{I'}C_{I'}}$ has a decomposition of the form (9). It is then easy to see that the terms W_A and W_B compatible with parties A or B acting first are in fact compatible, since C has no outgoing system, with the causal orders $A \prec B \prec C$ and $B \prec A \prec C$, respectively, and thus contribute to the terms $W^{A\prec B\prec C}$ and $W^{B\prec A\prec C}$ in equation(7). For the term W_C compatible with C acting first, letting C project his systems $C_{II'} \coloneqq C_I \otimes C_{I'}$ onto the maximally entangled state effectively 'teleports' his system to A. By definition, the conditional bipartite process then shared by A and B must be causally separable, and must therefore have a decomposition of the form (3), which also leads to a decomposition of the form (7) for W_C.

In order to prove the nonequivalence between Araújo *et al* and OG's definitions of causal separability, we will now show that OG's CS and ECS are nonequivalent—i.e. that there can be 'activation of causal nonseparability' (according to OG's terminology)—in the scenario where party C has no outgoing system. Note that this scenario differs from that in which OG already gave an example of activation of causal nonseparability: they indeed considered a tripartite case where C has no *incoming* system, rather than no *outgoing* system.

Consider for that the following process matrix:

$$W^{\text{act.}} := \frac{1}{8}\Big[\mathbb{1}(\mathbb{1}\mathbb{1} - \hat{z}\hat{z})\mathbb{1}\mathbb{1} + \frac{\sqrt{3}}{4}\mathbb{1}(\hat{x}\hat{x} + \hat{y}\hat{y})(\hat{z}\mathbb{1} + \mathbb{1}\hat{z}) + \frac{1}{2}\hat{z}(\hat{z}\mathbb{1} - \mathbb{1}\hat{z})\mathbb{1}\mathbb{1} + \frac{1}{4}\hat{x}(\hat{x}\hat{y} - \hat{y}\hat{x})(\hat{z}\mathbb{1} - \mathbb{1}\hat{z})\Big], \quad (10)$$

where the subsystems are written, for convenience, in the order $C_I A_I B_I A_O B_O$ (i.e. $W^{\text{act.}} \in C_I \otimes A_I \otimes B_I \otimes A_O \otimes B_O$). Here, as in the other examples presented in this paper, $\hat{x}$, $\hat{y}$, $\hat{z}$ denote the Pauli matrices, $\mathbb{1}$ denotes the 2×2 identity matrix and tensor products between all matrices are implicit.

We note first that $W^{\text{act.}}$ is compatible with Charlie acting first—i.e. with the order $C \prec \{A, B\}$[3]. (Indeed, it satisfies equation (18) given later, for $A_k = C$.) Any CP map applied by Charlie—i.e. since C has no outgoing system, any element of a positive-operator valued measure (POVM) in his qubit incoming space C_I—can be written as $M_{\vec{c}} = \mathbb{1} + \vec{c} \cdot \vec{\sigma}$, where $\vec{\sigma} := (\hat{x}, \hat{y}, \hat{z})$ and $\vec{c} := (c_x, c_y, c_z)$ is a three-dimensional real vector with $|\vec{c}| \leqslant 1$, so that $M_{\vec{c}} \geqslant 0$ (and where we ignore the trace-nonincreasing constraint, and indeed the overall normalisation of $M_{\vec{c}}$, since it is irrelevant for our argument). The resulting conditional matrix for parties A and B (as defined in equation (8)) is then

$$(W^{\text{act.}})_{|M_{\vec{c}}} := \text{Tr}_{C_I}[M_{\vec{c}}^{C_I} \otimes \mathbb{1}^{A_I B_I A_O B_O} \cdot W^{\text{act.}}] = \frac{1}{2} W_{|M_{\vec{c}}}^{A \prec B} + \frac{1}{2} W_{|M_{\vec{c}}}^{B \prec A} \quad (11)$$

with (written in the order $A_I B_I A_O B_O$)

$$W_{|M_{\vec{c}}}^{A \prec B} := \frac{1}{4}\Big[(\mathbb{1}\mathbb{1} - \hat{z}\hat{z})\mathbb{1}\mathbb{1} + \frac{\sqrt{3}}{2}(\hat{x}\hat{x} + \hat{y}\hat{y})\hat{z}\mathbb{1} + \frac{c_z}{2}(\hat{z}\mathbb{1} - \mathbb{1}\hat{z})\mathbb{1}\mathbb{1} + \frac{c_x}{2}(\hat{x}\hat{y} - \hat{y}\hat{x})\hat{z}\mathbb{1}\Big] \quad (12)$$

and with $W_{|M_{\vec{c}}}^{B \prec A}$ of a similar form, obtained from $W_{|M_{\vec{c}}}^{A \prec B}$ by changing $\hat{z}^{A_O}\mathbb{1}^{B_O}$ to $\mathbb{1}^{A_O}\hat{z}^{B_O}$ and c_x to $-c_x$.

Note that $W_{|M_{\vec{c}}}^{A \prec B}$ and $W_{|M_{\vec{c}}}^{B \prec A}$ are valid, causally ordered process matrices, compatible with $A \prec B$ and $B \prec A$, respectively (their eigenvalues are found to be 0, $\frac{1}{2}\left(1 \pm \sqrt{\frac{3 + c_x^2 + c_z^2}{4}}\right) \geqslant 0$ for $|\vec{c}| \leqslant 1$, and they satisfy the appropriate form of equation (19) given later). From equation (11) and the definition of causal separability in the bipartite case (Definition 1), we conclude that for any CP map (i.e. here, any POVM element) $M_{\vec{c}}$ applied by Charlie, $(W^{\text{act.}})_{|M_{\vec{c}}}$ is a (bipartite) causally separable process. Therefore, according to OG's definition 3, $W^{\text{act.}}$ is a tripartite CS process (with a single term in the decomposition (9),corresponding to C first).

A crucial feature of the decomposition (11) is that $(W^{\text{act.}})_{|M_{\vec{c}}}$, $W_{|M_{\vec{c}}}^{A \prec B}$ and $W_{|M_{\vec{c}}}^{B \prec A}$ all depend on Charlie's operation $M_{\vec{c}}$. Even though any valid process matrix in $C_I A_I B_I A_O B_O$ (including $W_{|M_{\vec{c}}}^{A \prec B}$ and $W_{|M_{\vec{c}}}^{B \prec A}$) is compatible with C acting last (since C has no outgoing system), the decomposition (11) still does not allow us to obtain a decomposition of the form $W^{\text{act.}} = \frac{1}{2}W^{A \prec B \prec C} + \frac{1}{2}W^{B \prec A \prec C}$ for $W^{\text{act.}}$ (or even with different weights $q, 1 - q$), as in equation (7). Indeed, such a decomposition for $W^{\text{act.}}$, with $W^{A \prec B \prec C}$ and $W^{B \prec A \prec C}$ valid process matrices compatible with the indicated causal order, does not exist. This can be shown using Araújo *et al*'s technique of 'witnesses of causal nonseparability' [14, 15]: one can construct a witness for $W^{\text{act.}}$, and we give one explicitly in appendix C.

Since, as stated above, the existence of such a decomposition (as in definition 2) would be equivalent in the scenario considered here to OG's ECS (Definition 4), this implies that although $W^{\text{act.}}$ is CS according to OG's definition (see above), it is not ECS. This provides an explicit example of 'activation of causal nonseparability' in that scenario.

Hence, OG's CS does not reduce (contrary to OG's ECS) to Araújo *et al*'s definition of causal separability in this particular scenario. Definitions 2 and 3 of causal separability are therefore inconsistent. Our aim now is to rectify this inconsistency.

3.4. Our choice of definition

To fix this, we now propose our own definition of multipartite causal separability, which indeed resolves the inconsistency pointed out above, and which we argue is a more natural definition for general multipartite scenarios. Similarly to OG, we choose a recursive definition, based on the concept of a conditional process matrix and very much in the spirit of the recursive definitions that have been given for multipartite causal correlations [17, 22]. For a process matrix to be compatible with a definite causal order, there should, in any run of the experiment, be a designated party that acts first (which party this is can be determined probabilistically,

[3] As C has no outgoing system, $W^{\text{act.}}$ is also compatible with C acting last (see appendix A.2.3). But to prove that $W^{\text{act.}}$ is CS (according to OG's definition) as we do below we need to consider C acting first.

just like in the bipartite case) and the conditional process matrix for the remaining parties, which depends on the action of the first party, should again be causally separable for any CP map that the first party applies.

For several reasons, we consider it important to allow extensions with extra incoming systems, similar to OG's *extensible* causal separability. Firstly, the whole process matrix framework is constructed so as to allow for shared ancillary systems between the parties. For consistency, we should thus take into account such extensions with shared incoming quantum states when defining causal (non)separability. Indeed, entanglement is a very different resource from causal nonseparability: entangled systems do not by themselves allow signalling between parties, and should be able to be distributed between parties prior to an experiment without 'activating' causal nonseparability. (Note, however, that entanglement can still play a crucial role in causal nonseparability, as e.g. in the quantum switch, where the control and target systems can end up being entangled after the parties' operations.) While a 'resource theory' for causal nonseparability has not yet been developed, it is reasonable to expect that providing additional shared (entangled) incoming states should be a free operation in such an approach. These considerations lead us to propose the following definition.

Definition 5 (*N*-partite causal separability). For $N = 1$, any process matrix is causally separable. For $N \geqslant 2$, an N-partite process matrix W is said to be *causally separable* if and only if, for any extension $A_{I'}^{\mathcal{N}}$ of the parties' incoming systems and any ancillary quantum state $\rho \in A_{I'}^{\mathcal{N}}$, $W \otimes \rho$ can be decomposed as

$$W \otimes \rho = \sum_{k \in \mathcal{N}} q_k \, W_{(k)}^{\rho}, \tag{13}$$

with $q_k \geqslant 0$, $\sum_k q_k = 1$, and where for each k, $W_{(k)}^{\rho} \in A_{II'O}^{\mathcal{N}}$ is a process matrix compatible with party A_k acting first, and is such that for any CP map $M_k \in A_{II'O}^{k}$ applied by party A_k, the conditional $(N-1)$-partite process matrix[4] $(W_{(k)}^{\rho})_{|M_k} \coloneqq \mathrm{Tr}_k[M_k \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W_{(k)}^{\rho}]$ is itself causally separable.

Note that there is a subtle difference between our definition here and that of OG's ECS (Definition 4). We indeed require all conditional process matrices appearing at all levels of the recursive decomposition to remain causally separable under extension with arbitrary ancillary states, while OG impose this *a priori* only for the original process matrix. In fact, although *prima facie* different, these definitions turn out to be equivalent; the proof of this is given in appendix D.

From definition 5 we recover the natural, intuitive definition of Araújo *et al* [14] in the particular tripartite case where one party has a trivial outgoing system—a case of practical relevance, as the quantum switch is the first example of a causally nonseparable process that has been demonstrated and studied in laboratory experiments [24, 25, 27]. One can also readily verify that process matrices that are causally separable by definition 5 cannot generate noncausal correlations (as defined in [17, 22]); an explicit proof is given in appendix E.

From now on, whenever we talk about causal (non)separability we will refer to our definition 5.

4. Characterising multipartite causal (non)separability

With the definition of causal (non)separability given above, we now turn to addressing the question of how to characterise causally separable process matrices in terms of simple conditions and how to demonstrate multipartite causal nonseparability in practice.

For that we will start by reviewing the characterisations of valid process matrices and of process matrices compatible with fixed causal orders, before recalling the characterisations of causally separable process matrices in the bipartite and tripartite cases, where we will give conditions for causal separability that are both necessary and sufficient. We will then present a generalisation to the N-partite case which, for $N \geqslant 4$, gives two conditions, one necessary and one sufficient, whose coincidence remains an open question.

In this section we will not concern ourselves with the normalisation of process matrices (which can always be imposed later). Our characterisations will then be given in terms of linear subspaces of matrices (e.g. the spaces $\mathcal{L}^{\mathcal{N}}$ and $\mathcal{L}^{A_1 \prec \cdots \prec A_N}$ introduced already in section 2); when adding the requirement of positive semidefiniteness, the corresponding sets of (nonnormalised) process matrices will thus be closed convex cones of positive semidefinite matrices. This will allow the conditions we give to be checked efficiently with semidefinite programming (SDP) techniques. In particular, by generalising the techniques used for the bipartite and restricted tripartite cases in [14, 15], we will extend the idea of witnesses of causal nonseparability to the multipartite case and show how multipartite witnesses can be constructed efficiently, allowing this causal

[4] Note that compared to equation (8), we take here $A_{II'O}^{k} \coloneqq A_{IO}^{k} \otimes A_{I'}^{k}$, $M_k \coloneqq M_{a_k|x_k}^{A_{II'O}^{k}}$, $\mathrm{Tr}_k \coloneqq \mathrm{Tr}_{A_{II'O}^{k}}$ and $\mathbb{1}^{\mathcal{N}\setminus k} \coloneqq \bigotimes_{j \in \mathcal{N}\setminus k} \mathbb{1}^{A_{II'O}^{j}}$ in the definition of the conditional matrix.

nonseparability to be verified experimentally by having each party perform appropriately chosen measurements [25, 27].

Following [14], we adopt the following notation, which will be used heavily throughout the rest of the paper:

$$ {}_X W \coloneqq (\mathrm{Tr}_X W) \otimes \frac{\mathbb{1}^X}{d_X}, \quad {}_1 W \coloneqq W, \quad {}_{[\sum_X \alpha_X X]} W \coloneqq \sum_X \alpha_X \, {}_X W, \tag{14} $$

with d_X the dimension of the Hilbert space of system X (note that $W \to {}_X W$ defines a CPTP map). In particular, constraints of the form ${}_{[1-X]} W = 0$ (which will appear regularly) therefore mean that W is of the form $W = \Omega \otimes \frac{\mathbb{1}^X}{d_X}$ (with $\Omega = \mathrm{Tr}_X W$).

4.1. Valid process matrices and compatibility with a fixed causal order

Recall from section 2 that the conditions for a process matrix W to be valid arise from requiring that the generalised Born rule (4) should give valid probability distributions, even when the parties share arbitrary ancillary systems. The fact that these probabilities should be nonnegative imposes that W must be positive semidefinite, while the requirement that these probabilities must sum to 1 implies that any valid (but, once again, not necessarily normalised) W must be in a subspace $\mathcal{L}^{\mathcal{N}}$ of $A_{IO}^{\mathcal{N}}$ [2, 14]. In appendix A.1 we recall the proof (following [14]) that this subspace can be characterised as follows:

$$ W \in \mathcal{L}^{\mathcal{N}} \Leftrightarrow \forall \ \mathcal{X} \subsetneq \mathcal{N}, \mathcal{X} \neq \varnothing, \mathrm{Tr}_{\mathcal{N}\setminus\mathcal{X}} W \in \mathcal{L}^{\mathcal{X}} \quad \text{and} \quad \prod_{i\in\mathcal{N}} {}_{[1-A_O^i]} W = 0 \tag{15} $$

$$ \Leftrightarrow \forall \ \mathcal{X} \subseteq \mathcal{N}, \mathcal{X} \neq \varnothing, \prod_{i\in\mathcal{X}} {}_{[1-A_O^i]} A_{IO}^{\mathcal{N}\setminus\mathcal{X}} W = 0. \tag{16} $$

Written in the form of equation (15), the validity constraint for W says that all reduced matrices $\mathrm{Tr}_{\mathcal{N}\setminus\mathcal{X}} W$ shared by the parties of any strict subset $\mathcal{X}$ of $\mathcal{N}$ (obtained after tracing out the parties that are not in $\mathcal{X}$) must be valid, and that W must further satisfy the additional constraint that $\prod_{i\in\mathcal{N}} {}_{[1-A_O^i]} W = 0$. The form of equation (16) expresses explicitly all the (linearly independent) constraints that these recursive validity conditions imply on[5] W. Denoting by $\mathcal{P}$ the convex cone of positive semidefinite matrices, the set of valid process matrices is then the convex cone

$$ \mathcal{W} = \mathcal{P} \cap \mathcal{L}^{\mathcal{N}}. \tag{17} $$

In order to discuss the causal separability of process matrices, it is necessary to also characterise the subspaces of such matrices that are compatible with certain fixed causal relations between (subsets of) parties. Such causal relations, as for the particular cases of fixed causal orders discussed in the previous sections, are understood via the notion of signalling: if a (group of) parties is in the causal future of some others, then there is no way for them to signal to those earlier parties.

We first consider the case of process matrices that are compatible with a given party A_k acting first[6]: regardless of the operation performed by the other parties $A_{k'}$ (for all $k' \neq k$), the marginal probability distribution for A_k obtained from (4) must not depend on the CPTP maps $M_{x_{k'}}^{A_{IO}^{k'}}$ chosen by those other parties. As already mentioned in the previous section and shown in appendix A.2, a given process matrix W satisfies this condition if and only if, whatever CP map M_k is applied by A_k, the conditional process matrix $W_{|M_k}$, as defined in equation (8), is a valid $(N-1)$-partite process matrix for the remaining parties in $\mathcal{N}\setminus A_k$.

We can in fact ignore here the assumption that $M_k \geqslant 0$, and the above constraint is equivalent to imposing that $W_{|M_k} \in \mathcal{L}^{\mathcal{N}\setminus A_k}$ for any $M_k \in A_{IO}^k$. Such a constraint defines a linear subspace of $A_{IO}^{\mathcal{N}}$. Taking its intersection with the subspace $\mathcal{L}^{\mathcal{N}}$, we denote the linear subspace of valid process matrices compatible with party A_k first by $\mathcal{L}^{A_k \prec (\mathcal{N}\setminus A_k)}$. We find, using equation (16) above (and after removing redundant constraints; see equation (A13) in appendix A.2):

$$ \begin{aligned} W \in \mathcal{L}^{A_k \prec (\mathcal{N}\setminus A_k)} &\Leftrightarrow W \in \mathcal{L}^{\mathcal{N}} \quad \text{and} \quad \forall \ M_k \in A_{IO}^k, \ W_{|M_k} \in \mathcal{L}^{\mathcal{N}\setminus A_k} \\ &\Leftrightarrow {}_{[1-A_O^k]} A_{IO}^{\mathcal{N}\setminus k} W = 0 \quad \text{and} \quad \forall \ \mathcal{X} \subseteq \mathcal{N}\setminus k, \mathcal{X} \neq 0, \prod_{i\in\mathcal{X}} {}_{[1-A_O^i]} A_{IO}^{\mathcal{N}\setminus k\setminus\mathcal{X}} W = 0. \end{aligned} \tag{18} $$

In appendix A.2 we also derive constraints for more general causal orders of the form $\mathcal{K}_1 \prec \mathcal{K}_2 \prec \cdots \prec \mathcal{K}_K$, for various disjoint subsets $\mathcal{K}_i$ of $\mathcal{N}$. Of particular interest is the specific case in which each $\mathcal{K}_i$ is a singleton, which gives constraints on a process matrix W being compatible with a fixed causal order such as $A_1 \prec A_2 \prec \cdots \prec A_N$. Such a W must be compatible with A_1 acting first (and must therefore satisfy

[5] Note that the constraint in equation (16) can also be written as $\prod_{i\in\mathcal{X}} {}_{[1-A_O^i]}(\mathrm{Tr}_{\mathcal{N}\setminus\mathcal{X}} W) = 0$. In this paper we generically use the form of equation (16) for ease of notation; it may be useful, however, to keep in mind that this type of constraint is in fact a constraint on the reduced matrix $\mathrm{Tr}_{\mathcal{N}\setminus\mathcal{X}} W$ shared by the parties in $\mathcal{X}$, as written more explicitly in equation (15).

[6] Note that a process matrix can be compatible with several different causal relations between parties. For example, if a matrix W does not allow any party to signal to another, then it is compatible with any party or group of parties acting first.

equation (18) for $k = 1$—in particular, the constraint on its third line); then, whatever CP map M_1 party A_1 applies, the resulting conditional process matrix $W_{|M_1}$ must then be a valid $(N-1)$-partite process matrix, compatible with party A_2 acting first (and must therefore satisfy equation (18) for $k = 2$—in particular, its third line—with $\mathcal{N}$ replaced by $\mathcal{N}\setminus\{1\}$); etc. By iterating this argument (up until the party A_N), we find that the linear subspace $\mathcal{L}^{A_1\prec\cdots\prec A_N}$ of process matrices compatible with the causal order $A_1 \prec \cdots \prec A_N$ is characterised by (see equation (A17)) [14, 30, 31]

$$W \in \mathcal{L}^{A_1\prec\cdots\prec A_N} \Leftrightarrow \forall\ k = 1, \ldots, N,\ {}_{[1-A_O^k]}A_{IO}^{(>k)}W = 0, \tag{19}$$

with $A_{IO}^{(>k)} = A_{IO}^{\{k+1,\ \ldots,\ N\}}$ (with $A_{IO}^{(>N)} = A_{IO}^{\varnothing} = 1$).

4.2. Bipartite and tripartite causally (non)separable process matrices

In the bipartite scenario, the above characterisation of the subspaces $\mathcal{L}^{A\prec B}$ and $\mathcal{L}^{B\prec A}$ allows us, from definition 1, to give the following explicit characterisation of causally separable process matrices.

Proposition 2 (Characterisation of bipartite causally separable process matrices). *A matrix $W \in A_{IO} \otimes B_{IO}$ is a valid bipartite causally separable process matrix if and only if it can be decomposed as*

$$W = W_{(A,B)} + W_{(B,A)}, \tag{20}$$

where, for each permutation (X, Y) of the two parties A and B, $W_{(X,Y)}$ is a positive semidefinite matrix satisfying

$${}_{[1-X_O]Y_{IO}}W_{(X,Y)} = 0,\ {}_{[1-Y_O]}W_{(X,Y)} = 0 \tag{21}$$

(i.e. $W_{(X,Y)}$ is a valid process matrix compatible with the causal order $X \prec Y$).

Note that, in contrast to equation (3) in definition 1, we did not write the weights q and $1-q$ explicitly in equation (20). Instead, for convenience and consistency with the characterisations of tripartite and N-partite causally separable processes which will follow, we decomposed W in terms of nonnormalised process matrices, writing $W_{(A,B)} = q\ W^{A\prec B}$ and $W_{(B,A)} = (1-q)\ W^{B\prec A}$.

As we discussed in section 3, the tripartite case of causal separability was already studied by Oreshkov and Giarmatzi under the name 'extensible causal separability' in [17]. In their Proposition 3.3 they provided a characterisation of tripartite (extensible) causal separability, albeit describing the constraints in a different way. In our approach, this characterisation can be expressed as follows:

Proposition 3 (Characterisation of tripartite causally separable process matrices). *A matrix $W \in A_{IO} \otimes B_{IO} \otimes C_{IO}$ is a valid tripartite causally separable process matrix (as per definition 5) if and only if it can be decomposed as*

$$\begin{aligned} W &= \overbrace{W_{(A)}}^{} + \overbrace{W_{(B)}}^{} + \overbrace{W_{(C)}}^{} \\ &= \overbrace{W_{(A,B,C)} + W_{(A,C,B)}}^{} + \overbrace{W_{(B,A,C)} + W_{(B,C,A)}}^{} + \overbrace{W_{(C,A,B)} + W_{(C,B,A)}}^{} \end{aligned} \tag{22}$$

where, for each permutation of the three parties (X, Y, Z), $W_{(X,Y,Z)}$ and $W_{(X)} := W_{(X,Y,Z)} + W_{(X,Z,Y)}$ are positive semidefinite matrices satisfying

$${}_{[1-X_O]Y_{IO}Z_{IO}}W_{(X)} = 0, \tag{23}$$

$${}_{[1-Y_O]Z_{IO}}W_{(X,Y,Z)} = 0, \quad {}_{[1-Z_O]}W_{(X,Y,Z)} = 0. \tag{24}$$

The proof of this characterisation was sketched in [17] using a somewhat different terminology to what we employ; in particular, they express causal constraints in terms of restrictions of what terms are 'allowed' in a Hilbert–Schmidt basis decomposition of a matrix (see appendix A.4). We give a more detailed proof in appendix B.1.2, which is again based on a 'teleportation technique' (see lemma B1 in appendix B), similar in spirit to the one briefly sketched in section 3.3.

Let us break down and analyse the terms appearing in the decomposition (22) to understand better this characterisation.

From the constraints in equation (24) it follows that, that for each party X, the matrix $W_{(X)} (= W_{(X,Y,Z)} + W_{(X,Z,Y)})$ satisfies ${}_{[1-Y_O]Z_{IO}}W_{(X)} = {}_{[1-Z_O]Y_{IO}}W_{(X)} = {}_{[1-Y_O][1-Z_O]}W_{(X)} = 0$. Together with equation (23) and the fact that $W_{(X)}$ is positive semidefinite, this implies that $W_{(X)}$ is a valid tripartite process matrix compatible with party X acting first (since it satisfies equation (18) for $A_k = X$). W is thus decomposed in equation (22) as a sum of three valid process matrices, which ensures in particular that it is itself a valid process matrix.

On the other hand, the matrices $W_{(X,Y,Z)}$ in the decomposition (22) are not necessarily valid process matrices. Nevertheless, the constraints (24) imply that whatever the CP map M_X applied by the first party X, the

conditional process matrix $(W_{(X,Y,Z)})_{|M_X} := \mathrm{Tr}_X[M_X \otimes \mathbb{1}^{YZ} \cdot W_{(X,Y,Z)}]$ is a valid bipartite process matrix, compatible with the causal order $Y \prec Z$ (indeed, it satisfies equation (19) for this causal order: e.g. $_{[1-Y_O]Z_{IO}}[(W_{(X,Y,Z)})_{|M_X}] = (_{[1-Y_O]Z_{IO}} W_{(X,Y,Z)})_{|M_X} = 0$).

The fact that the matrices $W_{(X,Y,Z)}$ are not necessarily valid process matrices, and thus that equation (22) does not simply decompose W into a combination of process matrices compatible with fixed causal orders, is a consequence of the possibility of dynamical (but still well-defined, albeit not *fixed*) causal orders (recall the discussion at the end of section 2.3). In section 4.5 we will consider in more detail a concrete example of a process matrix allowing for such dynamical causal orders.

4.3. General multipartite causally (non)separable process matrices

As we will see below, it is possible to generalise the decomposition of proposition 3 to the case of N-partite causal separability. While the generalisation clearly provides a sufficient condition for causal separability, it turns out that the proof that it is also a necessary condition does not readily generalise. Indeed, the proof for the tripartite case in appendix B.1.2 relies on the fact that each term $W_{(X)}$ in equation (22) is the sum of only two 'base' terms, something that is not true in the natural generalisation of this decomposition. (To understand this better, we encourage the interested reader to look at the subtleties of that proof.)

For the general multipartite case, we therefore provide the following, separate, necessary and sufficient conditions. Since these arise from different considerations, we will present and discuss these individually. Indeed, although these coincide in the bipartite and tripartite cases, it remains an open question whether this is the case in general (or if one is both necessary and sufficient but not the other, or if neither are).

4.3.1. Necessary condition

The necessary condition we present here is based on the teleportation technique and is a generalisation of the use of this approach in the proof of the tripartite characterisation. The teleportation technique is more formally described in lemma B1 in appendix B, but we briefly outline how it leads to the necessary condition to help understand the condition itself. The idea is to consider, in equation (13) of definition 5, a specific shared incoming ancillary state, as well as specific operations M_k applied by the parties A_k, for which there is a straightforward relation between the forms of the respective N-partite process matrices in which A_k acts first, and the corresponding $(N-1)$-partite conditional process matrices that we obtain after A_k has operated. As the latter are by definition causally separable (and satisfy thus the necessary conditions for $(N-1)$-partite causal separability), this allows us to infer necessary conditions for the causal separability of the original N-partite process matrix.

More precisely, we provide, as ancillary incoming systems, a maximally entangled state between every pair of parties, defining an overall ancillary state ρ. If W is a causally separable process matrix, then, by definition, $W \otimes \rho$ can be decomposed into a sum of process matrices $W^{\rho}_{(k)}$ compatible with a given party A_k acting first (see equation (13) in definition 5); furthermore, as ρ is pure, one can write $W^{\rho}_{(k)} = W_{(k)} \otimes \rho$ with $W_{(k)}$ itself being compatible with A_k first. For each such process matrix $W_{(k)}$ the party A_k can then 'teleport' the part of $W_{(k)}$ on their systems A^k_{IO} to another party $A_{k'}$ by applying an appropriate CP map M_k. The effect is that the resulting $(N-1)$-partite conditional process matrix $(W^{\rho}_{(k)})_{|M_k}$ formally has the same form as $W_{(k)}$ (tensored with what is left over of the, now reduced, ancillary state ρ), except that the systems A^k_{IO} are instead attributed ('teleported') to the ancillary incoming system $A^{k'}_{I'}$ of $A_{k'}$. From the definition of causal separability, $(W^{\rho}_{(k)})_{|M_k}$ must itself be causally separable, so the necessary condition can be recursively applied to this $(N-1)$-partite process matrix until the base case of $N = 3$, given by proposition 3, is reached.

We give the full details of the proof of the necessary condition in appendix B.2.1. However, in order to state more formally the condition itself, let us introduce the following notation. For a given matrix $W \in A^{\mathcal{N}}_{IO}$, we denote by $W^{A^k_{IO} \to A^{k'}_{I'}} \in A^{\mathcal{N}\setminus k}_{IO} \otimes A^{k'}_{I'}$ the same matrix, where the systems A^k_{IO} are attributed to some other system $A^{k'}_{I'}$ (of the same dimension as A^k_{IO}). More formally,

$$W^{A^k_{IO} \to A^{k'}_{I'}} := \sum_{i,j} \mathrm{Tr}_k[|i\rangle\langle j|^{A^k_{IO}} \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W] \otimes |j\rangle\langle i|^{A^{k'}_{I'}}, \tag{25}$$

where $\{|i\rangle\}$ is an orthonormal basis of $\mathcal{H}^{A^k_I} \otimes \mathcal{H}^{A^k_O}$.

We then obtain the following recursive necessary condition:

Proposition 4 (Necessary condition for general multipartite causal separability). *An N-partite causally separable process matrix $W \in A^{\mathcal{N}}_{IO}$ (as per definition 5) must necessarily have a decomposition of the form*

$$W = \sum_{k \in \mathcal{N}} W_{(k)}, \tag{26}$$

where each $W_{(k)}$ is a valid process matrix compatible with party A_k acting first, and such that for each $k' \neq k$, $W_{(k)}^{A_{IO}^k \to A_{I'}^{k'}}$ is an $(N-1)$-partite causally separable process matrix.

Hence, any constraints satisfied by $(N-1)$-partite causally separable process matrices must also be satisfied by $W_{(k)}$ after re-attributing the system $A_{I'}^{k'}$ back to A_{IO}^k—i.e. after formally replacing $A_I^{k'}$ by $A_I^{k'} A_{I'}^{k'}$ and then $A_{I'}^{k'}$ by A_{IO}^k in the constraints written using the notation defined in equation (14).

The decomposition of equation (26) follows from that of equation (13) in our definition of causal separability, for the appropriate choice of ancillary state and CP maps, as described above (see appendix B.2.1).

To further clarify this condition, let us illustrate, in the fourpartite case (with parties A, B, C, D), how one can use it to obtain explicit constraints on causally separable process matrices. Proposition 4 implies that a fourpartite causally separable process matrix W must be decomposable as

$$W = W_{(A)} + W_{(B)} + W_{(C)} + W_{(D)}, \tag{27}$$

with each $W_{(X)}$ (for $X = A, B, C, D$) being a valid process matrix compatible with party X acting first—hence satisfying equation (18) for $A_k = X$.[7] For each X and every other party $Y \neq X$, the recursive constraint that $W_{(X)}^{X_{IO} \to Y_{I'}}$ is a tripartite causally separable process matrix further implies, according to proposition 3 (for the three parties Y, Z, $T \neq X$) and after re-attributing the system $Y_{I'}$ to X_{IO} (i.e. replacing Y_{IO} by $Y_{I'} Y_{IO}$ and then $Y_{I'}$ by X_{IO} in the constraints), that there must exist a decomposition of $W_{(X)}$ of the form[8]

$$\begin{aligned} W_{(X)} &= W_{(X,Y)}^{[X\to Y]} + W_{(X,Z)}^{[X\to Y]} + W_{(X,T)}^{[X\to Y]} \\ &= W_{(X,Y,Z,T)}^{[X\to Y]} + W_{(X,Y,T,Z)}^{[X\to Y]} + W_{(X,Z,Y,T)}^{[X\to Y]} \\ &\quad + W_{(X,Z,T,Y)}^{[X\to Y]} + W_{(X,T,Y,Z)}^{[X\to Y]} + W_{(X,T,Z,Y)}^{[X\to Y]}, \end{aligned} \tag{28}$$

where each term appearing in the decomposition is positive semidefinite, $W_{(X,Y)}^{[X\to Y]} = W_{(X,Y,Z,T)}^{[X\to Y]} + W_{(X,Y,T,Z)}^{[X\to Y]}$, etc, and with (for all $X \neq Y \neq Z \neq T$)

$$\begin{aligned} [1-Y_O] Z_{IO} T_{IO} W_{(X,Y)}^{[X\to Y]} &= [1-Z_O] X_{IO} Y_{IO} T_{IO} W_{(X,Z)}^{[X\to Y]} = 0, \\ [1-Z_O] T_{IO} W_{(X,Y,Z,T)}^{[X\to Y]} &= [1-T_O] W_{(X,Y,Z,T)}^{[X\to Y]} = 0, \\ [1-Y_O] T_{IO} W_{(X,Z,Y,T)}^{[X\to Y]} &= [1-T_O] W_{(X,Z,Y,T)}^{[X\to Y]} = 0, \\ [1-T_O] X_{IO} Y_{IO} W_{(X,Z,T,Y)}^{[X\to Y]} &= [1-Y_O] W_{(X,Z,T,Y)}^{[X\to Y]} = 0. \end{aligned} \tag{29}$$

Finally, we remark that the constraints obtained by considering teleporting each party X's system to just a single other party Y (i.e. by just demanding the existence of a decomposition of the above form for some other party Y, rather than for *all* other parties $Y \neq X$) yields conditions that are still necessary for the causal separability of W, but which are generally weaker than those given in proposition 4. Indeed, in appendix F.1 we give an example of a fourpartite process matrix which satisfies those weaker conditions but not all of those given above.

4.3.2. Sufficient condition

A sufficient condition for causal separability can be obtained by considering a stricter form of the recursive decomposition (26) in proposition 4. In particular, we demand that W has a decomposition into $W_{(k)}$ compatible with A_k acting first and such that each $W_{(k)}$ itself recursively satisfies the sufficient constraints for an $(N-1)$-partite process matrix without A_{IO}^k being traced out. One can easily verify that the decomposition (22) in the tripartite case is a generalisation of this kind from the bipartite case. In the fourpartite case described explicitly above, this means that for each party X there should be a single decomposition of the form (28) (i.e. no longer dependent on Y) such that the constraints (29) are satisfied without tracing out X_{IO} on the first and fourth lines. The fact that, unlike in the necessary conditions, we only consider a single (recursive) decomposition of each $W_{(k)}$ means that we can give a more explicit formulation for the sufficient condition.

Before stating the sufficient condition, let us introduce some more notations. Let Π denote the set of permutations (generically denoted by π) of $\mathcal{N}$. For an ordered subset $(k_1, \ldots, k_n)$ of $\mathcal{N}$ with n elements (with $1 \leqslant n \leqslant N$, $k_i \neq k_j$ for $i \neq j$), let $\Pi_{(k_1, \ldots, k_n)}$ be the set of permutations of $\mathcal{N}$ for which the element k_1 is first, k_2

[7] Note that the existence, for all Y, of a decomposition of the form of equation (28) satisfying equation (29) implies all the constraints of equation (18), except for the third line (i.e. $[1-X_O] Y_{IO} Z_{IO} T_{IO} W = 0$).

[8] Here the superscripts $[X \to Y]$ are simply labels to indicate that, for each matrix $W_{(X)}$, there are potentially different decompositions of the form (28) for each $Y \neq X$. (The sufficient condition below will in fact precisely be obtained by assuming that these decompositions do not depend on Y.)

is second, …, and k_n is n^{th}—i.e. $\Pi_{(k_1,\,\ldots,\,k_n)} = \{\pi \in \Pi | \pi(1) = k_1, \ldots, \pi(n) = k_n\}$. With these notations, we have the following sufficient condition, that directly generalises the decomposition of proposition 3.

Proposition 5 (Sufficient condition for general multipartite causal separability). *If a matrix $W \in A_{IO}^{\mathcal{N}}$ can be decomposed as a sum of $N!$ positive semidefinite operators $W_\pi \geqslant 0$ in the form*

$$W = \sum_{\pi \in \Pi} W_\pi, \tag{30}$$

such that for any ordered subset of parties $(k_1, \ldots, k_n)$ of $\mathcal{N}$ (with $1 \leqslant n \leqslant N$, $k_i \neq k_j$ for $i \neq j$), the partial sum

$$W_{(k_1,\,\ldots,\,k_n)} := \sum_{\pi \in \Pi_{(k_1,\,\ldots,\,k_n)}} W_\pi \tag{31}$$

satisfies

$$[1 - A_O^{k_n}] A_{IO}^{\mathcal{N}\setminus\{k_1,\,\ldots,\,k_n\}} W_{(k_1,\,\ldots,\,k_n)} = 0, \tag{32}$$

then W *is a valid causally separable process matrix (as per definition 5).*

This decomposition was also suggested independently by Oreshkov as a possible generalisation of proposition 3 [32] (although following the approach of [2, 17], Oreshkov expressed it differently, namely in terms of allowed terms in a Hilbert–Schmidt basis decomposition of the matrices $W_{(k_1,\,\ldots,\,k_n)}$; see appendix A.4). The proof that the condition above is indeed sufficient is given in appendix B.2.2. In order to understand it better, it is nonetheless worth discussing the form of the decomposition and the terms appearing within in a little more detail.

Firstly, note that one can easily show by induction (see appendix B.2.2), that if equation (32) is satisfied for all $(k_1, \ldots, k_n)$, then one also has, for all $(k_1, \ldots, k_n)$ with $1 \leqslant n < N$, that

$$\forall\ \mathcal{X} \subseteq \mathcal{N}\setminus\{k_1, \ldots, k_n\},\ \mathcal{X} \neq \varnothing,\ \prod_{i \in \mathcal{X}} [1 - A_O^i] A_{IO}^{\mathcal{N}\setminus\{k_1,\,\ldots,\,k_n\}\setminus\mathcal{X}} W_{(k_1,\,\ldots,\,k_n)} = 0. \tag{33}$$

Note also that since all $W_\pi \geqslant 0$, all $W_{(k_1,\,\ldots,\,k_n)} \geqslant 0$ as well.

For $n = 1$, equations (32) and (33) imply that each matrix $W_{(k_1)}$ ($\geqslant 0$) is a valid process matrix compatible with party A_{k_1} acting first; indeed, equation (18) is satisfied for $A_k = A_{k_1}$. As $W = \sum_{k_1} W_{(k_1)}$ according to equations (30)–(31), this ensures in particular that W is indeed a valid process matrix.

Note, however, that in general the matrices $W_{(k_1,\,\ldots,\,k_n)}$ for $n > 1$ are *not* valid processes matrices compatible with the causal order $A_{k_1} \prec \cdots \prec A_{k_n}$. Indeed, as we already observed in the tripartite case, $W_{(k_1,\,\ldots,\,k_n)}$ may not generally be a valid process matrix at all. Nevertheless, comparing with equation (18), one can see that equations (32) and (33) imply that whatever the CP maps $M_{k_1}, \ldots, M_{k_{n-1}}$ applied by the $n - 1$ parties $A_{k_1}, \ldots, A_{k_{n-1}}$, the conditional matrix $(W_{(k_1,\,\ldots,\,k_n)})_{|M_{k_1} \otimes \cdots \otimes M_{k_{n-1}}} := \mathrm{Tr}_{k_1,\,\ldots,\,k_{n-1}}[M_{k_1} \otimes \cdots \otimes M_{k_{n-1}} \otimes \mathbb{1}^{\mathcal{N}\setminus\{k_1,\,\ldots,\,k_{n-1}\}} \cdot W_{(k_1,\,\ldots,\,k_n)}]$ is a valid $(N - n + 1)$-partite process matrix, compatible with party A_{k_n} acting first.

As we have noted already, the condition of proposition 5 coincides, in the bipartite and tripartite cases, with those given in propositions 2 and 3, respectively. Indeed, for these cases, the necessary and sufficient conditions given here coincided. For four-or-more parties it remains an open question whether this is also the case. We performed several numerical searches for process matrices satisfying the necessary but not sufficient conditions (see appendix F.2) and failed to find any such examples, although the complexity of the numerical searches means that we caution against interpreting this as evidence that the conditions coincide in general. In appendix B.3, however, we show that they do coincide in the specific fourpartite case with $d_{D_O} = 1$. This is a rather restricted scenario (where any process matrix is compatible with D acting last), but nonetheless includes cases of interest such as the fourpartite variant of the quantum switch we discuss at the end of this section.

Finally, we note that the decomposition in proposition 5 has consequences beyond the definition of causal separability meriting additional interest: as we show elsewhere [33], it characterises precisely (i.e. providing a necessary and sufficient condition for) quantum circuits with classical control of causal order.

4.4. Witnesses of causal nonseparability

While the previous characterisations provide mathematical descriptions of causally (non)separable process matrices, an important problem is the ability to detect and certify causal nonseparability in practice. One approach that has been explored extensively is to show the violation of causal inequalities [2, 16, 22, 34, 35], which is indeed only possible (within the process matrix formalism) with causally nonseparable process matrices (see appendix E), and provides a device-independent certificate of noncausality. However, certain causally nonseparable process matrices are known not to violate any such inequalities—this is, e.g. the case for the quantum switch [14, 17].

Another approach, first introduced in [14] for the bipartite and restricted tripartite scenarios, and further studied in [15], is to construct *witnesses of causal nonseparability*—or '*causal witnesses*' for short. Here, we outline this approach before describing how the conditions given in the previous subsections allow us to construct causal witnesses for general multipartite scenarios. This will permit a full analysis of the examples in the following section, as well as the verification of certain results already claimed in previous sections. While the overall approach of causal witnesses—and their formulation as efficiently solvable SDP problems—in the general case mirrors that of the specific scenarios previously studied [14, 15], the validity of the generalisation rests on certain technical details which we prove in appendix G.

A causal witness is defined as a Hermitian operator S such that

$$\mathrm{Tr}[S \cdot W^{\mathrm{sep}}] \geqslant 0 \tag{34}$$

for all $W^{\mathrm{sep}} \in \mathcal{W}^{\mathrm{sep}}$, where $\mathcal{W}^{\mathrm{sep}} \subset \mathcal{W}$ is the set of causally separable process matrices. For any causally nonseparable W^{ns}, it is known that there exists a causal witness S such that $\mathrm{Tr}[S \cdot W^{\mathrm{ns}}] < 0$ [14, 15]. Given a process, a causal witness S can be 'measured' by having each party implement suitably chosen operations or measurements, providing a now device-dependent test of causal nonseparability. This approach has been used, e.g. to verify experimentally the causal nonseparability of two different implementations of the quantum switch [25, 27].

Propositions 2, 3, 4 and 5 allow for the characterisation of the convex cone $\mathcal{W}^{\mathrm{sep}}$ of causally separable processes—or, for the latter two propositions, outer and inner approximations $\mathcal{W}_+^{\mathrm{sep}}$ and $\mathcal{W}_-^{\mathrm{sep}}$ thereof—in terms of Minkowski sums and intersections of linear subspaces and of the cone of positive semidefinite operators $\mathcal{P}$. The set of causal witnesses is then precisely the dual cone of $\mathcal{W}^{\mathrm{sep}}$, $\mathcal{S} = (\mathcal{W}^{\mathrm{sep}})^*$ [14, 15]. A characterisation of $\mathcal{S}$ can, in general, be obtained from the description of $\mathcal{W}^{\mathrm{sep}}$ by using the following duality relations for any two nonempty closed convex cones $\mathcal{C}_1$ and $\mathcal{C}_2$ [36]:

$$(\mathcal{C}_1 + \mathcal{C}_2)^* = \mathcal{C}_1^* \cap \mathcal{C}_2^*, \quad (\mathcal{C}_1 \cap \mathcal{C}_2)^* = \mathcal{C}_1^* + \mathcal{C}_2^* \tag{35}$$

(where $\mathcal{C}_1 + \mathcal{C}_2 = \{c_1 + c_2 | c_1 \in \mathcal{C}_1,\ c_2 \in \mathcal{C}_2\}$ is the Minkowski sum of the two cones $\mathcal{C}_1$ and $\mathcal{C}_2$; note that all the cones we shall consider will be nonempty, closed and convex).

Since these cones are convex, the construction of causal witnesses (or of explicit decompositions of causally separable process matrices) can be efficiently performed with SDP, as first described in [14]; we will follow here the slightly different approach of [15]. The question of whether a given W is causally separable can be reformulated as the optimisation problem of how much white noise can be added to a process matrix before it becomes causally separable. Let $\mathbb{1}^\circ = \mathbb{1}^{A_{IO}^{\mathcal{N}}} \big/ \prod_{k \in \mathcal{N}} d_{A_I^k}$ be the 'white noise' process matrix (which corresponds to each each party just receiving a fully mixed state $\mathbb{1}^{A_I^k} \big/ d_{A_I^k}$, and is causally separable), and consider the noisy process matrix

$$W(r) = \frac{1}{1+r}(W + r\mathbb{1}^\circ). \tag{36}$$

Since the normalisation is irrelevant to membership of $\mathcal{W}^{\mathrm{sep}}$, determining whether W is causally separable can be thus phrased as the SDP optimisation problem

$$\begin{aligned} &\min r \\ &\text{s.t.} \quad W + r\mathbb{1}^\circ \in \mathcal{W}^{\mathrm{sep}}, \end{aligned} \tag{37}$$

which can be efficiently solved using standard software by writing $\mathcal{W}^{\mathrm{sep}}$ in terms of SDP constraints (see [15], the examples below and appendix G for further details). The solution to this problem, r^*, gives the *random robustness* $\max(r^*, 0)$ of W, and a value $r^* > 0$ implies that W is causally nonseparable [14, 15].

Equation (37) is known as the primal problem, and is related to the dual problem

$$\begin{aligned} &\min \mathrm{Tr}[S \cdot W] \\ &\text{s.t.} \quad S \in \mathcal{S} \quad \text{and} \quad \mathrm{Tr}[S \cdot \mathbb{1}^\circ] = 1, \end{aligned} \tag{38}$$

defined over the dual cone $\mathcal{S}$ of $\mathcal{W}^{\mathrm{sep}}$ [14, 15]. The optimal solution S^* is a witness of the causal nonseparability of W whenever $\mathrm{Tr}[S^* \cdot W] < 0$. The Strong Duality theorem for SDP problems moreover relates these two problems, stating that their solutions satisfy

$$r^* = -\mathrm{Tr}[S^* \cdot W]. \tag{39}$$

In appendix G we show that (38) is indeed the dual of (37) and that the Strong Duality theorem is indeed applicable for arbitrary scenarios, as well as giving some further details. This implies in particular that the witness S^* thus obtained is optimal when W is subject to white noise, in the sense that it witnesses the causal nonseparability of all noisy process matrices $W(r)$ with r sufficiently small ($r < r^*$) so as for $W(r)$ to remain causally nonseparable.

For more than three parties, the witnesses in the set $\mathcal{S}_+ = (\mathcal{W}_+^{\mathrm{sep}})^*$ obtained from the cone $\mathcal{W}_+^{\mathrm{sep}} \supseteq \mathcal{W}^{\mathrm{sep}}$ arising from the necessary condition of proposition 4 are also valid witnesses of $\mathcal{W}^{\mathrm{sep}}$ since $\mathcal{S}_+ \subseteq \mathcal{S}$. On the other hand, by solving the primal SDP problem over the cone $\mathcal{W}_-^{\mathrm{sep}}$ arising from the sufficient condition in proposition 5, one can show the causal separability of any $W \in \mathcal{W}_-^{\mathrm{sep}} \subseteq W^{\mathrm{sep}}$ (through the construction of an explicit causally

separable decomposition for W of the form given in proposition 5). Recalling the claim that such process matrices correspond precisely to quantum circuits with classical control of orders [33], the dual cone $\mathcal{S}_-$ is thus the set of 'witnesses for no classical control of causal order' (which can thus be found by solving the dual SDP problem).

In appendix G we give some concrete characterisations of the cones $\mathcal{W}^{\text{sep}}$ and $\mathcal{S}$ for different scenarios in terms of SDP constraints, and which are relevant for the examples that we shall now give in the following section.

4.5. Examples

In the bipartite scenario and restricted tripartite scenario in which C has no outgoing system, several examples of causally nonseparable process matrices have previously been formulated and studied in detail [2, 14, 17]. The characterisations of the cones of causal witnesses that we give in appendix G.2 for these bipartite and restricted tripartite scenarios (see equations (G3) and(G5)) are equivalent to those given in [14, 15], and can readily be used to verify the causal nonseparability of these examples, following the approach just outlined.

As mentioned already in section 3.1, the quantum switch is a particularly interesting example of a causally nonseparable process matrix in the second of these scenarios. In that same scenario we have in fact also already looked at another explicit example: the process matrix $W^{\text{act.}}$ (10) introduced in section 3.3 to show the 'activation of causal nonseparability' under OG's definition of causal separability. An explicit witness from the cone (G5) is given in appendix C, equation (C2), which could thus have been equally well found with the approach of [14, 15].

Another example of 'activation of causal nonseparability' under OG's terminology was given in [17] in the different tripartite case in which one party, say now A, has only a nontrivial outgoing system, and can thus always be seen as acting first. A witness for this example can be found by solving the dual SDP problem (38) using now the cone of witnesses (G7) corresponding to this restricted tripartite scenario.

Of more novel interest is the fourpartite scenario, in which causal separability has not previously been characterised. A particularly interesting and simple example here is a fourpartite version of the quantum switch, in which a party A(lice) has no incoming system ($d_{A_I} = 1$) and always acts first, while another party D(orothy) has no outgoing system ($d_{D_O} = 1$) and always acts last. Let us describe more precisely this version of the quantum switch.

The switch is composed of two qubits: a control qubit and a target qubit. Initially, Alice prepares the control qubit in some state of her choosing (in general as a function of her input x). (Note that it is here that the fourpartite switch described here differs from the tripartite one, where the control qubit is in a fixed superposition.) The target qubit, initially prepared (externally to the four parties) in some state $|\psi\rangle$, is then sent to Bob and Charlie, who act in an order that depends on the state of the control qubit: if it is $|0\rangle$ then Bob acts before Charlie ($B \prec C$), while if it is $|1\rangle$ then Charlie acts before Bob ($C \prec B$). If it is in a superposition, then Bob and Charlie can instead be seen to act in a superposition of different orders. Finally, both qubits are sent to Dorothy who can perform a measurement on them (for simplicity, we will consider that D simply ignores the target qubit and thus will trace it out, as this will not change the discussion that follows)[9]. Labelling the relevant incoming and outgoing systems (where the superscripts indicate control and target qubits) $A^c_O, B^t_I, B^t_O, C^t_I, C^t_O, D^t_I, D^c_I$, the process matrix for the quantum switch can be written [14, 17, 29]

$$\begin{aligned} W^{\text{switch}} &= \text{Tr}_{D^t_I}|w\rangle\langle w| \\ &\quad \text{with} \quad |w\rangle = |0\rangle^{A^c_O}|\psi\rangle^{B^t_I}|\mathbb{1}\rangle\!\rangle^{B^t_O C^t_I}|\mathbb{1}\rangle\!\rangle^{C^t_O D^t_I}|0\rangle^{D^c_I} \\ &\quad + |1\rangle^{A^c_O}|\psi\rangle^{C^t_I}|\mathbb{1}\rangle\!\rangle^{C^t_O B^t_I}|\mathbb{1}\rangle\!\rangle^{B^t_O D^t_I}|1\rangle^{D^c_I}, \end{aligned} \tag{40}$$

where $|\mathbb{1}\rangle\!\rangle := |0\rangle|0\rangle + |1\rangle|1\rangle$ is the pure CJ representation (in the computational basis $\{|0\rangle, |1\rangle\}$) of an identity qubit channel. Note that, while Alice has control over the causal order of the other parties, this switch differs from a classical dynamical control of causal order in that she has coherent quantum control over the control qubit (and thus the causal orders).

In this particular restricted fourpartite scenario, our necessary and sufficient conditions for the causal separability of a process matrix W coincide and reduce to the existence of a decomposition of the form $W = W_{(A,B,C,D)} + W_{(A,C,B,D)}$ with $W_{(A,B,C,D)}, W_{(A,C,B,D)} \geqslant 0$ (which need not be valid process matrices) satisfying $_{[1-B_O]C_{IO}D_I}W_{(A,B,C,D)} = {}_{[1-C_O]D_I}W_{(A,B,C,D)} = {}_{[1-C_O]B_{IO}D_I}W_{(A,C,B,D)} = {}_{[1-B_O]D_I}W_{(A,C,B,D)} = 0$ (and with $_{[1-A_O]B_{IO}C_{IO}D_I}W = 0$ to ensure, with the previous constraints, that W is valid); see proposition B7 in appendix B.3. These conditions thus characterise precisely the cone $\mathcal{W}^{\text{sep}}$ in the scenario considered here, and

[9] We note that the quantum switch was also described as a fourpartite process in [25], with one party acting first, and one acting last. However, in that reference the first party was controlling the target qubit, rather than the control qubit as we consider here. In that case (with the first party controlling the target qubit), the random robustness is increased to 2.767. One could also have here a first party that controls both the target and control qubits (as in [29]), which further increases the tolerable white noise to 4.686; for simplicity we do not consider this possibility, as our goal here is just to illustrate the role of the control qubit. Note also that Rubino *et al* [25] used yet another definition of causal nonseparability, different from the ones discussed in section 3, which did not allow for dynamical causal orders. As argued before and discussed in [17, 22], such a definition is however too restrictive to really characterise processes that are compatible with a well-defined causal order, as one would like the notion of causal separability to do. Nevertheless, it turns out that the witness constructed and experimentally tested in [25] is not only a witness for fixed (nondynamical) causal orders, but also witnesses causal nonseparability as per our definition 5.

the dual cone of causal witnesses $\mathcal{S}$ is then readily obtainable (see equation (G11) for the explicit characterisation).

The causal nonseparability of W^{switch} can thus be verified by solving the dual SDP problem (38) and thereby obtaining a witness of its causal nonseparability. Doing so, we find that (up to numerical precision) the random robustness of W^{switch} of 2.343 (note that this does not depend on the choice of initial state of the target qubit, so in solving the SDP problem numerically we can take, e.g. $|\psi\rangle = |0\rangle$). In experimental efforts to measure a witness and verify the causal nonseparability of a process matrix, one may only have access to a restricted set of operations for the parties. Many natural such constraints can also be imposed as SDP constraints, as described in [15], allowing one to find implementable causal witnesses. A particularly natural such constraint is to restrict B and C's operations to unitary operations (as in the experimental implementation of the tripartite switch in [24, 27]); we find that the tolerable white noise on W^{switch} to witness its causal nonseparability is reduced, under such a restriction, to 0.746.

It is important to note that if we trace out the last party from W^{switch} (i.e. D_I^c in addition to D_I^t), we obtain

$$\text{Tr}_D\, W^{\text{switch}} = |0\rangle\langle 0|^{A_O^c}|\psi\rangle\langle\psi|^{B_I^t}|\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{B_O^t C_I^t}\mathbb{1}^{C_O^t} + |1\rangle\langle 1|^{A_O^c}|\psi\rangle\langle\psi|^{C_I^t}|\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{C_O^t B_I^t}\mathbb{1}^{B_O^t}, \tag{41}$$

which is causally separable since it is of the form of equation (22) with just the first two terms being nonzero: $\text{Tr}_D\, W^{\text{switch}} = W_{(A)} = W_{(A,B,C)} + W_{(A,C,B)}$, with $W_{(A,B,C)}$, $W_{(A,C,B)}$ (defined as the first and second terms in equation (41) above, respectively) and $W_{(A)}$ satisfying the constraints of equations (23)–(24). This was also the case with the original tripartite version of the quantum switch (in which the control qubit is in the fixed state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$). There, one is left with a simple probabilistic mixture of channels in two different directions after tracing out the last party [14, 17]. In contrast here, equation (41) is not compatible with any probabilistic mixture of fixed causal orders: indeed, $W_{(A,B,C)}$ and $W_{(A,C,B)}$ are not valid process matrices, as ${}_{[1-A_O]B_{IO}C_{IO}}W_{(A,B,C)} = -{}_{[1-A_O]B_{IO}C_{IO}}W_{(A,C,B)} \neq 0$ (these terms cancel in the sum $W_{(A,B,C)} + W_{(A,C,B)}$, so that ${}_{[1-A_O]B_{IO}C_{IO}}W_{(A)} = 0$ as required for $W_{(A)}$ to be a valid process matrix). Rather, $\text{Tr}_D\, W^{\text{switch}}$ is a 'classical switch' in which A can incoherently control the causal order between B and C, which thus allows for dynamical causal orders.

5. Discussion

In this paper we studied the question of how to generalise the concept of causal (non)separability to the multipartite case. We reviewed several definitions that had been proposed for multipartite scenarios in previous works, namely the definition of causal separability introduced by Araújo *et al* [14] for a particular tripartite situation, and Oreshkov and Giarmatzi's definitions of causal separability (CS) and extensible causal separability (ECS) [17] for the general multipartite case. We established the equivalence between Araújo *et al*'s (restricted) definition of causal separability and Oreshkov and Giarmatzi's definition of ECS in the particular tripartite situation considered by Araújo *et al*, thus linking two *a priori* different definitions for that case. Moreover, by showing that ECS and CS are different in that scenario, we found that the two definitions of causal separability proposed by Araújo *et al* [14] and by Oreshkov and Giarmatzi [17] were inconsistent, a problem that thus needed to be addressed.

We proposed a new general definition of N-partite causal nonseparability, similar in spirit to the recursive definitions that have been proposed for multipartite causal correlations [17, 22], and more consistent with the fact that the process matrix framework always allows for parties to share additional ancillary systems. Our definition thus avoids some unwanted features of the definition of CS in [17], such as the 'activation' of causal nonseparability by shared entanglement. Moreover, we showed that our definition, although *a priori* different, in fact reduces to the notion of ECS proposed in [17], which also reduces to the definition of Araújo *et al* [14] in the particular restricted scenario considered there.

We then focused on characterising causally separable process matrices, giving (in the general multipartite case) two conditions—one necessary and one sufficient (Propositions 4 and 5, respectively)—for a given process matrix to be causally separable. These conditions allowed us to characterise the corresponding sets of process matrices through SDP constraints, and to generalise the tool of witnesses for causal nonseparability to the multipartite case. In the bipartite and tripartite cases, our necessary and sufficient conditions coincide and reduce to those previously described [2, 14, 17]. The principal open question raised by this work is whether this also holds in the general N-partite case with $N \geqslant 4$, or whether one of the two is both necessary and sufficient (or if one could derive yet another distinct condition, that would we both necessary and sufficient).

As we show elsewhere, our sufficient condition characterises precisely the processes that can be realised as a quantum circuit with classical control of causal order [33]. If that condition is in fact also necessary, this would thus confirm the conjecture of Oreshkov and Giarmatzi, that causally separable process matrices (or 'extensibly causally separable processes' using their terminology) are those realisable by such 'classically controlled quantum circuits' [17]. This would provide more solid founding for our understanding of the notion of causal separability, which would then indeed correspond to our intuition (quantum circuits with possibly dynamical

causal orders that are classically controlled). Furthermore, the proof in [33] would also provide a general explicit construction to realise any given causally separable process matrix in practice.

However, the forms of our necessary and sufficient condition, and the fact that the proof for the necessity of the conditions in the tripartite case does not generalise straightforwardly to more parties, indeed leave open the possibility that our sufficient condition may turn out to not be necessary. If this is the case, it would mean that there exist causally separable process matrices that are not realisable as classically controlled quantum circuits—and which we would not currently know how to realise experimentally. It would certainly be interesting to understand what kind of situations such process matrices correspond to—and if (and how) they can be realised quantum mechanically. This question is reminiscent of the open problem of whether process matrices that allow for the violation of causal inequalities are realisable with 'standard' quantum mechanics. Here the question would concern even less extreme situations: causally separable process matrices.

Another question that arises naturally in the multipartite case is whether a given phenomenon is *genuinely* multipartite, in the sense that its occurrence truly requires the coordinated action of a certain number of parties. It would be important for our understanding of multipartite process matrices to define a notion of 'genuinely multipartite causal nonseparability', similar to the concept of 'genuinely multipartite noncausality' for correlations [35] and analogous to the notions of genuinely multipartite entanglement [37] and nonlocality [38–40]. It would then also be interesting to study whether the definition can be refined to give a hierarchy of degrees of causal nonseparability, similar to the approach in [35] for correlations, and whether the characterisation of the corresponding process matrices and the construction of 'witnesses of genuinely multipartite causal nonseparability' are still possible with SDP techniques. These questions are left for further research.

Finally, our clarification of the definition of causal separability in the N-partite scenario, as well as characterisations of causally separable process matrices, should be helpful in the study of causal nonseparability as a computational and information processing resource [4–13]. Indeed, since multipartite scenarios offer significantly richer structure, understanding these scenarios is a prerequisite to grasping the full possibilities of causal indefiniteness as a resource, and an important step towards developing a resource theory of noncausality [41, 42].

Acknowledgments

We acknowledge fruitful discussions with F Costa, C Giarmatzi and O Oreshkov, and financial support from the 'Retour Post-Doctorants' (ANR-13-PDOC-0026) and the 'Investissements d'avenir' (ANR-15-IDEX-02) programs of the French National Research Agency.

Appendix A. Characterisation of process matrices

In this first appendix we show how the valid process matrices, as well as those compatible with a given causal order, can be characterised. We then discuss some properties of process matrices, and alternative characterisations.

Recall that in the construction of the process matrix framework (as in section 2.1), the CJ isomorphism [20, 21] is used to represent the parties' operations. Different versions of the CJ isomorphism exist; following [2], one may for instance define the CJ representation of a linear map $\mathcal{M}\colon A_I \to A_O$ as

$$M := [\mathcal{I} \otimes \mathcal{M}(|\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|)]^{\mathrm{T}} = \left[\sum_{m,m'} |m\rangle\langle m'| \otimes \mathcal{M}(|m\rangle\langle m'|)\right]^{\mathrm{T}} \in A_I \otimes A_O, \tag{A1}$$

where $\mathcal{I}$ is the identity channel, $\{|m\rangle\}_m$ is a fixed orthonormal basis of $\mathcal{H}^{A_I}$, $|\mathbb{1}\rangle\!\rangle := \sum_m |m\rangle \otimes |m\rangle$, and T denotes transposition in the chosen basis $\{|m\rangle\}_m$ of $\mathcal{H}^{A_I}$ and some fixed basis of $\mathcal{H}^{A_O}$. From its CJ representation M it is easy to recover the map $\mathcal{M}$, using $\mathcal{M}(\rho) = [\mathrm{Tr}_{A_I}[M \cdot \rho \otimes \mathbb{1}^{A_O}]]^{\mathrm{T}}$; see for instance appendix A in [14] for more details. Referring to this isomorphism, in the following we always identify linear maps with their CJ representation; recall in particular that a linear map is CP if and only if its CJ representation is positive semidefinite, and it is trace preserving if and only if its CJ representation satisfies $\mathrm{Tr}_{A_O} M = \mathbb{1}^{A_I}$.

A.1. Valid process matrices

A given matrix $W \in A_{IO}^{\mathcal{N}}$ defines a valid N-partite process matrix if and only if it generates nonnegative and normalised probabilities $P(\vec{a}|\vec{x})$ through the generalised Born rule of equation (4)—including in the case where an ancillary quantum state ρ in some extension $A_{I'}^{\mathcal{N}}$ of the parties' incoming spaces is attached to W (and thus

shared among the N parties), and the parties' operations are allowed to act on their joint incoming systems $A^k_{II'} \coloneqq A^k_I \otimes A^k_{I'}$.

The constraint that the probabilities in equation (4) are nonnegative for any set of CP maps—i.e. any positive semidefinite matrices $M^{A^k_{IO}}_{a_k|x_k}$—and that this remains the case when attaching any ancillary quantum state $\rho \in A^{\mathcal{N}}_{I'}$ (and for any $M^{A^k_{II'O}}_{a_k|x_k} \geqslant 0$), translates into the constraint that W must be positive semidefinite [2, 14].

As for normalisation, the constraint is that equation (4) must give a total probability equal to 1 for any set of CPTP maps—i.e. any positive semidefinite matrices $M^{A^k_{IO}}_{x_k}$ satisfying ${}_{A^k_O}M^{A^k_{IO}}_{x_k} = \frac{\mathbb{1}}{d_{A^k_O}}$, using the 'trace-out-and-replace' notation ${}_X\cdot$ defined in equation (14). It is easy to see that the constraint of positive semidefiniteness does not play any role here; furthermore, note that for any matrix $M \in A_{IO}$, $M' \coloneqq \frac{\mathbb{1}}{d_{A_O}} + {}_{[1-A_O]}M$ satisfies ${}_{A_O}M' = \frac{\mathbb{1}}{d_{A_O}}$ and that any $M \in A_{IO}$ satisfying ${}_{A_O}M = \frac{\mathbb{1}}{d_{A_O}}$ is of the form $M = \frac{\mathbb{1}}{d_{A_O}} + {}_{[1-A_O]}M$. The normalisation constraint thus translates into the constraint that

$$\mathrm{Tr}\left[\left(\frac{\mathbb{1}}{d_{A^1_O}} + {}_{[1-A^1_O]}M_1\right) \otimes \cdots \otimes \left(\frac{\mathbb{1}}{d_{A^N_O}} + {}_{[1-A^N_O]}M_N\right) \cdot W\right] = 1 \tag{A2}$$

for any set of matrices $M_1 \in A^1_{IO}, \ldots, M_N \in A^N_{IO}$. Expanding this constraint, and using the fact that the map ${}_{[1-A^k_O]}\cdot$ is self adjoint with respect to the trace (Hilbert–Schmidt) inner product (i.e. $\mathrm{Tr}[{}_{[1-A^k_O]}M \cdot W] = \mathrm{Tr}[M \cdot {}_{[1-A^k_O]}W]$), one finds that this is equivalent to

$$\mathrm{Tr}\, W = \prod_{k\in\mathcal{N}} d_{A^k_O} \qquad \text{and} \tag{A3}$$

$$\forall\ \mathcal{X} \subseteq \mathcal{N},\ \mathcal{X} \neq \varnothing,\ \prod_{i\in\mathcal{X}} {}_{[1-A^i_O]}\mathrm{Tr}_{\mathcal{N}\setminus\mathcal{X}}\, W = 0. \tag{A4}$$

Note that for simplicity we did not explicitly attach an ancillary state ρ to W here; doing so would have led to the same conclusion. Full details for this whole argument can be found in appendix B of [14].

We shall in general ignore the normalisation constraint (A3) when talking about valid process matrices. The $2^N - 1$ linear constraints of equation (A4) define a linear subspace of $A^{\mathcal{N}}_{IO}$, which we denote by $\mathcal{L}^{\mathcal{N}}$, the subspace of valid process matrices: explicitly (noting that the constraints $\prod_{i\in\mathcal{X}}{}_{[1-A^i_O]}\mathrm{Tr}_{\mathcal{N}\setminus\mathcal{X}}\, W = 0$ are equivalent to $\prod_{i\in\mathcal{X}}{}_{[1-A^i_O]}A^{\mathcal{N}\setminus\mathcal{X}}_{IO} W = 0$),

$$W \in \mathcal{L}^{\mathcal{N}} \Leftrightarrow \forall\ \mathcal{X} \subseteq \mathcal{N},\ \mathcal{X} \neq \varnothing,\ \prod_{i\in\mathcal{X}} {}_{[1-A^i_O]}A^{\mathcal{N}\setminus\mathcal{X}}_{IO} W = 0, \tag{A5}$$

as in equation (16) of the main text. It is furthermore straightforward to check that this is equivalent to the following recursive characterisation, as in equation (15):

$$W \in \mathcal{L}^{\mathcal{N}} \Leftrightarrow \forall\ \mathcal{X} \subsetneq \mathcal{N},\ \mathcal{X} \neq \varnothing,\ \mathrm{Tr}_{\mathcal{N}\setminus\mathcal{X}}\, W \in \mathcal{L}^{\mathcal{X}} \text{ and } \prod_{i\in\mathcal{N}} {}_{[1-A^i_O]}W = 0. \tag{A6}$$

Summing up, the set of (nonnormalised) valid process matrices is the convex cone $\mathcal{W} = \mathcal{P} \cap \mathcal{L}^{\mathcal{N}}$, where $\mathcal{P}$ is the cone of positive semidefinite matrices.

A.2. Compatibility with fixed causal orders

Let us now analyse the constraints imposed on process matrices by requiring that they are compatible with a given causal order.

A.2.1. Causal order between two subsets of parties. Consider two nonempty disjoint subsets of parties $\mathcal{K}_1, \mathcal{K}_2 \subsetneq \mathcal{N}$. We say that the correlation $P(\vec{a}|\vec{x})$ is compatible with the causal order $\mathcal{K}_1 \prec \mathcal{K}_2$ if and only if there is no signalling from the parties in $\mathcal{K}_2$ to the parties in $\mathcal{K}_1$—i.e. the marginal probability distribution for the outputs of parties in $\mathcal{K}_1$ does not depend on the inputs of parties in $\mathcal{K}_2$: $P(\vec{a}_{\mathcal{K}_1}|\vec{x}) = P(\vec{a}_{\mathcal{K}_1}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}_2})$ for all $\vec{x}$, $\vec{a}_{\mathcal{K}_1}$. We then say that a (valid) process matrix W is compatible with the causal order $\mathcal{K}_1 \prec \mathcal{K}_2$ if and only if it only generates correlations (through the generalised Born rule (4), possibly allowing for extensions of W with some ancillary state ρ) that are compatible with $\mathcal{K}_1 \prec \mathcal{K}_2$.

Formally, this means (ignoring again for simplicity the possibility of attaching an ancillary state ρ; as before, the same reasoning also goes through if we allow for this possibility) that whatever the CP maps $M^{A^{k_1}_{IO}}_{a_{k_1}|x_{k_1}}$ applied by the parties in $\mathcal{K}_1$ and whatever the CPTP maps $M^{A^{k_2}_{IO}}_{x_{k_2}}$ and $M^{A^{k_3}_{IO}}_{x_{k_3}}$ (such that ${}_{A^{k_2/3}_O}M^{A^{k_2/3}_{IO}}_{x_{k_2/3}} = \frac{\mathbb{1}}{d_{A^{k_2/3}_O}}$) applied by the parties in $\mathcal{K}_2$ and in $\overline{\mathcal{K}_{12}} \coloneqq \mathcal{N}\setminus(\mathcal{K}_1 \cup \mathcal{K}_2)$ (which may be empty), respectively, one must have

$$\mathrm{Tr}\left[\bigotimes_{k_1\in\mathcal{K}_1} M^{A_{IO}^{k_1}}_{a_{k_1}|x_{k_1}} \bigotimes_{k_2\in\mathcal{K}_2} M^{A_{IO}^{k_2}}_{x_{k_2}} \bigotimes_{k_3\in\overline{\mathcal{K}_{12}}} M^{A_{IO}^{k_3}}_{x_{k_3}} \cdot W\right] = \mathrm{Tr}\left[\bigotimes_{k_1\in\mathcal{K}_1} M^{A_{IO}^{k_1}}_{a_{k_1}|x_{k_1}} \bigotimes_{k_2\in\mathcal{K}_2} \frac{\mathbb{1}}{d_{A_O^{k_2}}} \bigotimes_{k_3\in\overline{\mathcal{K}_{12}}} M^{A_{IO}^{k_3}}_{x_{k_3}} \cdot W\right] \tag{A7}$$

(i.e. the probabilities should be the same if the parties in $\mathcal{K}_2$ apply the CPTP maps $\frac{\mathbb{1}}{d_{A_O^{k_2}}}$ instead of $M^{A_{IO}^{k_2}}_{x_{k_2}}$). As in the previous subsection, the constraint of positive semidefiniteness of the CJ matrices M_k does not play any role here, and we can equivalently write the above constraint as

$$\begin{aligned}&\mathrm{Tr}\left[\bigotimes_{k_1\in\mathcal{K}_1} M_{k_1} \bigotimes_{k_2\in\mathcal{K}_2}\left(\frac{\mathbb{1}}{d_{A_O^{k_2}}} + [1-A_O^{k_2}]M_{k_2}\right) \bigotimes_{k_3\in\overline{\mathcal{K}_{12}}}\left(\frac{\mathbb{1}}{d_{A_O^{k_3}}} + [1-A_O^{k_3}]M_{k_3}\right)\cdot W\right]\\ &= \mathrm{Tr}\left[\bigotimes_{k_1\in\mathcal{K}_1} M_{k_1} \bigotimes_{k_2\in\mathcal{K}_2}\frac{\mathbb{1}}{d_{A_O^{k_2}}} \bigotimes_{k_3\in\overline{\mathcal{K}_{12}}}\left(\frac{\mathbb{1}}{d_{A_O^{k_3}}} + [1-A_O^{k_3}]M_{k_3}\right)\cdot W\right]\end{aligned} \tag{A8}$$

for any matrices $M_k \in A_{IO}^k$. Expanding this constraint in a similar way as above (or as it was done in more details in appendix B of [14]), we find that it is equivalent to the following $2^{N-|\mathcal{K}_1|} - 2^{N-|\mathcal{K}_1|-|\mathcal{K}_2|}$ linear constraints:

$$\forall\ \mathcal{X}_2 \subseteq \mathcal{N}\backslash\mathcal{K}_1,\ \mathcal{X}_2\cap\mathcal{K}_2 \neq \varnothing,\ \prod_{i\in\mathcal{X}_2}[1-A_O^i]A_{IO}^{\mathcal{N}\backslash\mathcal{K}_1\backslash\mathcal{X}_2}W = 0. \tag{A9}$$

Combining these conditions with those from equation (A5) to ensure that W is a valid process matrix, and removing redundant constraints[10], one can then characterise the subspace $\mathcal{L}^{\mathcal{K}_1\prec\mathcal{K}_2}$ of (valid) process matrices compatible with the causal order $\mathcal{K}_1 \prec \mathcal{K}_2$ through the following $2^{N-|\mathcal{K}_2|} - 1 + 2^{N-|\mathcal{K}_1|} - 2^{N-|\mathcal{K}_1|-|\mathcal{K}_2|}$ constraints:

$$\begin{aligned}W \in \mathcal{L}^{\mathcal{K}_1\prec\mathcal{K}_2} \Leftrightarrow\ &\forall\ \mathcal{X}_1 \subseteq \mathcal{N}\backslash\mathcal{K}_2,\ \mathcal{X}_1 \neq \varnothing,\ \prod_{i\in\mathcal{X}_1}[1-A_O^i]A_{IO}^{\mathcal{N}\backslash\mathcal{X}_1}W = 0\\ &\text{and}\ \forall\ \mathcal{X}_2 \subseteq \mathcal{N}\backslash\mathcal{K}_1,\ \mathcal{X}_2\cap\mathcal{K}_2 \neq \varnothing,\ \prod_{i\in\mathcal{X}_2}[1-A_O^i]A_{IO}^{\mathcal{N}\backslash\mathcal{K}_1\backslash\mathcal{X}_2}W = 0.\end{aligned} \tag{A10}$$

Let us assume now that $\mathcal{K}_1$ and $\mathcal{K}_2$ cover the full set $\mathcal{N}$, so that $\mathcal{K}_1 \cup \mathcal{K}_2 = \mathcal{N}$. The characterisation above then simplifies to the following $2^{|\mathcal{K}_1|} + 2^{|\mathcal{K}_2|} - 2$ constraints:

$$\begin{aligned}&[\text{For } \mathcal{K}_1 \cup \mathcal{K}_2 = \mathcal{N}:]\\ &W \in \mathcal{L}^{\mathcal{K}_1\prec\mathcal{K}_2} \Leftrightarrow\ \forall\ \mathcal{X}_1 \subseteq \mathcal{K}_1,\ \mathcal{X}_1 \neq \varnothing,\ \prod_{i\in\mathcal{X}_1}[1-A_O^i]A_{IO}^{\mathcal{K}_1\backslash\mathcal{X}_1}A_{IO}^{\mathcal{K}_2}W = 0\\ &\qquad\text{and}\ \forall\ \mathcal{X}_2 \subseteq \mathcal{K}_2,\ \mathcal{X}_2 \neq \varnothing,\ \prod_{i\in\mathcal{X}_2}[1-A_O^i]A_{IO}^{\mathcal{K}_2\backslash\mathcal{X}_2}W = 0.\end{aligned} \tag{A11}$$

Comparing these constraints with equation (A5), one can see that the third line of equation (A11) is equivalent to imposing that the reduced process $\mathrm{Tr}_{\mathcal{K}_2} W$ is in $\mathcal{L}^{\mathcal{K}_1}$, the subspace of valid $|\mathcal{K}_1|$-partite process matrices for parties in $\mathcal{K}_1$; while the fourth line is equivalent (using the fact that $W = 0$ if and only if $\mathrm{Tr}_{\mathcal{K}_1}[M_{\mathcal{K}_1}\otimes\mathbb{1}^{\mathcal{K}_2}\cdot W] = 0$ for all $M_{\mathcal{K}_1}$) to imposing that whatever CP maps $M_{\mathcal{K}_1} \in A_{IO}^{\mathcal{K}_1}$ applied by the parties in $\mathcal{K}_1$, the conditional matrix $W_{|M_{\mathcal{K}_1}} := \mathrm{Tr}_{\mathcal{K}_1}[M_{\mathcal{K}_1}\otimes\mathbb{1}^{\mathcal{K}_2}\cdot W]$ must be in the subspace $\mathcal{L}^{\mathcal{K}_2}$ of valid process matrices for the parties in $\mathcal{K}_2$ (note that $M_{\mathcal{K}_1}$ may or may not be of a product form $\bigotimes_{k_1\in\mathcal{K}_1} M_{k_1}$ here, and that its complete positiveness is in fact irrelevant)[11]. We thus equivalently have the following characterisation:

$$[\text{For } \mathcal{K}_1 \bigcup \mathcal{K}_2 = \mathcal{N}:]\ W \in \mathcal{L}^{\mathcal{K}_1\prec\mathcal{K}_2} \Leftrightarrow\ \mathrm{Tr}_{\mathcal{K}_2} W \in \mathcal{L}^{\mathcal{K}_1}\quad\text{and}\ \forall\ M_{\mathcal{K}_1} \in A_{IO}^{\mathcal{K}_1},\ W_{|M_{\mathcal{K}_1}} \in \mathcal{L}^{\mathcal{K}_2}. \tag{A12}$$

These constraints are indeed quite intuitive: they simply correspond to the fact that for a process matrix correlation $P(\vec{a}|\vec{x})$ to be compatible with the causal order $\mathcal{K}_1 \prec \mathcal{K}_2$ (with $\mathcal{K}_1 \cup \mathcal{K}_2 = \mathcal{N}$), the probability distributions $P(\vec{a}_{\mathcal{K}_1}|\vec{x}_{\mathcal{K}_1})$ and $P(\vec{a}_{\mathcal{K}_2}|\vec{x},\ \vec{a}_{\mathcal{K}_1}) := P(\vec{a}|\vec{x})/P(\vec{a}_{\mathcal{K}_1}|\vec{x}_{\mathcal{K}_1})$ can be calculated from $\mathrm{Tr}_{\mathcal{K}_2} W$ and $W_{|M_{\mathcal{K}_1}}$, and must be well-defined.

In particular, for $\mathcal{K}_1 = \{A_k\}$ (a singleton of just one party coming first) and $\mathcal{K}_2 = \mathcal{N}\backslash A_k$, equation (A11) becomes

[10] One can easily see that the constraints from equation (A5) for which $\mathcal{X}\cap\mathcal{K}_2 \neq \varnothing$ are already implied by those of equation (A9): indeed, defining $\mathcal{X}_1 := \mathcal{X}\cap\mathcal{K}_1$ and $\mathcal{X}_2 := \mathcal{X}\cap(\mathcal{N}\backslash\mathcal{K}_1) \subseteq \mathcal{N}\backslash\mathcal{K}_1$, in such a case one has $\mathcal{X}_2\cap\mathcal{K}_2 \neq \varnothing$ and $\prod_{i\in\mathcal{X}}[1-A_O^i]A_{IO}^{\mathcal{N}\backslash\mathcal{X}}W = \prod_{i\in\mathcal{X}_1}[1-A_O^i]A_{IO}^{\mathcal{K}_1\backslash\mathcal{X}_1}(\prod_{i\in\mathcal{X}_2}[1-A_O^i]A_{IO}^{\mathcal{N}\backslash\mathcal{K}_1\backslash\mathcal{X}_2}W) = 0$ according to equation (A9). Only the constraints from equation (A5) for which $\mathcal{X}\cap\mathcal{K}_2 = \varnothing$, i.e. $\mathcal{X} \subseteq \mathcal{N}\backslash\mathcal{K}_2$ as in the second line of equation (A10) (where $\mathcal{X}$ was renamed $\mathcal{X}_1$), are nonredundant.

[11] Although the constraints in the fourth line of equation (A11) are written exactly as those that would define $\mathcal{L}^{\mathcal{K}_2}$, we emphasise that they apply here to some matrix $W \in A_{IO}^{\mathcal{N}}$, rather than to $W \in A_{IO}^{\mathcal{K}_2}$ as in the definition of $\mathcal{L}^{\mathcal{K}_2}$. This is why they must of course not directly be interpreted as implying that $W \in \mathcal{L}^{\mathcal{K}_2}$, but $W_{|M_{\mathcal{K}_1}} \in \mathcal{L}^{\mathcal{K}_2}$ for all $M_{\mathcal{K}_1}$, as in equation (A12).

$$W \in \mathcal{L}^{A_k \prec (\mathcal{N}\setminus A_k)} \Leftrightarrow {}_{[1-A_O^k]}A_{IO}^{\mathcal{N}\setminus k} W = 0 \quad \text{and} \ \forall \ \mathcal{X} \subseteq \mathcal{N}\setminus k,\ \mathcal{X} \neq 0,\ \prod_{i\in\mathcal{X}} {}_{[1-A_O^i]}A_{IO}^{\mathcal{N}\setminus k\setminus\mathcal{X}} W = 0, \tag{A13}$$

as in equation (18) of the main text. For $\mathcal{K}_1 = \mathcal{N}\setminus A_k$ and $\mathcal{K}_2 = \{A_k\}$ (a singleton of just one party coming last), equation (A11) becomes

$$W \in \mathcal{L}^{(\mathcal{N}\setminus A_k) \prec A_k} \Leftrightarrow \ \forall \ \mathcal{X} \subseteq \mathcal{N}\setminus k,\ \mathcal{X} \neq \varnothing,\ \prod_{i\in\mathcal{X}} {}_{[1-A_O^i]}A_{IO}^{\mathcal{N}\setminus k\setminus\mathcal{X}} A_{IO}^{k} W = 0 \quad \text{and} \quad {}_{[1-A_O^k]} W = 0. \tag{A14}$$

A.2.2. Causal order between several subsets of parties. Consider now K disjoint subsets $\mathcal{K}_i$ of $\mathcal{N}$. Generalising the idea of causal order between two subsets of parties, we say that the correlation $P(\vec{a}|\vec{x})$ is compatible with the causal order $\mathcal{K}_1 \prec \mathcal{K}_2 \prec \cdots \prec \mathcal{K}_K$ if and only if there is no signalling from 'future parties' to 'past parties'—i.e. if for any $k = 1, \ldots, K-1$, the outputs of parties in $\mathcal{K}_{(\leqslant k)} := \bigcup_{i=1}^{k} \mathcal{K}_i$ do not depend on the inputs of the parties in $\mathcal{K}_{(>k)} := \bigcup_{j=k+1}^{K} \mathcal{K}_j$: $P(\vec{a}_{\mathcal{K}_{(\leqslant k)}}|\vec{x}) = P(\vec{a}_{\mathcal{K}_{(\leqslant k)}}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}_{(>k)}})$ for all $\vec{x}$, $\vec{a}_{\mathcal{K}_{(\leqslant k)}}$, or equivalently, the correlation is compatible with the causal order $\mathcal{K}_{(\leqslant k)} \prec \mathcal{K}_{(>k)}$ for all $k = 1, \ldots, K-1$.

As before, we then say that a process matrix W is compatible with the causal order $\mathcal{K}_1 \prec \cdots \prec \mathcal{K}_K$ if and only if it only generates correlations that are compatible with that order. Similarly to $\mathcal{L}^{\mathcal{K}_1 \prec \mathcal{K}_2}$, we define the subspace

$$\mathcal{L}^{\mathcal{K}_1 \prec \cdots \prec \mathcal{K}_K} := \bigcap_{k=1}^{K-1} \mathcal{L}^{\mathcal{K}_{(\leqslant k)} \prec \mathcal{K}_{(>k)}} \tag{A15}$$

of (valid) process matrices compatible with the causal order $\mathcal{K}_1 \prec \cdots \prec \mathcal{K}_K$.

In the case where the subsets $\mathcal{K}_i$ define a full partition of $\mathcal{N}$ (i.e. where $\bigcup_{i=1}^{K} \mathcal{K}_i = \mathcal{N}$), we easily obtain, from equations (A11) and (A12) (after removing redundant constraints as in footnote 10), that

$$\begin{aligned} &[\text{For } \bigcup_{i=1}^{K} K_i = N:]\ W \in L^{K_1 \prec \cdots \prec K_K} \iff \forall\ k = 1, \ldots, K, \forall\ M_{K_{(<k)}} \in A_{IO}^{K_{(<k)}},\ \mathrm{Tr}_{K_{(>k)}} W_{|M_{K_{(<k)}}} \in L^{K_k} \\ &\iff \forall\ k = 1, \ldots, K, \forall\ X_k \subseteq K_k, X_k \neq \varnothing,\ \prod_{i\in X_k} [1-A_O^i] A_{IO}^{K_k/X_k} A_{IO}^{K_{(>k)}} W = 0 \end{aligned} \tag{A16}$$

with $\mathcal{K}_{(<k)} := \bigcup_{i=1}^{k-1} \mathcal{K}_i$ for $k = 2, \ldots, K$, $\mathcal{K}_{(<1)} = \mathcal{K}_{(>K)} = \varnothing$, $W_{|M_{\mathcal{K}_{(<1)}}} = W$.

In particular, in the case where all subsets $\mathcal{K}_k$ are singletons—i.e. $\mathcal{K}_k = \{A_{\pi(k)}\}$ for some permutation π of $\mathcal{N}$—we find that the subspace $\mathcal{L}^{A_{\pi(1)} \prec \cdots \prec A_{\pi(N)}}$ of process matrices compatible with the causal order $A_{\pi(1)} \prec A_{\pi(2)} \prec \cdots \prec A_{\pi(N)}$ is characterised by [14, 30, 31]

$$W \in \mathcal{L}^{A_{\pi(1)} \prec \cdots \prec A_{\pi(N)}} \Leftrightarrow \ \forall\ k = 1, \ldots, N,\ {}_{[1-A_O^{\pi(k)}]} A_{IO}^{\pi(>k)} W = 0 \tag{A17}$$

with $\pi(>k) = \{\pi(k+1), \ldots, \pi(N)\}$ for $k = 1, \ldots, N-1$, $\pi(>N) = \varnothing$ (as in equation (19) when π is the identity permutation).

A.2.3. Particular cases with $d_{A_I^f} = 1$ or $d_{A_O^\ell} = 1$. Suppose there exists a party A_f which has a trivial incoming space, i.e. such that $d_{A_I^f} = 1$. The constraints of equation (A5) can be written, depending on whether $A_f \in \mathcal{X}$ or $A_f \notin \mathcal{X}$ and renaming $\mathcal{X}\setminus A_f \to \mathcal{X}$ in the former case, in the forms ${}_{[1-A_O^f]} \prod_{i\in\mathcal{X}} {}_{[1-A_O^i]} A_{IO}^{\mathcal{N}\setminus A_f\setminus\mathcal{X}} W = 0$ for all $\mathcal{X} \subseteq \mathcal{N}\setminus A_f$ and ${}_{A_O^{A_f}} \prod_{i\in\mathcal{X}} {}_{[1-A_O^i]} A_{IO}^{\mathcal{N}\setminus A_f\setminus\mathcal{X}} W = 0$ for all $\mathcal{X} \subseteq \mathcal{N}\setminus A_f$, $\mathcal{X} \neq \varnothing$, respectively. Summing up these two constraints in the case where $\mathcal{X} \neq \varnothing$ (and keeping the first one for the case where $\mathcal{X} = \varnothing$), we find that $\mathcal{L}^{\mathcal{N}}$ is characterised by the same constraints as those characterising $\mathcal{L}^{A_f \prec (\mathcal{N}\setminus A_f)}$ in equation (A13), namely

$$\begin{aligned} &[\text{For } d_{A_I^f} = 1:]\ W \in \mathcal{L}^{\mathcal{N}} \Leftrightarrow W \in \mathcal{L}^{A_f \prec (\mathcal{N}\setminus A_f)} \\ &\Leftrightarrow {}_{[1-A_O^f]} A_{IO}^{\mathcal{N}\setminus A_f} W = 0 \ \text{ and } \forall\ \mathcal{X} \subseteq \mathcal{N}\setminus A_f,\ \mathcal{X} \neq \varnothing,\ \prod_{i\in\mathcal{X}} [1-A_O^i] A_{IO}^{\mathcal{N}\setminus A_f\setminus\mathcal{X}} W = 0. \end{aligned} \tag{A18}$$

Hence, in that case any valid process matrix is compatible with party A_f acting first. This corresponds indeed to the natural intuition that, because party A_f does not receive any physical system from anyone, they do not need to wait for any other party to act before them.

In the case where several parties in $\mathcal{A}_f := \{A_{f_1}, \ldots, A_{f_n}\}$ have trivial incoming spaces (such that $d_{A_I^{f_j}} = 1$ for all $j = 1, \ldots, n$), equation (A18) easily generalises to[12]

[12] This implies in particular that for an extension of the outgoing systems only, $W \otimes W^{\text{ext.}}$ with $W^{\text{ext.}} \in A_{O'}^{\mathcal{N}}$ is valid only if $W^{\text{ext.}}$ is proportional to the identity operator: see footnote 12.

$$[\text{For } d_{A_I^{f_j}} = 1,\ \forall\ A_{f_j} \in \mathcal{A}_f\text{:}]\ W \in \mathcal{L}^{\mathcal{N}} \Leftrightarrow W \in \mathcal{L}^{A_{f_j} \prec (\mathcal{N}\setminus A_{f_j})}\ \forall\ j$$
$$\Leftrightarrow {}_{[1-A_O^{\mathcal{A}_f}]}A_{IO}^{\mathcal{N}\setminus\mathcal{A}_f} W = 0 \quad \text{and } \forall\ \mathcal{X} \subseteq \mathcal{N}\setminus\mathcal{A}_f,\ \mathcal{X} \neq \varnothing,\ \prod_{i\in\mathcal{X}} {}_{[1-A_O^i]}A_{IO}^{\mathcal{N}\setminus\mathcal{A}_f\setminus\mathcal{X}} W = 0. \tag{A19}$$

Instead of $d_{A_I^f} = 1$, suppose now that there exists a party A_ℓ which has a trivial outgoing space, i.e. such that $d_{A_O^\ell} = 1$, and consider the causal order $(\mathcal{N}\setminus A_\ell) \prec A_\ell$. Note that in this case the map $W \to {}_{A_O^\ell} W$ reduces to the identity, so that any constraint of the form ${}_{[1-A_O^\ell]\ldots} W = 0$ is trivially satisfied. The nontrivial constraints from equations (A5) and (A14) then reduce to the same set of constraints, namely,

$$[\text{For } d_{A_O^\ell} = 1\text{:}]\ W \in \mathcal{L}^{\mathcal{N}} \Leftrightarrow W \in \mathcal{L}^{(\mathcal{N}\setminus A_\ell) \prec A_\ell}$$
$$\Leftrightarrow \forall\ \mathcal{X} \subseteq \mathcal{N}\setminus A_\ell,\ \mathcal{X} \neq \varnothing,\ \prod_{i\in\mathcal{X}} {}_{[1-A_O^i]}A_{IO}^{\mathcal{N}\setminus A_\ell\setminus\mathcal{X}} A_I^\ell W = 0. \tag{A20}$$

Hence, similarly to the previous case, here any valid process matrix is compatible with party A_ℓ acting last. This is again rather intuitive: as party A_ℓ sends no physical system out and cannot signal to anyone, then they can always come last—see, e.g. the motivation for only considering fixed orders with party C coming last in Araújo *et al*'s definition of causal separability [14].

If now several parties in $\mathcal{A}_\ell \coloneqq \{A_{\ell_1}, \ldots, A_{\ell_n}\}$ have trivial outgoing spaces (such that $d_{A_O^{\ell_j}} = 1$ for all $j = 1, \ldots, n$), then one can easily check that any process matrix is compatible with all those parties acting last, with any causal order among them: for any permutation π of $\{1, \ldots, n\}$,

$$[\text{For } d_{A_O^{\ell_j}} = 1,\ \forall\ A_{\ell_j} \in \mathcal{A}_\ell\text{:}]\ W \in \mathcal{L}^{\mathcal{N}} \Leftrightarrow W \in \mathcal{L}^{(\mathcal{N}\setminus\mathcal{A}_\ell) \prec A_{\ell_{\pi(1)}} \prec \cdots \prec A_{\ell_{\pi(n)}}}$$
$$\Leftrightarrow \forall\ \mathcal{X} \subseteq \mathcal{N}\setminus\mathcal{A}_\ell,\ \mathcal{X} \neq \varnothing,\ \prod_{i\in\mathcal{X}} {}_{[1-A_O^i]}A_{IO}^{\mathcal{N}\setminus\mathcal{A}_\ell\setminus\mathcal{X}} A_I^{\mathcal{A}_\ell} W = 0. \tag{A21}$$

It is worth emphasising that no similar property holds for several parties in $\mathcal{A}_f = \{A_{f_1}, \ldots, A_{f_n}\}$ having trivial incoming spaces, as considered previously: any process matrix is compatible in that case with any causal order $A_{f_j} \prec (\mathcal{N}\setminus A_{f_j})$ (as in equation (A19)), but not necessarily with $A_{f_1} \prec \cdots \prec A_{f_n} \prec (\mathcal{N}\setminus\mathcal{A}_f)$ (or with any other permutation of the first parties).[13]

To finish here, note, furthermore, that if a party A_k has both $d_{A_I^k} = d_{A_O^k} = 1$, then clearly one can just ignore it: in such a case, $W \in \mathcal{L}^{\mathcal{N}} \Leftrightarrow W \in \mathcal{L}^{\mathcal{N}\setminus A_k}$.

A.2.4. Comment on our use of the notation $\prec$. Let us comment briefly here on our use of the notation $\prec$. Recall that for two disjoint nonempty subsets $\mathcal{K}_1$ and $\mathcal{K}_2$ of $\mathcal{N}$, a probability distribution P is said to be compatible with the causal order $\mathcal{K}_1 \prec \mathcal{K}_2$ if and only if $P(\vec{a}_{\mathcal{K}_1}|\vec{x}) = P(\vec{a}_{\mathcal{K}_1}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}_2})$ for all $\vec{x}$, $\vec{a}_{\mathcal{K}_1}$. It should be noted that the relation 'compatible with $\mathcal{K}_1 \prec \mathcal{K}_2$' thus defined is not transitive, and therefore it does not define a partial order between events. For instance, $P(a, b, c|x, y, z) \coloneqq \delta_{a,z}\ \delta_{b,0}\ \delta_{c,0}$ (with δ the Kronecker delta and a, z taking at least two different values) is compatible with $A \prec B$ and $B \prec C$, but not with $A \prec C$. This justifies why, considering more subsets, we defined the notation $\mathcal{K}_1 \prec \mathcal{K}_2 \prec \cdots \prec \mathcal{K}_K$ to formally mean $\mathcal{K}_{(\leqslant k)} \prec \mathcal{K}_{(>k)}$—rather than just $\mathcal{K}_k \prec \mathcal{K}_{k+1}$—for all $k = 1, \ldots, K-1$.

We note also that the notation $\prec$ was used differently in [17], where it denoted a strict partial order (and was hence transitive). Our use of the notation $\prec$ here is consistent e.g. with that of [14, 15, 18, 22, 27, 34, 35, 43, 44], and would instead correspond to the notation $\not\succcurlyeq$ in [17] (also used in [2]).

A.3. Operations on process matrices

In this section we clarify how process matrices behave in general, with respect to their validity and their compatibility with fixed causal orders, when tracing out subsystems or attaching extensions, and when tracing out, adding or grouping parties.

A.3.1. Tracing out subsystems/Attaching extensions. Suppose that the incoming and outgoing spaces of N parties can be decomposed as $A_{IO}^{\mathcal{N}} \otimes A_{I'O'}^{\mathcal{N}}$ (possibly with some trivial spaces A_I^k, A_O^k, $A_{I'}^k$ or $A_{O'}^k$), and consider a given matrix $W' \in A_{II'OO'}^{\mathcal{N}}$. If W' is a valid process matrix, then so is $W \coloneqq \mathrm{Tr}_{A_{I'O'}^{\mathcal{N}}} W'$; similarly, if W' is compatible with a causal order $\mathcal{K}_1 \prec \mathcal{K}_2$, then so is W. Both properties are quite intuitive:[14] clearly, ignoring some parts of

[13] We use here, in particular, the fact that ${}_{[1-A_O^{f_j}]}A_{IO}^{\mathcal{N}\setminus\mathcal{A}_f} W = 0\ \forall\ j$ is equivalent to ${}_{[1-A_O^{\mathcal{A}_f}]}A_{IO}^{\mathcal{N}\setminus\mathcal{A}_f} W = 0$. Note that if all incoming spaces are trivial, i.e. $\mathcal{A}_f = \mathcal{N}$, then ${}_{[1-A_O^{\mathcal{N}}]} W = 0$ implies that the only valid process matrices are those proportional to the identity operator $\mathbb{1}^{A_O^{\mathcal{N}}}$.

[14] Note indeed, in a similar fashion, that while compatibility of a probability distribution P with the orders $(\mathcal{K}_1 \cup \mathcal{K}_2) \prec \mathcal{K}_3$ and $(\mathcal{K}_1 \cup \mathcal{K}_3) \prec \mathcal{K}_2$ implies compatibility with $\mathcal{K}_1 \prec (\mathcal{K}_2 \cup \mathcal{K}_3)$, and therefore with both $\mathcal{K}_1 \prec \mathcal{K}_2 \prec \mathcal{K}_3$ and $\mathcal{K}_1 \prec \mathcal{K}_3 \prec \mathcal{K}_2$, it is not the case that compatibility with $\mathcal{K}_1 \prec (\mathcal{K}_2 \cup \mathcal{K}_3)$ and $\mathcal{K}_2 \prec (\mathcal{K}_1 \cup \mathcal{K}_3)$ necessarily implies $(\mathcal{K}_1 \cup \mathcal{K}_2) \prec \mathcal{K}_3$, and it therefore does also not necessarily imply compatibility with $\mathcal{K}_1 \prec \mathcal{K}_2 \prec \mathcal{K}_3$ or $\mathcal{K}_2 \prec \mathcal{K}_1 \prec \mathcal{K}_3$. As a counter-example, one can see for instance that $P(a, b, c|x, y, z) = \frac{1}{2}\delta_{a\oplus b,z}\ \delta_{c,0}$ (with binary inputs and outputs, where δ the Kronecker delta and $\oplus$denotes addition modulo 2) is compatible with both $A \prec \{B, C\}$ and $B \prec \{A, C\}$, but not with $\{A, B\} \prec C$.

the incoming and outgoing spaces cannot make a process matrix invalid, and cannot induce some signalling where there was none before. Note, however, that the converse is in general not true: if $W = \mathrm{Tr}_{A_{I'O'}^{\mathcal{N}}} W'$ is a valid process matrix (or is compatible with $\mathcal{K}_1 \prec \mathcal{K}_2$), this does not guarantee in general that W' is also a valid process matrix (or is compatible with $\mathcal{K}_1 \prec \mathcal{K}_2$).

There is nevertheless a case, where the validity of a process matrix W ensures that a 'larger' matrix W' (defined on more subsystems) is valid: namely, when one attaches to W some ancillary state ρ. Indeed, in constructing the framework of process matrices, it is always assumed that one can consider some extensions of the incoming spaces of each party, and distribute (possibly entangled) ancillary states shared by all parties. Hence, by definition, if a matrix $W \in A_{IO}^{\mathcal{N}}$ is a valid process matrix, then for any quantum state (i.e. any positive semidefinite matrix, up to normalisation) ρ in any extension $A_{I'}^{\mathcal{N}}$, the matrix $W' = W \otimes \rho \in A_{II'O}^{\mathcal{N}}$ defines a valid process matrix. Similarly, if W is compatible with a given causal order $\mathcal{K}_1 \prec \mathcal{K}_2$ between two disjoint subsets of parties, then so is $W \otimes \rho$.

One may then wonder if instead of attaching an ancillary state $\rho \in A_{I'}^{\mathcal{N}}$ to the incoming spaces, one could attach any other positive semidefinite matrix $W^{\text{ext.}} \in A_{I'O'}^{\mathcal{N}}$ in some extension of both incoming and outgoing spaces. It is clear, from the previous remarks on the partial trace of subsystems, that for a valid (nonzero) process matrix W, a necessary condition for $W' \coloneqq W \otimes W^{\text{ext.}}$ to define a valid process matrix is that $W^{\text{ext.}}$ itself is also a valid process matrix[15]. For two parties and more, this condition is however not sufficient (as noted also in [41, 42]): for instance, $W = \frac{1}{2}(\mathbb{1}^{A_O B_I} + \hat{z}^{A_O}\hat{z}^{B_I})$ and $W^{\text{ext.}} = \frac{1}{2}(\mathbb{1}^{A_{I'}B_{O'}} + \hat{z}^{A_{I'}}\hat{z}^{B_{O'}})$ are both valid, but $W \otimes W^{\text{ext.}}$ is not (due to the fact that W and $W^{\text{ext.}}$ allow for some signalling in two conflicting directions). Similarly, for a (nonzero) process matrix W compatible with $\mathcal{K}_1 \prec \mathcal{K}_2$, a necessary condition for $W' \coloneqq W \otimes W^{\text{ext.}}$ to be a process matrix compatible with $\mathcal{K}_1 \prec \mathcal{K}_2$ is that $W^{\text{ext.}}$ is also a process matrix compatible with $\mathcal{K}_1 \prec \mathcal{K}_2$. As before, this condition is however not sufficient for three parties and more.

A.3.2. Tracing out/adding/separating/grouping parties. In the previous observations we were keeping the set of parties under consideration $\mathcal{N}$ fixed. Let us now consider how process matrices behave when changing the set of parties.

Consider a nonempty subset $\mathcal{N}_0$ of $\mathcal{N}$. Clearly, if $W \in A_{IO}^{\mathcal{N}}$ is a valid N-partite process matrix, then its restriction to the parties in the subset $\mathcal{N}_0$, defined as $W_0 \coloneqq \mathrm{Tr}_{\mathcal{N}\setminus\mathcal{N}_0} W$, is a valid $|\mathcal{N}_0|$-partite process matrix. Similarly, if W is compatible with a causal order $\mathcal{K}_1 \prec \mathcal{K}_2$, then W_0 is compatible with the order $(\mathcal{K}_1 \cap \mathcal{N}_0) \prec (\mathcal{K}_2 \cap \mathcal{N}_0)$.[16]

Let $\mathcal{N}_1$ and $\mathcal{N}_2$ be two disjoint sets of parties. If $W_1 \in A_{IO}^{\mathcal{N}_1}$ and $W_2 \in A_{IO}^{\mathcal{N}_2}$ are two valid process matrices for the parties in $\mathcal{N}_1$ and $\mathcal{N}_2$, respectively, then so is $W = W_1 \otimes W_2 \in A_{IO}^{\mathcal{N}_1 \cup \mathcal{N}_2}$ for all parties in $\mathcal{N}_1 \cup \mathcal{N}_2$. (Note however that if $\mathcal{N}_1$ and $\mathcal{N}_2$ are not disjoint, this may not hold any more, as in the case with $\mathcal{N}_1 = \mathcal{N}_2 = \mathcal{N}$ considered in the previous subsection.) If say W_1 is compatible with a causal order $\mathcal{K}_1 \prec \mathcal{K}_1'$ (with $\mathcal{K}_1$, $\mathcal{K}_1'$ two disjoint nonempty subsets of $\mathcal{N}_1$), then so is W. Furthermore, for any nonempty subsets $\mathcal{K}_1 \subseteq \mathcal{N}_1$ and $\mathcal{K}_2 \subseteq \mathcal{N}_2$, W is compatible with both orders $\mathcal{K}_1 \prec \mathcal{K}_2$ and $\mathcal{K}_2 \prec \mathcal{K}_1$.

Suppose now that the incoming and outgoing spaces of a party, say A_N, can be factorised into $A_{IO}^{N} = A_{IO}^{N^{(1)}} \otimes A_{IO}^{N^{(2)}}$. One can then virtually 'separate' A_N into two new parties, $A_{N^{(1)}}$ and $A_{N^{(2)}}$,with incoming and outgoing spaces $A_{IO}^{N^{(1)}}$ and $A_{IO}^{N^{(2)}}$, respectively, and thus consider the new set of $N+1$ parties $\mathcal{N}' = \{A_1, \ldots, A_{N-1}, A_{N^{(1)}}, A_{N^{(2)}}\}$. If $W \in A_{IO}^{\mathcal{N}}$ is a valid N-partite process matrix,one can then verify that when considering the set $\mathcal{N}'$, $W \in A_{IO}^{\mathcal{N}'}$ is also is a valid $(N+1)$-partite process matrix, i.e. $W \in \mathcal{L}^{\mathcal{N}'}$. If W is compatible with a causal order $\mathcal{K}_1 \prec \mathcal{K}_2$, then W is also compatible with $\mathcal{K}_1' \prec \mathcal{K}_2'$, where $\mathcal{K}_i'$ is obtained from $\mathcal{K}_i$ (like $\mathcal{N}'$ from $\mathcal{N}$) by replacing A_N by $A_{N^{(1)}}$, $A_{N^{(2)}}$ (when $A_N \in K_i$).[17]

Conversely, for a given set of $N \geqslant 2$ parties $\mathcal{N}$, let us finally consider a set $\mathcal{N}'$ obtained from $\mathcal{N}$ by now grouping two or more of the parties, e.g. $\mathcal{N}' = \{A_1, \ldots, A_{N-2}, \{A_{N-1}, A_N\}\}$, where $\{A_{N-1}, A_N\}$ is considered to form a new effective party. Then W is not necessarily a valid $(N-1)$-partite process matrix for the parties in $\mathcal{N}'$. The reason for this is that valid process matrices are required to give valid probability distributions only for product operations of the parties; if two parties are grouped together and perform a joint operation, that may no longer yield valid probabilities. An explicit counterexample is for instance $W = \frac{1}{2}(\mathbb{1}^{A_O B_I} + \hat{z}^{A_O}\hat{z}^{B_I})$, which

[15] They can be verified straightforwardly using for instance equations (A5) and (A10), respectively, by writing (in the first case) $\Pi_{i\in\mathcal{X}}[\mathbb{1}-A_O^i]A_{IO}^{\mathcal{N}\setminus\mathcal{X}} W = \mathrm{Tr}_{A_{I'O'}^{\mathcal{N}}}[\Pi_{i\in\mathcal{X}}[\mathbb{1}-A_O^i]A_{IO}^{\mathcal{N}\setminus\mathcal{X}} W'] = \mathrm{Tr}_{A_{I'O'}^{\mathcal{N}}}[\Pi_{i\in\mathcal{X}}[\mathbb{1}-A_{OO'}^i]A_{II'OO'}^{\mathcal{N}\setminus\mathcal{X}} W'] = 0$ (and by noting that $W' \geqslant 0$ implies $W \geqslant 0$).

[16] Again, both properties can easily be checked by using for instance equations (A5) and (A10), and the fact that $W \geqslant 0$ implies $W_0 \geqslant 0$.

[17] Both properties are straightforward when recalling that the validity and compatibility with a fixed order constraints are obtained by imposing certain conditions for all operations $M \in A_{IO}^{N}$ of party A_N: clearly, these constraints are also satisfied if A_N is separated into two parties $A_{N^{(1)}}$ and $A_{N^{(2)}}$, and M takes the form $M = M^{(1)} \otimes M^{(2)} \in A_{IO}^{N^{(1)}} \otimes A_{IO}^{N^{(2)}}$ (and by noting that if $M^{(1)}$ and $M^{(2)}$ are CPTP, then so is M). These properties can also be verified formally using the characterisations of equations (A5) or (A10), after noting in particular that $[\mathbb{1}-A_O^{N^{(1)}}A_O^{N^{(2)}}]\ldots W = 0$ is equivalent to $[\mathbb{1}-A_O^{N^{(1)}}]\ldots W = [\mathbb{1}-A_O^{N^{(2)}}]\ldots W = 0$.

represents a (dephasing) channel from A to B and is indeed a valid bipartite process matrix, but not a valid single-partite process if A and B are grouped together (as ${}_{[1-A_O]}W \neq 0$; e.g. the joint CPTP map $M = \frac{1}{2}(\mathbb{1}^{A_O B_I} - \hat{z}^{A_O}\hat{z}^{B_I})$ gives $\mathrm{Tr}[M \cdot W] = 0$).[18]

A.4. Allowed and forbidden terms in the Hilbert–Schmidt basis decomposition of process matrices

In [2, 17], the constraints characterising the set of valid process matrices and the set of process matrices compatible with a fixed causal order between two complementary subsets were formulated in a different way, namely by specifying which terms can appear in the decomposition of the corresponding operators in a Hilbert–Schmidt basis. To complete this appendix, we now establish the connection between these two alternative characterisations, and we prove their equivalence.

A Hilbert–Schmidt basis of some space of linear operators X (acting on a d_X-dimensional Hilbert space) is given by a set of generalised Pauli matrices, i.e. a set of Hermitian operators $\{\sigma_\mu^X\}_{\mu=0}^{d_X^2-1}$, with $\sigma_0^X = \mathbb{1}^X$, $\mathrm{Tr}[\sigma_\mu^X \sigma_\nu^X] = d_X \delta_{\mu,\nu}$ for all $\mu, \nu = 0, \ldots, d_X^2 - 1$, and $\mathrm{Tr}[\sigma_\mu^X] = 0$ for $\mu \geqslant 1$. In such a basis, a process matrix $W \in A_{IO}^1 \otimes A_{IO}^2 \otimes A_{IO}^3 \otimes \cdots$ can be expanded as

$$W = \sum_{\mu_1,\nu_1,\mu_2,\nu_2,\ldots} w_{\mu_1\nu_1\mu_2\nu_2\mu_3\nu_3\ldots}\, \sigma_{\mu_1}^{A_I^1}\sigma_{\nu_1}^{A_O^1}\sigma_{\mu_2}^{A_I^2}\sigma_{\nu_2}^{A_O^2}\sigma_{\mu_3}^{A_I^3}\sigma_{\nu_3}^{A_O^3}\cdots$$
$$\text{with } w_{\mu_1\nu_1\mu_2\nu_2\mu_3\nu_3\ldots} \in \mathbb{R} \quad \forall\ \mu_1, \nu_1, \mu_2, \nu_2, \mu_3, \nu_3, \ldots. \tag{A22}$$

The approach of [2, 17] looks at which terms $\sigma_{\mu_1}^{A_I^1}\sigma_{\nu_1}^{A_O^1}\sigma_{\mu_2}^{A_I^2}\sigma_{\nu_2}^{A_O^2}\sigma_{\mu_3}^{A_I^3}\sigma_{\nu_3}^{A_O^3}\cdots$ can appear with a nonzero coefficient $w_{\mu_1\nu_1\mu_2\nu_2\mu_3\nu_3\ldots}$ (i.e. are 'allowed') in the above decomposition. According to Proposition 3.1 of [17], a Hermitian operator W is in the linear subspace $\mathcal{L}^{\mathcal{N}}$ of valid process matrices if and only if, in addition to the identity term $\mathbb{1}^{\mathcal{N}}$, it contains only terms for which at least one party A_k has a nontrivial operator $\sigma_{\mu_k} \neq \mathbb{1}$ on A_I^k and the identity operator $\mathbb{1}$ on A_O^k.[19]

To see that this statement is indeed equivalent to our own characterisation of $\mathcal{L}^{\mathcal{N}}$, let us first verify that all terms of this kind fulfil all the constraints of equation (A5). This is clearly the case for the identity $\mathbb{1}^{\mathcal{N}}$, since ${}_{[1-A_O^i]}\mathbb{1}^{\mathcal{N}} = 0$ for any party A_i. Consider then some generic Hilbert–Schmidt term T_k of the form $\cdots\sigma_{\mu_k}^{A_I^k}\mathbb{1}^{A_O^k}\cdots$ (with $\sigma_{\mu_k} \neq \mathbb{1}$). Such a term satisfies ${}_{[1-A_O^k]}T_k = {}_{A_{IO}^k}T_k = 0$, so that for any $\mathcal{X} \subseteq \mathcal{N}$, $\mathcal{X} \neq \varnothing$ we have $\prod_{i\in\mathcal{X}}[1-A_O^i]A_{IO}^{\mathcal{N}\setminus\mathcal{X}}T_k = 0$, whether $k \in \mathcal{X}$ or $k \in \mathcal{N}\setminus\mathcal{X}$. By linearity, any operator W whose Hilbert–Schmidt decomposition (A22) only contains the identity or such terms T_k thus satisfies all the constraints (A5). Conversely, suppose that the Hilbert–Schmidt decomposition of W contains a term F (with a nonzero weight) that is 'forbidden' according to Proposition 3.1 of [17], that is, a term such that for all parties A_k, there is either a nontrivial operator $\sigma_{\nu_k} \neq \mathbb{1}$ on A_O^k, or an identity operator on both A_I^k and A_O^k (and where there is at least one party for which the former is true). Consider then the nonempty subset $\mathcal{X} \subseteq \mathcal{N}$ of parties A_i for which $\sigma_{\nu_i}^{A_O^i} \neq \mathbb{1}^{A_O^i}$ in F. For $i \in \mathcal{X}$, one thus has ${}_{[1-A_O^i]}F = F$, while for $j \in \mathcal{N}\setminus\mathcal{X}$, ${}_{A_{IO}^j}F = F$; all in all, $\prod_{i\in\mathcal{X}}[1-A_O^i]A_{IO}^{\mathcal{N}\setminus\mathcal{X}}F = F$. By the linear independence of all Hilbert–Schmidt terms, W then cannot satisfy $\prod_{i\in\mathcal{X}}[1-A_O^i]A_{IO}^{\mathcal{N}\setminus\mathcal{X}}W = 0$, and thus violates the constraints of equation (A5).

The process matrices that are compatible with the causal order $\mathcal{K}_1 \prec \mathcal{K}_2$, with $\mathcal{K}_1 \cup \mathcal{K}_2 = \mathcal{N}$, were likewise characterised in [17] in terms of allowed terms in a Hilbert–Schmidt basis decomposition. The following terminology was used: the *restriction* of a Hilbert–Schmidt term onto certain subsystems is the part of the term corresponding to the respective subsystems—for example, the restriction of the term $\sigma_{\mu_1}^{A_I^1}\sigma_{\nu_1}^{A_O^1}\sigma_{\mu_2}^{A_I^2}\sigma_{\nu_2}^{A_O^2}\sigma_{\mu_3}^{A_I^3}\sigma_{\nu_3}^{A_O^3}\cdots$ onto the subsystems $A_I^2 \otimes A_O^2$ is just $\sigma_{\mu_2}^{A_I^2}\sigma_{\nu_2}^{A_O^2}$. According to Proposition 3.2 in [17], the (valid) process matrices that do not allow signalling from $\mathcal{K}_2$ to $\mathcal{K}_1$ are those that contain only Hilbert–Schmidt terms whose restriction to the (incoming and outgoing systems of) parties in $\mathcal{K}_2$ are of the allowed type for a $|\mathcal{K}_2|$-partite process matrix for those parties—that is, terms with either the identity operator $\mathbb{1}^{A_I^{k_2}}\mathbb{1}^{A_O^{k_2}}$ for all parties in $\mathcal{K}_2$, or for which there is some party $A_{k_2} \in \mathcal{K}_2$ with a nontrivial generalised Pauli operator $\sigma_{\mu_{k_2}} \neq \mathbb{1}$ on $A_I^{k_2}$ and the identity operator on $A_O^{k_2}$.

To see that this proposition is indeed equivalent to our characterisation of $\mathcal{L}^{\mathcal{K}_1\prec\mathcal{K}_2}$ given by equation (A11) or (A12), note that the restriction of a Hilbert–Schmidt term T to $\mathcal{K}_2$ is precisely obtained, up to a multiplicative

[18] Nevertheless, from equations (A19) and (A21) one can see that parties who all have trivial incoming spaces, or parties who all have trivial outgoing spaces, can be grouped together without changing the validity of the process matrix in question.

[19] As clarified in [45], valid process matrix can indeed only contain terms that, except for the identity, do not appear in the Hilbert–Schmidt decomposition of $\bigotimes_{k\in\mathcal{N}} M_{x_k}^{A_{IO}^k}$, for any CPTP maps $M_{x_k}^{A_{IO}^k}$ (as otherwise it is always possible to find some CPTP maps $M_{x_k}^{A_{IO}^k}$ that give nonnormalised probabilities via the generalised Born rule of equation (4)). Given the constraint $\mathrm{Tr}_{A_O^k} M_{x_k}^{A_{IO}^k} = \mathbb{1}^{A_I^k}$ for CPTP maps, one can see that the only forbidden terms in any $M_{x_k}^{A_{IO}^k}$ are of the form $\sigma_{\mu_k}^{A_I^k}\mathbb{1}^{A_O^k}$ with $\sigma_{\mu_k} \neq \mathbb{1}$. Thus, only terms that contain $\sigma_{\mu_k}^{A_I^k}\mathbb{1}^{A_O^k}$ for at least one value of k cannot appear in the Hilbert–Schmidt decomposition of $\bigotimes_{k\in\mathcal{N}} M_{x_k}^{A_{IO}^k}$, and are thus allowed (in addition to the identity) in the decomposition of a process matrix.

factor (which may be 0), by taking the partial trace $T_{|M_{\mathcal{K}_1}} := \mathrm{Tr}_{\mathcal{K}_1}[M_{\mathcal{K}_1} \otimes \mathbb{1}^{\mathcal{K}_2} \cdot T]$, for any $M_{\mathcal{K}_1} \in A_{IO}^{\mathcal{K}_1}$. Hence, imposing that all Hilbert–Schmidt terms T in the decomposition of $W = \sum_T w_T\, T$ have restrictions to $\mathcal{K}_2$ that are in $\mathcal{L}^{\mathcal{K}_2}$, as in the characterisation of [17] just recalled, is equivalent to imposing that for any $M_{\mathcal{K}_1} \in A_{IO}^{\mathcal{K}_1}$, $W_{|M_{\mathcal{K}_1}} = \sum_T w_T\, T_{|M_{\mathcal{K}_1}}$ only have terms $T_{|M_{\mathcal{K}_1}} \in \mathcal{L}^{\mathcal{K}_2}$, i.e. that $W_{|M_{\mathcal{K}_1}}$ itself is in $\mathcal{L}^{\mathcal{K}_2}$, as imposed in equation (A12). Note that Proposition 3.2 in [17] pre-supposed that the process matrix under consideration was valid. If this is not pre-supposed, one must in addition impose, according to the previous characterisation, that for Hilbert–Schmidt terms in the decomposition of W whose restriction to $\mathcal{K}_2$ is the identity operator $\mathbb{1}^{\mathcal{K}_2}$, there must either also be the identity operator for all parties in $\mathcal{K}_1$, or there must be some party $A_{k_1} \in \mathcal{K}_1$ with a restriction $\sigma_{\mu_{k_1}}^{A_I^{k_1}} \mathbb{1}^{A_O^{k_1}} \neq \mathbb{1}^{A_I^{k_1}} \mathbb{1}^{A_O^{k_1}}$—in other words, one must impose that $\mathrm{Tr}_{\mathcal{K}_2} W \in \mathcal{L}^{\mathcal{K}_1}$, so that one also recovers the first constraint of equation (A12).

Appendix B. Characterisation of causally separable process matrices

In this appendix we prove the propositions that allow us to characterise causally separable process matrices in terms of simple conditions. We start by proving the first part of proposition 1, namely that in the particular tripartite scenario with $d_{C_O} = 1$, Araújo *et al*'s definition of causal separability (Definition 2) is equivalent to OG's notion of extensible causal separability (Definition 4), and thus also to our definition 5. Then we provide the proofs for the characterisation of general tripartite causally separable process matrices (Proposition 3) as well as for the necessary condition (Proposition 4) and the sufficient condition (Proposition 5) in the general N-partite case. Note that all the special cases follow from propositions 4 and 5, and we could just give the proofs of those two general propositions. However, for pedagogical reasons we start with the simpler proofs, which may entail some repetition in the arguments, but allows for greater clarity in presenting the core ideas.

All of the proofs below (of increasing complexity) make use of the same type of argument to prove the necessity of the respective conditions. This argument is based on the 'teleportation technique' that follows from the lemma below. Before stating it, let us introduce some further notation. For two Hilbert spaces $\mathcal{H}^X$, $\mathcal{H}^{X'}$ with the same dimension d, and denoting by $\{|i\rangle^{X^{(\prime)}}\}_{i=1}^{d}$ an orthonormal basis of either $\mathcal{H}^X$ or $\mathcal{H}^{X'}$, we will consider the maximally entangled state $|\Phi^+\rangle^{X/X'} := \frac{1}{\sqrt{d}} \sum_i |i\rangle^X \otimes |i\rangle^{X'}$. We also recall that for a given matrix $W \in A_{IO}^{\mathcal{N}}$, we denote by $W^{A_{IO}^k \to A_{I'}^{k'}}$ the matrix in $(\bigotimes_{j \in \mathcal{N}\setminus k} A_{IO}^j) \otimes A_{I'}^{k'}$ that has formally the same form as W, except that party A_k's system A_{IO}^k is now attributed to an extension $A_{I'}^{k'}$ of party $A_{k'}$'s incoming space. Formally (recalling equation (25)),

$$W^{A_{IO}^k \to A_{I'}^{k'}} := \sum_{i,j} \mathrm{Tr}_k[|i\rangle\langle j|^{A_{IO}^k} \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W] \otimes |j\rangle\langle i|^{A_{I'}^{k'}}, \tag{B1}$$

where $\{|i\rangle\}$ is an orthonormal basis of $\mathcal{H}^{A_I^k} \otimes \mathcal{H}^{A_O^k}$.

Lemma B1 ('Teleportation technique'). *Consider a process matrix $W \in \mathcal{L}^{\mathcal{N}}$, to which one attaches a maximally entangled state $|\Phi^+\rangle\langle\Phi^+|^{A_{I''}^k/A_{I'}^{k'}}$ shared by parties A_k and $A_{k'}$, with dimensions $d_{A_{I''}^k} = d_{A_{I'}^{k'}} = d_{A_{IO}^k}$, and possibly some other ancillary state $\tilde{\rho}$ in some further extension $A_{I'}^{\mathcal{N}\setminus k'} \otimes A_{I''}^{k'}$. Consider then the case where party A_k applies the CP map represented by the positive semidefinite CJ matrix $M_k = |\Phi^+\rangle\langle\Phi^+|^{A_{IO}^k/A_{I''}^k} \otimes \mathbb{1}^{A_{I'}^k}$. The resulting conditional matrix for the other $N-1$ parties is then*

$$\begin{aligned}
&(W \otimes |\Phi^+\rangle\langle\Phi^+|^{A_{I''}^k/A_{I'}^{k'}} \otimes \tilde{\rho})_{|M_k = |\Phi^+\rangle\langle\Phi^+|^{A_{IO}^k/A_{I''}^k} \otimes \mathbb{1}^{A_{I'}^k}} \\
&\quad := \mathrm{Tr}_k[|\Phi^+\rangle\langle\Phi^+|^{A_{IO}^k/A_{I''}^k} \otimes \mathbb{1}^{A_{I'}^k} \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W \otimes |\Phi^+\rangle\langle\Phi^+|^{A_{I''}^k/A_{I'}^{k'}} \otimes \tilde{\rho}] \\
&\quad = \frac{1}{(d_{A_{IO}^k})^2}\, W^{A_{IO}^k \to A_{I'}^{k'}} \otimes \mathrm{Tr}_k[\tilde{\rho}].
\end{aligned} \tag{B2}$$

Proof. For clarity, let us write explicitly as superscripts the spaces in which the various operators act. We have:

$$\begin{aligned}
&(W^{A_{IO}^{\mathcal{N}}} \otimes |\Phi^+\rangle\langle\Phi^+|^{A_{I''}^k/A_{I'}^{k'}} \otimes \tilde{\rho}^{A_{I'}^{\mathcal{N}\setminus k'} A_{I''}^{k'}})_{|M_k = |\Phi^+\rangle\langle\Phi^+|^{A_{IO}^k/A_{I''}^k} \otimes \mathbb{1}^{A_{I'}^k}} \\
&= \mathrm{Tr}_{A_{II'I''O}^k}[|\Phi^+\rangle\langle\Phi^+|^{A_{IO}^k/A_{I''}^k} \otimes \mathbb{1}^{A_{I'}^k} \otimes \mathbb{1}^{A_{II'O}^{\mathcal{N}\setminus k} A_{I''}^{k'}} \cdot W^{A_{IO}^{\mathcal{N}}} \otimes |\Phi^+\rangle\langle\Phi^+|^{A_{I''}^k/A_{I'}^{k'}} \otimes \tilde{\rho}^{A_{I'}^{\mathcal{N}\setminus k'} A_{I''}^{k'}}] \\
&= \frac{1}{(d_{A_{IO}^k})^2} \sum_{i,i',j,j'} \mathrm{Tr}_{A_{II''O}^k}[|i\rangle\langle i'|^{A_{IO}^k} \otimes |i\rangle\langle i'|^{A_{I''}^k} \otimes \mathbb{1}^{A_{IO}^{\mathcal{N}\setminus k}} \otimes \mathbb{1}^{A_{I'}^{k'}} \cdot W^{A_{IO}^{\mathcal{N}}} \otimes |j\rangle\langle j'|^{A_{I''}^k} \otimes |j\rangle\langle j'|^{A_{I'}^{k'}}] \otimes \mathrm{Tr}_{A_{I'}^k}[\tilde{\rho}^{A_{I'}^{\mathcal{N}\setminus k'} A_{I''}^{k'}}] \\
&= \frac{1}{(d_{A_{IO}^k})^2} \sum_{i,j} \mathrm{Tr}_{A_{IO}^k}[|i\rangle\langle j|^{A_{IO}^k} \otimes \mathbb{1}^{A_{IO}^{\mathcal{N}\setminus k}} \cdot W^{A_{IO}^{\mathcal{N}}}] \otimes |j\rangle\langle i|^{A_{I'}^{k'}} \otimes \mathrm{Tr}_{A_{I'}^k}[\tilde{\rho}^{A_{I'}^{\mathcal{N}\setminus k'} A_{I''}^{k'}}] = \frac{1}{(d_{A_{IO}^k})^2}\, W^{A_{IO}^k \to A_{I'}^{k'}} \otimes \mathrm{Tr}_k[\tilde{\rho}].
\end{aligned} \tag{B3}$$

□

We shall also use the following facts in (some of) the proofs below:

Proposition B1. *Without loss of generality, each $W^{\rho}_{(k)}$ in Definition (5) can be taken to be of the form $W_{(k)} \otimes \rho$. Equation (13) then implies the direct decomposition $W = \sum_{k\in\mathcal{N}} q_k\, W_{(k)}$, with each $W_{(k)} \in A^{\mathcal{N}}_{IO}$ being a process matrix compatible with party A_k acting first (and such that for any CP map $M_k \in A^{k}_{II'O}$, $(W_{(k)} \otimes \rho)_{|M_k}$ is causally separable).*

Proof. If ρ is pure, then from the extremality of pure states it follows that $W^{\rho}_{(k)} = W_{(k)} \otimes \rho$. If ρ is mixed one can first purify it by introducing an additional incoming system for some arbitrary party, obtain the appropriate decomposition (13) for its purification, and then trace out the additional incoming space just introduced to reach the same conclusion. As $W^{\rho}_{(k)}$ is compatible with A_k acting first and $W_{(k)} = \mathrm{Tr}_{A^{\mathcal{N}}_{I'}} W^{\rho}_{(k)}$, then $W_{(k)}$ itself must also be compatible with A_k acting first (see remarks in appendix A.3.1). □

Proposition B2. *In a scenario where the parties' incoming spaces are decomposed as $A^{\mathcal{N}}_{I} \otimes A^{\mathcal{N}}_{I'}$ (possibly with some trivial spaces A^{k}_{I} or $A^{k}_{I'}$), if a process matrix $W \in A^{\mathcal{N}}_{II'O}$ is causally separable, then so is $\mathrm{Tr}_{I'}\, W \in A^{\mathcal{N}}_{IO}$ (with $\mathrm{Tr}_{I'}$:=$\mathrm{Tr}_{A^{\mathcal{N}}_{I'}}$).*

Proof. For $N = 1$ party any process matrix is by definition causally separable, so that the claim is trivial.

Suppose the claim holds true in the $(N-1)$-partite case. If $W \in A^{\mathcal{N}}_{II'O}$ is causally separable then by definition 5, for any extension $A^{\mathcal{N}}_{I''}$ of the parties' incoming systems and any ancillary quantum state $\rho \in A^{\mathcal{N}}_{I''}$, $W \otimes \rho$ has a decomposition of the form $W \otimes \rho = \sum_k q_k W^{\rho}_{(k)}$ with each $W^{\rho}_{(k)}$ a valid process matrix compatible with party A_k first, and such that for any possible CP map $M_k{}' \in A^{k}_{II'I''O}$ applied by party A_k, the conditional $(N-1)$-partite process matrix $(W^{\rho}_{(k)})_{|M_k{}'} := \mathrm{Tr}_k[M_k{}' \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W^{\rho}_{(k)}]$ is itself causally separable.

One then has $\mathrm{Tr}_{I'}\, W \otimes \rho = \sum_k q_k (\mathrm{Tr}_{I'}\, W^{\rho}_{(k)})$, with $\mathrm{Tr}_{I'}\, W^{\rho}_{(k)}$ a valid process matrix compatible with party A_k first (see remarks in appendix A.3.1). For any possible CP map $M_k \in A^{k}_{II''O}$ applied by party A_k, one has $(\mathrm{Tr}_{I'}\, W^{\rho}_{(k)})_{|M_k} := \mathrm{Tr}_k[M_k \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot (\mathrm{Tr}_{I'}\, W^{\rho}_{(k)})] = \mathrm{Tr}_{I'}\mathrm{Tr}_k[M_k \otimes \mathbb{1}^{A^{k}_{I'}} \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W^{\rho}_{(k)}] = \mathrm{Tr}_{I'}[(W^{\rho}_{(k)})_{|M_k{}'=M_k\otimes\mathbb{1}^{A^{k}_{I'}}}]$. As stated above, $(W^{\rho}_{(k)})_{|M_k{}'=M_k\otimes\mathbb{1}^{A^{k}_{I'}}}$ is causally separable, and by the induction hypothesis so is $\mathrm{Tr}_{I'}[(W^{\rho}_{(k)})_{|M_k{}'=M_k\otimes\mathbb{1}^{A^{k}_{I'}}}] = (\mathrm{Tr}_{I'}\, W^{\rho}_{(k)})_{|M_k}$. We thus have a valid causally separable decomposition of $\mathrm{Tr}_{I'}\, W \otimes \rho$ for any extension ρ, which proves that $\mathrm{Tr}_{I'}\, W$ is causally separable, and which thus proves, by induction, the claim of proposition B2. □

Again, this property is quite intuitive: clearly, discarding some parts of the incoming systems cannot induce some causal nonseparability where there was none previously. As for the similar statements for valid process matrices and for process matrices compatible with a fixed causal order discussed in appendix A.3.1, the converse is not necessarily true: if $\mathrm{Tr}_{I'}\, W$ is a causally separable process matrix, then W may not necessarily be causally separable[20]—unless W is of the product form $W = W_0^{A^{\mathcal{N}}_{IO}} \otimes \rho^{A^{\mathcal{N}}_{I'}}$, in which case by our definition 5 if W_0 is a causally separable process matrix then so is $W = W_0 \otimes \rho$.

B.1. Tripartite causally separable process matrices

B.1.1. Particular tripartite case with $d_{C_O} = 1$. Let us start by considering the tripartite case where party C has no outgoing system (or equivalently, a trivial outgoing system, i.e. $d_{C_O} = 1$). The following proposition directly implies (after proper re-normalisation with appropriate weights $q, 1-q$) the first part of proposition 1, namely the equivalence in that case between Araújo *et al*'s causal separability and OG's extensible causal separability (which, we recall, is what we simply call causal separability here).

Proposition B3 (Characterisation of tripartite causally separable process matrices with $d_{C_O} = 1$). *In a tripartite scenario where party C has no outgoing system, a matrix $W \in A_{IO} \otimes B_{IO} \otimes C_I$ is a valid tripartite causally separable process matrix (as per definition 5) if and only if it can be decomposed as*

$$W = W_{(A,B,C)} + W_{(B,A,C)}, \tag{B4}$$

where, for each permutation (X, Y) of the two parties A and B, $W_{(X,Y,C)}$ is a positive semidefinite matrix satisfying

$$[1-X_O]\, Y_{IO} C_I W_{(X,Y,C)} = 0,\ [1-Y_O] C_I W_{(X,Y,C)} = 0 \tag{B5}$$

(i.e. $W_{(X,Y,C)}$ is a valid process matrix compatible with the causal order $X \prec Y \prec C$).

[20] As a counterexample, consider some causally nonseparable bipartite process matrix $W \in A_{I'O} \otimes B_{IO}$. The process matrix $\mathrm{Tr}_{A_{I'}} W \in A_O \otimes B_{IO}$ is then compatible with the order $A \prec B$ and thus causally separable (see appendix A.2.3), although W is not.

Proof. Consider a causally separable process matrix $W \in A_{IO}B_{IO}C_I$. Let us then introduce an extension $A_{I'} \otimes C_{I''}$ of parties A and C's incoming spaces, of dimensions $d_{A_{I'}} = d_{C_{I''}} = d_{C_I}$, and consider attaching to W the maximally entangled ancillary state $\rho = |\Phi^+\rangle\langle\Phi^+|^{C_{I''}/A_{I'}}$.

As W is assumed to be causally separable, according to definition 5 and proposition B1 it must be decomposable as

$$W = W_{(A)} + W_{(B)} + W_{(C)}, \tag{B6}$$

where each term $W_{(X)} \in A_{IO}B_{IO}C_I$ is a (nonnormalised) process matrix compatible with party X acting first, and such that whatever CP map that party applies to their share of $W_{(X)} \otimes \rho$, the resulting conditional process matrix for the other two parties is causally separable.

As $W_{(A)}$ is compatible both with A first and with C last (since $d_{C_O} = 1$, see appendix A.2.3), it is compatible with the fixed causal order $A \prec B \prec C$; formally, it satisfies ${}_{[1-A_O]B_{IO}C_I}W_{(A)} = {}_{[1-B_O]C_I}W_{(A)} = 0$, see equation (A17). Similarly, $W_{(B)}$ is compatible with the order $B \prec A \prec C$.

Consider now the term $W_{(C)}$. Letting party C act first on $W_{(C)} \otimes \rho = W_{(C)} \otimes |\Phi^+\rangle\langle\Phi^+|^{C_{I''}/A_{I'}}$ and project his incoming systems onto the maximally entangled state $|\Phi^+\rangle^{C_I/C_I''}$, according to lemma B1 (with a trivial extra ancillary state $\tilde{\rho}$), parties A and B are then left with the conditional process matrix

$$(W_{(C)} \otimes \rho)_{|M_C = |\Phi^+\rangle\langle\Phi^+|^{C_I/C_I''}} \propto W_{(C)}^{C_I \to A_{I'}}. \tag{B7}$$

By assumption this conditional process matrix must be a (bipartite) causally separable process matrix: according to proposition 2, there must therefore exist a decomposition for $W_{(C)}^{C_I \to A_{I'}}$ of the form

$$W_{(C)}^{C_I \to A_{I'}} = W_{(C),(A,B)}^{C_I \to A_{I'}} + W_{(C),(B,A)}^{C_I \to A_{I'}} \tag{B8}$$

with $W_{(C),(A,B)}^{C_I \to A_{I'}}$, $W_{(C),(B,A)}^{C_I \to A_{I'}} \in A_{II'O}B_{IO}$ two process matrices compatible with the fixed orders $A \prec B$ and $B \prec A$, respectively—i.e. satisfying

$$\begin{aligned} {}_{[1-A_O]B_{IO}}W_{(C),(A,B)}^{C_I \to A_{I'}} &= {}_{[1-B_O]}W_{(C),(A,B)}^{C_I \to A_{I'}} = 0, \\ {}_{[1-B_O]A_{II'O}}W_{(C),(B,A)}^{C_I \to A_{I'}} &= {}_{[1-A_O]}W_{(C),(B,A)}^{C_I \to A_{I'}} = 0. \end{aligned} \tag{B9}$$

Recall now that $W_{(C)}^{C_I \to A_{I'}}$ is formally the same matrix as $W_{(C)}$, except that system C_I is replaced by $A_{I'}$. Changing back $A_{I'}$ into C_I in equation (B8), we obtain the decomposition

$$W_{(C)} = W_{(C),(A,B)} + W_{(C),(B,A)} \tag{B10}$$

with two positive semidefinite matrices $W_{(C),(A,B)}$, $W_{(C),(B,A)} \in A_{IO}B_{IO}C_I$ satisfying

$$\begin{aligned} {}_{[1-A_O]B_{IO}}W_{(C),(A,B)} &= {}_{[1-B_O]}W_{(C),(A,B)} = 0, \\ {}_{[1-B_O]A_{IO}C_I}W_{(C),(B,A)} &= {}_{[1-A_O]}W_{(C),(B,A)} = 0, \end{aligned} \tag{B11}$$

as implied by equation (B9) after replacing $A_{I'}$ by C_I. These constraints further imply that

$$\begin{aligned} {}_{[1-A_O]B_{IO}C_I}W_{(C),(A,B)} &= {}_{[1-B_O]C_I}W_{(C),(A,B)} = 0, \\ {}_{[1-B_O]A_{IO}C_I}W_{(C),(B,A)} &= {}_{[1-A_O]C_I}W_{(C),(B,A)} = 0, \end{aligned} \tag{B12}$$

i.e. that $W_{(C),(A,B)}$ and $W_{(C),(B,A)}$ are process matrices compatible with the fixed causal orders $A \prec B \prec C$ and $B \prec A \prec C$, respectively (see equation (A17)).

From equations (B6) and (B10), and by defining $W_{(A,B,C)} \coloneqq W_{(A)} + W_{(C),(A,B)} \geqslant 0$ and $W_{(B,A,C)} \coloneqq W_{(B)} + W_{(C),(B,A)} \geqslant 0$, we thus find that W indeed has a decomposition of the form of equation (B4), with each term satisfying the constraints of equation (B5).

Conversely, any process matrix W that can be decomposed as in equation (B4), with process matrices $W_{(A,B,C)}$ and $W_{(B,A,C)}$ satisfying the constraints of equation (B5)—i.e. being compatible with the causal orders $A \prec B \prec C$ and $B \prec A \prec C$—is clearly causally separable, which concludes the proof of proposition B3. □

B.1.2. General tripartite causally separable process matrices. We now turn to proving proposition 3, which characterises causal separability in the general tripartite scenario where all three parties have nontrivial incoming and outgoing systems.

Proof. Consider a causally separable process matrix $W \in A_{IO}B_{IO}C_{IO}$. Let us introduce here an extension $A_{I'} \otimes A_{I''} \otimes B_{I'} \otimes B_{I''} \otimes C_{I'} \otimes C_{I''}$ of all three parties' incoming spaces, with dimensions $d_{A_{I''}} = d_{B_{I'}} = d_{A_{IO}}$, $d_{B_{I''}} = d_{C_{I'}} = d_{B_{IO}}$ and $d_{C_{I''}} = d_{A_{I'}} = d_{C_{IO}}$, and let us attach to W the state $\rho = |\Phi^+\rangle\langle\Phi^+|^{A_{I''}/B_{I'}} \otimes |\Phi^+\rangle\langle\Phi^+|^{B_{I''}/C_{I'}} \otimes |\Phi^+\rangle\langle\Phi^+|^{C_{I''}/A_{I'}}$.

According to definition 5 and proposition B1, W must be decomposable as

$$W = W_{(A)} + W_{(B)} + W_{(C)}, \tag{B13}$$

where each term $W_{(X)} \in A_{IO}B_{IO}C_{IO}$ is a process matrix compatible with party X acting first—so that it satisfies in particular ${}_{[1-X_O]Y_{IO}Z_{IO}}W_{(X)} = 0$ (with $X \neq Y \neq Z$, see equation (A13)), as in equation (23)—and such that whatever that party does on $W_{(X)} \otimes \rho$, the resulting conditional process matrix for the other two parties is causally separable.

Consider the first term in equation (B13). Letting party A act first on $W_{(A)} \otimes \rho$ and perform the operation described by the CJ operator $M_A = |\Phi^+\rangle\langle\Phi^+|^{A_{IO}/A_{I''}} \otimes \mathbb{1}^{A_{I'}} \geqslant 0$, we find, using lemma B1 (with $\tilde{\rho} = \mathrm{Tr}_{A_{I''}B_{I'}}[\rho]$), that the remaining parties $B,\ C$ are left with the conditional process matrix

$$(W_{(A)} \otimes \rho)_{|M_A=|\Phi^+\rangle\langle\Phi^+|^{A_{IO}/A_{I''}}\otimes\mathbb{1}^{A_{I'}}} \propto W_{(A)}^{A_{IO}\to B_{I'}} \otimes \mathrm{Tr}_{A_{I'I''}B_{I'}}[\rho]. \tag{B14}$$

By assumption this conditional process matrix—and therefore $W_{(A)}^{A_{IO}\to B_{I'}}$ itself (according to proposition B2, after tracing out $\mathrm{Tr}_{A_{I'I''}B_{I'}}[\rho]$ completely)—must be a (bipartite) causally separable process matrix: there must therefore exist a decomposition of the form

$$W_{(A)}^{A_{IO}\to B_{I'}} = W_{(A,B,C)}^{A_{IO}\to B_{I'}} + W_{(A,C,B)}^{A_{IO}\to B_{I'}}, \tag{B15}$$

where $W_{(A,B,C)}^{A_{IO}\to B_{I'}},\ W_{(A,C,B)}^{A_{IO}\to B_{I'}} \in B_{II'O}C_{IO}$ are bipartite processes compatible with the causal orders $B \prec C$ and $C \prec B$, respectively. After re-attributing the system $B_{I'}$ to A_{IO}, we obtain a decomposition for $W_{(A)}$,

$$W_{(A)} = W_{(A,B,C)} + W_{(A,C,B)}, \tag{B16}$$

with the positive semidefinite matrices $W_{(A,B,C)},\ W_{(A,C,B)} \in A_{IO}B_{IO}C_{IO}$ satisfying the following constraints, obtained (as we did in the previous subsection) after replacing $B_{I'}$ by A_{IO} in the constraints satisfied by $W_{(A,B,C)}^{A_{IO}\to B_{I'}}$ and $W_{(A,C,B)}^{A_{IO}\to B_{I'}}$:

$$\begin{aligned} {}_{[1-B_O]C_{IO}}W_{(A,B,C)} = {}_{[1-C_O]}W_{(A,B,C)} = 0, \\ {}_{[1-C_O]A_{IO}B_{IO}}W_{(A,C,B)} = {}_{[1-B_O]}W_{(A,C,B)} = 0. \end{aligned} \tag{B17}$$

Furthermore, since $W_{(A)}$ is compatible with A acting first it satisfies in particular ${}_{[1-C_O]B_{IO}}W_{(A)} = 0$ (see equation (A13)), and because of equation (B17), we also have ${}_{[1-C_O]B_{IO}}W_{(A,B,C)} = 0$. Given that $W_{(A,C,B)} = W_{(A)} - W_{(A,B,C)}$, we have ${}_{[1-C_O]B_{IO}}W_{(A,C,B)} = 0$ as well[21]. Together with equation (B17), we thus find that all constraints of equation (24) for $X = A$ are satisfied. One can similarly show that they are satisfied for $X = B,\ C$, which proves (since we noted before that equation (23) is also satisfied) that the decomposition of proposition 3 is indeed a necessary condition for any causally separable process matrix W.

Conversely, suppose a matrix $W \in A_{IO} \otimes B_{IO} \otimes C_{IO}$ has a decomposition of the form (22) that satisfies equations (23)–(24). Then as we noted right after proposition 3, each term $W_{(X)}$ is a valid process matrix, compatible with party X acting first. For any CP map M_X applied by party X on its share of $W_{(X)}$, the resulting conditional process matrix for the other two parties $Y,\ Z$ is

$$(W_{(X)})_{|M_X} = \mathrm{Tr}_X[M_X \otimes \mathbb{1}^{Y_{IO}Z_{IO}} \cdot W_{(X)}] = \mathrm{Tr}_X[M_X \otimes \mathbb{1}^{Y_{IO}Z_{IO}} \cdot (W_{(X,Y,Z)} + W_{(X,Z,Y)})] = (W_{(X,Y,Z)})_{|M_X} + (W_{(X,Z,Y)})_{|M_X} \tag{B18}$$

with $(W_{(X,Y,Z)})_{|M_X} \geqslant 0$ satisfying

$$\begin{aligned} {}_{[1-Y_O]Z_{IO}}[(W_{(X,Y,Z)})_{|M_X}] = [{}_{[1-Y_O]Z_{IO}}W_{(X,Y,Z)}]_{|M_X} = 0, \\ {}_{[1-Z_O]}[(W_{(X,Y,Z)})_{|M_X}] = [{}_{[1-Z_O]}W_{(X,Y,Z)}]_{|M_X} = 0 \end{aligned} \tag{B19}$$

(and similarly for $(W_{(X,Z,Y)})_{|M_X}$), as follows from equation (24). This shows that $(W_{(X,Y,Z)})_{|M_X}$ and $(W_{(X,Z,Y)})_{|M_X}$ are valid bipartite process matrices compatible with the orders $Y \prec Z$ and $Z \prec Y$, respectively, so that $(W_{(X)})_{|M_X}$ is causally separable. Note that for any ancillary state ρ, $W \otimes \rho$ also has a decomposition as in equation (22), obtained simply by attaching the ancillary state to every individual term in the decomposition of W. Therefore, the same reasoning as above applies, which implies that W is causally separable. This thus shows that the

[21] Note that this is the step where the tripartite proof does not generalise straightforwardly to $N \geqslant 4$ parties. In particular, we cannot use the same argument to prove that the constraints (29) that appear in our necessary condition are satisfied without tracing out the X_{IO} on the first and fourth lines (as one would need if the terms in equation (28) were to satisfy equation (32) and thus specify a decomposition satisfying also our sufficient condition for causal separability). One indeed obtains e.g. ${}_{[1-Z_O]Y_{IO}T_{IO}}W_{(X,Z)}^{[X\to Y]} = {}_{[1-Z_O]Y_{IO}T_{IO}}(W_{(X)} - W_{(X,Y,Z,T)}^{[X\to Y]} - W_{(X,Y,T,Z)}^{[X\to Y]} - W_{(X,T,Y,Z)}^{[X\to Y]} - W_{(X,T,Z,Y)}^{[X\to Y]}) = -({}_{[1-Z_O]Y_{IO}T_{IO}}W_{(X,T,Z,Y)}^{[X\to Y]})$ which, *a priori*, may still be nonzero.

decomposition of proposition 3 is also a sufficient condition for a matrix W to represent a causally separable process matrix, which concludes the proof of that proposition. □

Let us mention here that proposition B3, for the particular tripartite case where $d_{C_O} = 1$, could also be obtained as a corollary of the general tripartite case considered by proposition 3. Indeed, in the case where $d_{C_O} = 1$ the matrices $W_{(A)}$ and $W_{(B)}$ in equation (22) are compatible with the fixed causal orders $A \prec B \prec C$ and $B \prec A \prec C$, respectively (as they are compatible with both A or B first, and C last); furthermore, the matrix $W_{(C,A,B)}$ satisfies ${}_{[1-A_O]B_{IO}}W_{(C,A,B)} = {}_{[1-B_O]}W_{(C,A,B)} = 0$ and therefore ${}_{[1-A_O]B_{IO}C_I}W_{(C,A,B)} = {}_{[1-B_O]C_I}W_{(C,A,B)} = 0$, which implies that it is also compatible with $A \prec B \prec C$; and similarly, the matrix $W_{(C,B,A)}$ is also compatible with $B \prec A \prec C$. The decomposition of equation (22) thus provides a decomposition in the form $W = \tilde{W}_{(A,B,C)} + \tilde{W}_{(B,A,C)}$ with $\tilde{W}_{(A,B,C)} = W_{(A)} + W_{(C,A,B)}$ and $\tilde{W}_{(B,A,C)} = W_{(B)} + W_{(C,B,A)}$ satisfying the constraints of equation(B5).

B.1.3. Particular tripartite case with $d_{A_I} = 1$. Another particular tripartite case of interest is one where one party, say now A, has no *incoming* space (or a trivial one, with $d_{A_I} = 1$). The following characterisation is also obtained as a corollary of the general tripartite case above.

Proposition B4 (Characterisation of tripartite causally separable process matrices with $d_{A_I} = 1$). *In a tripartite scenario where party A has no incoming system, a matrix $W \in A_O \otimes B_{IO} \otimes C_{IO}$ is a valid tripartite causally separable process matrix (as per definition 5) if and only if*

$$ {}_{[1-A_O]B_{IO}C_{IO}}W = 0 \tag{B20}$$

and W can be decomposed as

$$ W = W_{(A,B,C)} + W_{(A,C,B)}, \tag{B21}$$

where, for each permutation (X, Y) of the two parties B and C, $W_{(A,X,Y)}$ is a positive semidefinite matrix satisfying

$$ {}_{[1-X_O]Y_{IO}}W_{(A,X,Y)} = 0, \ {}_{[1-Y_O]}W_{(A,X,Y)} = 0. \tag{B22}$$

Note already that contrary to the decomposition of proposition B3, the two summands $W_{(A,B,C)}$ and $W_{(A,C,B)}$ above are not necessarily valid process matrices: indeed, they are not required to satisfy ${}_{[1-A_O]B_{IO}C_{IO}}W_{(A,X,Y)} = 0$ (only their sum must satisfy equation (B20)). This allows for dynamical causal orders, where A (incoherently) controls the causal order between the next parties B and C.

Proof. According to proposition B7, a causally separable process matrix $W \in A_O \otimes B_{IO} \otimes C_{IO}$ must have a decomposition of the form (22) that satisfies the constraints (23)–(24).

In particular, the constraints ${}_{[1-A_O]}W_{(B,C,A)} = {}_{[1-C_O]A_O}W_{(B,C,A)} = 0$ imply ${}_{[1-C_O]}W_{(B,C,A)} = 0$. The constraints on $W_{(B,A,C)}$ and $W_{(B,C,A)}$ in turn imply that ${}_{[1-A_O]C_{IO}}W_{(B)} = {}_{[1-C_O]}W_{(B)} = 0$, which, together with ${}_{[1-B_O]A_OC_{IO}}W_{(B)} = 0$, further imply that ${}_{[1-A_O]B_{IO}C_{IO}}W_{(B)} = {}_{[1-B_O]C_{IO}}W_{(B)} = 0$. Similarly, one also has ${}_{[1-A_O]B_{IO}C_{IO}}W_{(C)} = {}_{[1-C_O]B_{IO}}W_{(C)} = {}_{[1-B_O]}W_{(C)} = 0$.

It then follows, since $W = W_{(A)} + W_{(B)} + W_{(C)}$ and ${}_{[1-A_O]B_{IO}C_{IO}}W_{(A)} = 0$, that ${}_{[1-A_O]B_{IO}C_{IO}}W = 0$ as well; furthermore, by defining $\tilde{W}_{(A,B,C)} = W_{(A,B,C)} + W_{(B)}$ and $\tilde{W}_{(A,C,B)} = W_{(A,C,B)} + W_{(C)}$, we obtain the decomposition $W = \tilde{W}_{(A,B,C)} + \tilde{W}_{(A,C,B)}$ of the form (B21), with $\tilde{W}_{(A,B,C)}$ and $\tilde{W}_{(A,C,B)}$ satisfying the constraints (B22).

Conversely, it is clear that the decomposition of proposition B4 is a particular case of that of proposition 3 (with $W_{(B)} = W_{(C)} = 0$), so that any process matrix that can be decomposed as in equation (B21) and satisfies equations (B20) and (B22) is causally separable according to proposition 3. □

B.1.4. Allowed and forbidden terms in a Hilbert–Schmidt basis decomposition. We note that [17] already provided a characterisation of general tripartite causally separable process matrices—or 'extensibly' causally separable process matrices, in their terminology. Let us prove here the equivalence with our own characterisation (Proposition 3) explicitly.

According to Proposition 3.3 in [17], every tripartite (extensibly) causally separable process matrix $W \in A_{IO} \otimes B_{IO} \otimes C_{IO}$ can be written in the form

$$ W = W_{(A)} + W_{(B)} + W_{(C)}, \tag{B23}$$

where each $W_{(X)}$ contains only Hilbert–Schmidt terms (see appendix A.4) that are allowed in a process matrix compatible with party X acting first as per Proposition 3.2 in [17]—i.e. $W_{(X)} \in \mathcal{L}^{X \prec \{Y,Z\}}$ in our language—and

has the form

$$W_{(X)} = \Omega_{(X,Y,Z)} \otimes \mathbb{1}^{Z_O} + \Omega_{(X,Z,Y)} \otimes \mathbb{1}^{Y_O}, \tag{B24}$$

where $\Omega_{(X,Y,Z)} \in X_{IO} \otimes Y_{IO} \otimes Z_I$ and $\Omega_{(X,Z,Y)} \in X_{IO} \otimes Y_I \otimes Z_{IO}$ are positive semidefinite. (We changed here the notations of [17] to match ours; note in particular that unlike in [17], we again ignore the normalisation constraints in the decomposition of W.)

Any such process matrix can thus be decomposed as in proposition 3, with $W_{(X,Y,Z)} = \Omega_{(X,Y,Z)} \otimes \mathbb{1}^{Z_O} \geqslant 0$ (which also implies $W_{(X)} = W_{(X,Y,Z)} + W_{(X,Z,Y)} \geqslant 0$). As $W_{(X)} \in \mathcal{L}^{X \prec \{Y,Z\}}$, it satisfies in particular ${}_{[1-X_O]Y_{IO}Z_{IO}}W_{(X)} = 0$ (see equation (A13)), as in equation (23). It is furthermore immediate to see that each $W_{(X,Y,Z)} = \Omega_{(X,Y,Z)} \otimes \mathbb{1}^{Z_O}$ satisfies the second constraint in (24), i.e. ${}_{[1-Z_O]}W_{(X,Y,Z)} = 0$. Finally, one has ${}_{[1-Y_O]Z_{IO}}W_{(X)} = 0$ (see again equation (A13)) and ${}_{[1-Y_O]}W_{(X,Z,Y)} = 0$, which implies ${}_{[1-Y_O]Z_{IO}}W_{(X,Y,Z)} = {}_{[1-Y_O]Z_{IO}}(W_{(X)} - W_{(X,Z,Y)}) = 0$, i.e. the first constraint in equation (24). Thus, any process matrix that satisfies the characterisation of proposition 3.3 from [17] also satisfies that of our proposition 3.Conversely, let W be a process matrix that has a decomposition as in proposition 3. As discussed after that proposition, the conditions (23)–(24) imply that $W_{(X)} \in \mathcal{L}^{X \prec \{Y,Z\}}$. Furthermore, $W_{(X,Y,Z)} \geqslant 0$ and ${}_{[1-Z_O]}W_{(X,Y,Z)} = 0$ implies that $W_{(X,Y,Z)}$ is of the form $\Omega_{(X,Y,Z)} \otimes \mathbb{1}^{Z_O}$ with $\Omega_{(X,Y,Z)} \geqslant 0$, so that $W_{(X)} = W_{(X,Y,Z)} + W_{(X,Z,Y)}$ is of the form (B24). This indeed establishes the equivalence of Proposition 3.3 in [17] with our proposition 3.

As emphasised before, the matrices $W_{(X,Y,Z)} = \Omega_{(X,Y,Z)} \otimes \mathbb{1}^{Z_O}$ need not be valid process matrices (the only requirement is that $\Omega_{(X,Y,Z)} \geqslant 0$). Both individual summands in equation (B24) can thus contain terms that are forbidden in a process matrix compatible with party X acting first, as long as these terms cancel out in the sum. More precisely, in addition to the terms that are allowed in a process matrix with X first, $W_{(X,Y,Z)}$ and $W_{(X,Z,Y)}$ can contain terms of the form $\sigma_\mu^{X_I}\sigma_\nu^{X_O}\mathbb{1}^{Y_{IO}Z_{IO}}$ with $\sigma_\nu^{X_O} \neq \mathbb{1}^{X_O}$ (i.e. $\nu \geqslant 1$). Any other term that is forbidden in a process matrix with X first has a nontrivial σ operator on either Y_O or Z_O, and thus cannot appear in $\Omega_{(X,Y,Z)}$ or $\Omega_{(X,Z,Y)}$, and cannot be cancelled out in equation (B24). In the explicit example of equation (41) given at the end of section 4.5, for instance, on can check indeed that $W_{(A,B,C)}$ and $W_{(A,C,B)}$ contain the Hilbert–Schmidt term $\hat{z}^{A_O^c}\mathbb{1}^{B_{IO}^i C_{IO}^i}$, which come with opposite signs and cancel out in the sum.

B.2. General multipartite causally separable process matrices

B.2.1. Necessary condition for causal separability. Let us now prove the necessary condition for general multipartite causal separability given by proposition 4.

Proof. Consider an N-partite causally separable process matrix $W \in A_{IO}^{\mathcal{N}}$. Let us introduce now, for each party A_k, an extension $\bigotimes_{k' \in \mathcal{N}\setminus k}(A^k_{I'_{(\leftarrow k')}} \otimes A^k_{I''_{(\rightarrow k')}})$ of their incoming space, with dimensions $d_{A^k_{I'_{(\leftarrow k')}}} = d_{A^{k'}_{IO}}$ and $d_{A^k_{I''_{(\rightarrow k')}}} = d_{A^k_{IO}}$, and let us attach to W the state $\rho = \bigotimes_{k,k' \in \mathcal{N}, k \neq k'}(|\Phi^+\rangle\langle\Phi^+|^{A^k_{I''_{(\rightarrow k')}}/A^{k'}_{I'_{(\leftarrow k)}}} \otimes |\Phi^+\rangle\langle\Phi^+|^{A^{k'}_{I''_{(\rightarrow k)}}/A^k_{I'_{(\leftarrow k')}}})$—i.e. we provide each pair of parties with two maximally entangled states, which will allow us to use the teleportation technique in either direction.

According to definition 5 and proposition B1, W must be decomposable as

$$W = \sum_{k \in \mathcal{N}} W_{(k)}, \tag{B25}$$

where each term $W_{(k)}$ is a process matrix compatible with party A_k acting first, and such that whatever that party does on $W_{(k)} \otimes \rho$, the resulting conditional process matrix for the other $N-1$ parties is causally separable.

Consider, for a given $k \in \mathcal{N}$, the process matrix $W_{(k)} \otimes \rho$, and let party A_k perform, for a given $k' \neq k$, the CP map $M_k = |\Phi^+\rangle\langle\Phi^+|^{A^k_{IO}/A^k_{I'_{(\leftarrow k')}}} \otimes \mathbb{1}^{A^k_{I\text{rest}'}} \geqslant 0$, with $A^k_{I\text{rest}'} = A^k_{I'_{(\leftarrow k')}} \bigotimes_{k'' \in \mathcal{N}\setminus\{k,k'\}}(A^k_{I'_{(\leftarrow k'')}} \otimes A^k_{I''_{(\rightarrow k'')}})$. The resulting conditional process matrix for the other $N-1$ parties is then, according to lemma B1 (with $\tilde{\rho} = \mathrm{Tr}_{A^k_{I''_{(\rightarrow k')}} A^{k'}_{I'_{(\leftarrow k)}}}[\rho]$),

$$(W_{(k)} \otimes \rho)_{|M_k} \propto W_{(k)}^{A^k_{IO} \to A^{k'}_{I'_{(\leftarrow k)}}} \otimes \mathrm{Tr}_k \mathrm{Tr}_{A^{k'}_{I'_{(\leftarrow k)}}}[\rho]. \tag{B26}$$

As this conditional process matrix must be causally separable, it then follows from proposition B2 that $W_{(k)}^{A^k_{IO} \to A^{k'}_{I'_{(\leftarrow k)}}}$ itself must be causally separable, which concludes the proof of proposition 4 (where $W_{(k)}^{A^k_{IO} \to A^{k'}_{I'_{(\leftarrow k)}}}$ was simply denoted $W_{(k)}^{A^k_{IO} \to A^{k'}_{I'}}$). □

B.2.2. Sufficient condition for causal separability. Here we shall prove the sufficient condition for general multipartite causal separability given by proposition 5. Let us however first prove the claim that was made just

after that proposition, namely that if equation (32) is satisfied for all $(k_1, \ldots, k_n)$, then one also has, for all $(k_1, \ldots, k_n)$ with $1 \leqslant n < N$, that

$$\forall\ \mathcal{X} \subseteq \mathcal{N}\setminus\{k_1, \ldots, k_n\},\ \mathcal{X} \neq \varnothing,\ \prod_{i\in\mathcal{X}}[1-A_O^i]A_{IO}^{\mathcal{N}\setminus\{k_1, \ldots, k_n\}\setminus\mathcal{X}}W_{(k_1, \ldots, k_n)} = 0. \tag{B27}$$

Proof. This can be seen (by induction) as follows. Assume that equation (32) is satisfied for all $(k_1, \ldots, k_n)$.

Equation (B27) is indeed satisfied for all $(k_1, \ldots, k_n)$ for $n = N - 1$, as in that case $W_{(k_1, \ldots, k_{N-1})} = W_{(k_1, \ldots, k_{N-1}, k_N)}$ satisfies ${}_{[1-A_O^{k_N}]}W_{(k_1, \ldots, k_{N-1})} = 0$ by assumption (32).

Suppose then that for a given value of $n \geqslant 2$, equation (B27) is satisfied for all $(k_1, \ldots, k_n)$. Consider a given ordered subset of parties $(k_1, \ldots, k_{n-1})$ and a given nonempty subset $\mathcal{X} \subseteq \mathcal{N}\setminus\{k_1, \ldots, k_{n-1}\}$, and define for ease of notation the linear function $f^{\mathcal{X}}_{\{k_1, \ldots, k_{n-1}\}}(W) \coloneqq \prod_{i\in\mathcal{X}}[1-A_O^i]A_{IO}^{\mathcal{N}\setminus\{k_1, \ldots, k_{n-1}\}\setminus\mathcal{X}}W$.

Let then $k_n \in \mathcal{N}\setminus\{k_1, \ldots, k_{n-1}\}$. If $\mathcal{X} = \{k_n\}$ then $f_{\mathcal{X}}(W_{(k_1, \ldots, k_n)}) = {}_{[1-A_O^{k_n}]}A_{IO}^{\mathcal{N}\setminus\{k_1, \ldots, k_{n-1}, k_n\}}W_{(k_1, \ldots, k_n)} = 0$ according to equation (32). If on the other hand $\mathcal{X} \neq \{k_n\}$, then depending on whether $k_n \in \mathcal{X}$ or $k_n \notin \mathcal{X}$, we have $f_{\mathcal{X}}(W_{(k_1, \ldots, k_n)}) = {}_{[1-A_O^{k_n}]}(\prod_{i\in\mathcal{X}'}[1-A_O^i]A_{IO}^{\mathcal{N}\setminus\{k_1, \ldots, k_{n-1}, k_n\}\setminus\mathcal{X}'}W_{(k_1, \ldots, k_n)})$ with $\mathcal{X}' = \mathcal{X}\setminus k_n$, or $f_{\mathcal{X}}(W_{(k_1, \ldots, k_n)}) = {}_{A_{IO}^{k_n}(\prod_{i\in\mathcal{X}}[1-A_O^i]A_{IO}^{\mathcal{N}\setminus\{k_1, \ldots, k_{n-1}, k_n\}\setminus\mathcal{X}}W_{(k_1, \ldots, k_n)})}$. In both cases, $\mathcal{X}^{(\prime)} \subseteq \mathcal{N}\setminus\{k_1, \ldots, k_n\}$, $\mathcal{X}^{(\prime)} \neq \varnothing$, so that by the induction hypothesis $f_{\mathcal{X}}(W_{(k_1, \ldots, k_n)}) = 0$, which thus holds for all $k_n \in \mathcal{N}\setminus\{k_1, \ldots, k_{n-1}\}$. As $W_{(k_1, \ldots, k_{n-1})} = \sum_{k_n\in\mathcal{N}\setminus\{k_1, \ldots, k_{n-1}\}}W_{(k_1, \ldots, k_{n-1}, k_n)}$, we also have $f_{\mathcal{X}}(W_{(k_1, \ldots, k_{n-1})}) = 0$, which, by induction, concludes the proof of equation (B27) for all ordered subsets $(k_1, \ldots, k_n)$ with $1 \leqslant n < N$. □

In particular for the case $n = 1$, equation (B27) together with equation (32) imply that each matrix $W_{(k_1)}$ is a valid process matrix compatible with party A_{k_1} acting first (see equation (A13)).

Let us now prove proposition 5 by induction.

Proof. Clearly, it trivially holds for $N = 1$ (in which case equation (32) ensures in particular that W is a valid process matrix). (Note also that for $N = 2$ and 3, it reduces to the sufficient conditions of propositions 2 and 3, respectively.)

Suppose proposition 5 holds in the $(N - 1)$-partite case, and consider a matrix $W \in A_{IO}^{\mathcal{N}}$ that can be decomposed as in equation (30), with all partial sums $W_{(k_1, \ldots, k_n)}$ satisfying equation (32). Then we have

$$W = \sum_{k_1\in\mathcal{N}} W_{(k_1)} \tag{B28}$$

with (as noted above) each $W_{(k_1)}$ being a valid process matrix compatible with party A_{k_1} acting first.

Consider a CP map M_{k_1} applied by party A_{k_1} on $W_{(k_1)} = \sum_{\pi\in\Pi_{(k_1)}} W_\pi$. The resulting conditional process matrix for the remaining $N - 1$ parties is

$$(W_{(k_1)})_{|M_{k_1}} \coloneqq \mathrm{Tr}_{k_1}[M_{k_1} \otimes \mathbb{1}^{\mathcal{N}\setminus k_1} \cdot W_{(k_1)}] = \sum_{\pi\in\Pi_{(k_1)}} (W_\pi)_{|M_{k_1}} \tag{B29}$$

with $(W_\pi)_{|M_{k_1}} \coloneqq \mathrm{Tr}_{k_1}[M_{k_1} \otimes \mathbb{1}^{\mathcal{N}\setminus k_1} \cdot W_\pi]$. By denoting by $\Pi^{\mathcal{N}\setminus k_1}$ the set of permutations of $\mathcal{N}\setminus k_1$ (and by $\Pi^{\mathcal{N}\setminus k_1}_{(k_2, \ldots, k_n)}$ the set of those that start with $k_2, \ldots, k_n$), by writing any permutation π of $\mathcal{N}$ that starts with k_1 as $\pi = (k_1, \pi')$ with $\pi' \in \Pi^{\mathcal{N}\setminus k_1}$, and by defining $[(W_{(k_1)})_{|M_{k_1}}]_{\pi'} \coloneqq (W_{(k_1, \pi')})_{|M_{k_1}}$, we can re-write equation (B29) as

$$(W_{(k_1)})_{|M_{k_1}} = \sum_{\pi'\in\Pi^{\mathcal{N}\setminus k_1}} [(W_{(k_1)})_{|M_{k_1}}]_{\pi'}, \tag{B30}$$

in a similar form to equation (30). For $n = 2, \ldots, N$, and for any ordered subset of parties $(k_2, \ldots, k_n)$ of $\mathcal{N}\setminus\{k_1\}$, the partial sums

$$[(W_{(k_1)})_{|M_{k_1}}]_{(k_2, \ldots, k_n)} \coloneqq \sum_{\pi'\in\Pi^{\mathcal{N}\setminus k_1}_{(k_2, \ldots, k_n)}} [(W_{(k_1)})_{|M_{k_1}}]_{\pi'} = \sum_{\pi\in\Pi_{(k_1, k_2, \ldots, k_n)}} (W_\pi)_{|M_{k_1}} = (W_{(k_1, k_2, \ldots, k_n)})_{|M_{k_1}} \tag{B31}$$

then satisfy

$$\begin{aligned}{}_{[1-A_O^{k_n}]}A_{IO}^{\mathcal{N}\setminus\{k_1\}\setminus\{k_2, \ldots, k_n\}}[(W_{(k_1)})_{|M_{k_1}}]_{(k_2, \ldots, k_n)} &= {}_{[1-A_O^{k_n}]}A_{IO}^{\mathcal{N}\setminus\{k_1, k_2, \ldots, k_n\}}[(W_{(k_1, k_2, \ldots, k_n)})_{|M_{k_1}}] \\ &= ({}_{[1-A_O^{k_n}]}A_{IO}^{\mathcal{N}\setminus\{k_1, k_2, \ldots, k_n\}}W_{(k_1, k_2, \ldots, k_n)})_{|M_{k_1}} = 0\end{aligned} \tag{B32}$$

by assumption (32). Thus, equation (B30) provides a decomposition of the $(N - 1)$-partite process matrix $(W_{(k_1)})_{|M_{k_1}}$ of the same form as in equation (30), with positive semidefinite matrices $[(W_{(k_1)})_{|M_{k_1}}]_{\pi'}$ and with all partial sums satisfying the analogue constraints as those of equation (32). By the induction hypothesis, this implies that $(W_{(k_1)})_{|M_{k_1}}$ is causally separable.

Note that the exact same reasoning also goes through if instead of W we consider $W \otimes \rho$ with any ancillary state ρ. Indeed, $W \otimes \rho$ also has a decomposition as in equation (30) obtained simply by attaching the ρ to every individual term in the decomposition of W. This shows that W is causally separable, and by induction this proves that the decomposition of proposition 5 is indeed a sufficient condition for W to be causally separable. □

For clarity and to get some better intuition on how it generalises the characterisation of proposition 3 for the tripartite case, let us write the sufficient condition of proposition 5 explicitly in the fourpartite case:

Proposition B5 (Sufficient condition for fourpartite causally separable process matrices). *If a matrix* $W \in A_{IO} \otimes B_{IO} \otimes C_{IO} \otimes D_{IO}$ *can be decomposed as*

$$\begin{aligned} W &= \overbrace{\qquad\qquad W_{(A)} \qquad\qquad}^{} + W_{(B)} + W_{(C)} + W_{(D)} \\ &= \overbrace{W_{(A,B)}}^{} + \overbrace{W_{(A,C)}}^{} + \overbrace{W_{(A,D)}}^{} + \cdots + \cdots + \cdots \\ &= \overbrace{W_{(A,B,C,D)}+W_{(A,B,D,C)}}^{} + \overbrace{W_{(A,C,B,D)}+W_{(A,C,D,B)}}^{} + \overbrace{W_{(A,D,B,C)}+W_{(A,D,C,B)}}^{} + \cdots + \cdots + \cdots \end{aligned} \tag{B33}$$

with, for each permutation of the four parties (X, Y, Z, T), *positive semidefinite matrices* $W_{(X,Y,Z,T)}$, $W_{(X,Y)} \coloneqq W_{(X,Y,Z,T)} + W_{(X,Y,T,Z)}$ *and* $W_{(X)} \coloneqq W_{(X,Y)} + W_{(X,Z)} + W_{(X,T)}$ *satisfying*

$$[1-X_O]Y_{IO}Z_{IO}T_{IO}W_{(X)} = 0, \tag{B34}$$

$$[1-Y_O]Z_{IO}T_{IO}W_{(X,Y)} = 0, \tag{B35}$$

$$[1-Z_O]T_{IO}W_{(X,Y,Z,T)} = 0, \quad [1-T_O]W_{(X,Y,Z,T)} = 0, \tag{B36}$$

then W *is a valid fourpartite causally separable process matrix (as per our definition 5).*

It follows from equations (B35)–(B36) that for each party X, $W_{(X)}$ also satisfies $[1-Y_O]Z_{IO}T_{IO}W_{(X)} = [1-Y_O][1-Z_O]T_{IO}W_{(X)} = [1-Y_O][1-Z_O][1-T_O]W_{(X)} = 0$ for all $X \neq Y \neq Z \neq T$ (see equation (B27)). This, together with equation (B34) and the fact that $W_{(X)} \geqslant 0$, implies that $W_{(X)}$ is a valid process matrix, compatible with party X acting first (see equation (A13)).

Similarly, it follows from equation (B36) that for each pair of parties X, Y, $W_{(X,Y)}$ also satisfies $[1-Z_O]T_{IO}W_{(X,Y)} = [1-Z_O][1-T_O]W_{(X,Y)} = 0$ for all $X \neq Y \neq Z \neq T$. This, together with equation (B35) and the fact that $W_{(X,Y)} \geqslant 0$, implies that whatever CP map M_X party X applies, the conditional process matrix $(W_{(X,Y)})_{|M_X}$ is a valid tripartite process matrix for parties Y, Z, T, compatible with party Y acting first.

Finally, equation (B36) implies that whatever CP maps M_X, M_Y parties X and Y apply, the conditional matrix $(W_{(X,Y,Z,T)})_{|M_X \otimes M_Y}$ is a valid bipartite process matrix for parties Z, T, compatible with party Z acting first.

B.3. Fourpartite causally separable process matrices in the particular case with $d_{D_O} = 1$

Consider now a fourpartite situation where party D has no outgoing system (or a trivial one, with $d_{D_O} = 1$). It turns out that in such a case our sufficient condition above is also necessary, and it simplifies as follows (note the similarity with proposition 3).

Proposition B6 (Characterisation of fourpartite causally separable process matrices with $d_{D_O} = 1$). *In a fourpartite scenario where party* D *has no outgoing system, a matrix* $W \in A_{IO} \otimes B_{IO} \otimes C_{IO} \otimes D_I$ *is a valid fourpartite causally separable process matrix (as per definition 5) if and only if it can be decomposed as*

$$\begin{aligned} W &= \overbrace{\qquad W_{(A)} \qquad}^{} + \overbrace{\qquad W_{(B)} \qquad}^{} + \overbrace{\qquad W_{(C)} \qquad}^{} \\ &= \overbrace{W_{(A,B,C,D)} + W_{(A,C,B,D)}}^{} + \overbrace{W_{(B,A,C,D)} + W_{(B,C,A,D)}}^{} + \overbrace{W_{(C,A,B,D)} + W_{(C,B,A,D)}}^{} \end{aligned} \tag{B37}$$

where, for each permutation (X, Y, Z) *of the three parties* A, B *and* C, $W_{(X,Y,Z,D)}$ *and* $W_{(X)} \coloneqq W_{(X,Y,Z,D)} + W_{(X,Z,Y,D)}$ *are positive semidefinite matrices satisfying*

$$[1-X_O]Y_{IO}Z_{IO}D_I W_{(X)} = 0, \tag{B38}$$

$$[1-Y_O]Z_{IO}D_I W_{(X,Y,Z,D)} = 0, \quad [1-Z_O]D_I W_{(X,Y,Z,D)} = 0. \tag{B39}$$

Proof. According to the necessary condition of proposition 4, a causally separable process matrix $W \in A_{IO} \otimes B_{IO} \otimes C_{IO} \otimes D_I$ must have a decomposition of the form $W = W_{(A)} + W_{(B)} + W_{(C)} + W_{(D)}$ where each $W_{(X)}$ is a process matrix compatible with X first, such that for any $Y \neq X$, $W_{(X)}^{X_{IO} \to Y_{I'}}$ is causally separable.

Consider first $X = A$, and note already that as $W_{(A)}$ is compatible with A first, one has, from equation (A13),

$$[1-A_O]B_{IO}C_{IO}D_I W_{(A)} = 0. \tag{B40}$$

Taking now $Y = B$, we have that $W_{(A)}^{A_{IO}\to B_{I'}} \in B_{II'O} \otimes C_{IO} \otimes D_I$ is a tripartite causally separable process matrix in a scenario where one party (D) has no outgoing space. Using the characterisation of proposition B3, and re-attributing the system $B_{I'}$ back to A_{IO} (as we did, e.g. in the proof of proposition B3), we obtain that $W_{(A)}$ must have a decomposition of the form

$$W_{(A)} = W_{(A,B,C,D)} + W_{(A,C,B,D)} \tag{B41}$$

with $W_{(A,B,C,D)}$, $W_{(A,C,B,D)} \geqslant 0$ satisfying

$$\begin{aligned} {}_{[1-B_O]C_{IO}D_I}W_{(A,B,C,D)} = {}_{[1-C_O]D_I}W_{(A,B,C,D)} = 0, \\ {}_{[1-C_O]A_{IO}B_{IO}D_I}W_{(A,C,B,D)} = {}_{[1-B_O]D_I}W_{(A,C,B,D)} = 0. \end{aligned} \tag{B42}$$

The first line further implies that ${}_{[1-C_O]B_{IO}D_I}W_{(A,B,C,D)} = 0$; noting that ${}_{[1-C_O]B_{IO}D_I}W_{(A)} = 0$ as well (as $W_{(A)}$ is compatible with A first, see again equation (A13)) and that $W_{(A,C,B,D)} = W_{(A)} - W_{(A,B,C,D)}$, we also have

$${}_{[1-C_O]B_{IO}D_I}W_{(A,C,B,D)} = 0. \tag{B43}$$

Hence, the term $W_{(A)}$ can be decomposed as in equation (B37), with the corresponding constraints being satisfied. The same holds, in a similar way, for the terms $W_{(B)}$ and $W_{(C)}$.

Consider now $W_{(D)}$. Taking e.g. $Y = A$, we have that $W_{(D)}^{D_I\to A_{I'}} \in A_{II'O} \otimes B_{IO} \otimes C_{IO}$ is a tripartite causally separable process matrix, which must have a decomposition as in proposition 3. After relabelling the matrices $W_{(X,Y,Z)}$ from equation (22) in the decomposition thus obtained to $W_{(X,Y,Z,D)}^{D_I\to A_{I'}}$, re-attributing the system $A_{I'}$ back to D_I and applying the map ${}_{D_I}\cdot$ to all constraints of equations (23)–(24), we find that $W_{(D)}$ also has a decomposition as in equation (B37) that satisfies the constraints (B38)–(B39).

Altogether, W is thus a combination of terms that have a decomposition as in proposition B6; combining these decompositions, it directly follows that W itself has a decomposition of the form of equation (B37) that satisfies the required constraints.

Conversely, it is easy to see that if a matrix W has a decomposition of the form (B37), then it is also of the form (B33) (where all terms $W_{(X,Y,Z,T)}$ with $D \neq T$ are 0, and thus only the terms $W_{(X,Y,Z,D)} = W_{(X,Y)}$ remain). Furthermore, if the decomposition satisfies the constraints of equations (B38)–(B39), then it also satisfies those of equations (B34)–(B36). According to proposition 5, this implies that such a process matrix W is causally separable. □

One can further simplify the characterisation above in the particular fourpartite case where, in addition to one party (D) having no outgoing system, one also has a party (A) with no incoming system. We then obtain the following:

Proposition B7 (Characterisation of fourpartite causally separable process matrices with $d_{A_I} = 1$ and $d_{D_O} = 1$). *In a fourpartite scenario where party* A *has no incoming system and party* D *has no outgoing system, a matrix $W \in A_O \otimes B_{IO} \otimes C_{IO} \otimes D_I$ is a valid fourpartite causally separable process matrix (as per definition 5) if and only if*

$${}_{[1-A_O]B_{IO}C_{IO}D_I}W = 0 \tag{B44}$$

and W *can be decomposed as*

$$W = W_{(A,B,C,D)} + W_{(A,C,B,D)}, \tag{B45}$$

where, for each permutation (X, Y) of the two parties B and C, $W_{(A,X,Y,D)}$ is a positive semidefinite matrix satisfying

$${}_{[1-X_O]Y_{IO}D_I}W_{(A,X,Y,D)} = 0, \ {}_{[1-Y_O]D_I}W_{(A,X,Y,D)} = 0. \tag{B46}$$

We emphasise again that the two summands $W_{(A,B,C,D)}$ and $W_{(A,C,B,D)}$ above are not necessarily valid process matrices, thus allowing for dynamical causal orders. We omit the proof of proposition B7 here, as it follows that of proposition B4 very closely. We note, as an aside, that both propositions B3 and B4 could be obtained as corollaries of proposition B7 after removing one party. Namely, by imposing $d_{A_O} = 1$ above (which, in particular, makes equation (B44) trivial), ignoring A and relabelling $(B, C, D) \to (A, B, C)$ we obtain proposition B3; by imposing $d_{D_I} = 1$ instead in proposition B7 we directly obtain proposition B4.

To conclude this section, we further note that propositions B6 and B7 generalise straightforwardly to cases with more parties D, E,… that have no outgoing spaces (by simply replacing D_I by $D_I E_I \cdots$). Hence, we can give necessary and sufficient conditions for causal separability in any N-partite scenario in which at most three parties have nontrivial outgoing spaces.

Appendix C. Explicit witness of causal nonseparability for $W^{\text{act.}}$

In this appendix we provide an explicit witness of causal nonseparability for the process matrix $W^{\text{act.}}$ introduced in section 3.3.

According to equation (G5) in appendix G.2 below, in the tripartite scenario in which $W^{\text{act.}}$ is defined, where $d_{C_O} = 1$, the cone of causal witnesses can be characterised as

$$\mathcal{S} = \{S \in A_{IO}B_{IO}C_I | S = S_A^{(+)} + S_A^{(1)} + S_A^{(2)} \text{ with } S_A^{(+)} \geqslant 0,\ {}_{[1-A_O]B_{IO}C_I}S_A^{(1)} = S_A^{(1)},\ {}_{[1-B_O]C_I}S_A^{(2)} = S_A^{(2)},$$
$$S = S_B^{(+)} + S_B^{(1)} + S_B^{(2)} \text{ with } S_B^{(+)} \geqslant 0,\ {}_{[1-B_O]A_{IO}C_I}S_B^{(1)} = S_B^{(1)},\ {}_{[1-A_O]C_I}S_B^{(2)} = S_B^{(2)}\ \}. \quad \text{(C1)}$$

Using the approach of section 4.4, we obtained the following causal witness for $W^{\text{act.}}$, written, as in the definition (10) of $W^{\text{act.}}$, in the order $C_I A_I B_I A_O B_O$:

$$S^{\text{act.}} = \frac{1}{4}\Big[\mathbb{1}(\mathbb{1}\mathbb{1} - \hat{z}\hat{z})(\mathbb{1}\mathbb{1} - \hat{z}\hat{z}) - \frac{2}{3}\mathbb{1}(\hat{x}\hat{x} + \hat{y}\hat{y})(\mathbb{1}\hat{z} + \hat{z}\mathbb{1})$$
$$+ \frac{1}{\sqrt{3}}\hat{z}(\mathbb{1}\hat{z} - \hat{z}\mathbb{1})(\mathbb{1}\mathbb{1} - \hat{z}\hat{z}) + \frac{1}{\sqrt{3}}\hat{x}(\hat{x}\hat{y} - \hat{y}\hat{x})(\mathbb{1}\hat{z} - \hat{z}\mathbb{1}) + \frac{1}{3}\hat{y}(\hat{x}\hat{x} + \hat{y}\hat{y})(\mathbb{1}\hat{z} - \hat{z}\mathbb{1})\Big]. \quad \text{(C2)}$$

To see that $S^{\text{act.}}$ indeed defines a valid causal witness, one can verify that it admits decompositions as in equation (C1) above, with (still written in the same order) $S_A^{(1)} = 0$, $S_A^{(2)} = \frac{1}{4}[-\frac{4}{3}\mathbb{1}(\hat{x}\hat{x} + \hat{y}\hat{y})\mathbb{1}\hat{z}]$, $S_A^{(+)} = S^{\text{act.}} - S_A^{(2)}$ and similarly $S_B^{(1)} = 0$, $S_B^{(2)} = \frac{1}{4}[-\frac{4}{3}\mathbb{1}(\hat{x}\hat{x} + \hat{y}\hat{y})\hat{z}\mathbb{1}]$, $S_B^{(+)} = S^{\text{act.}} - S_B^{(2)}$. One can easily check that all constraints in equation (C1) are satisfied.

With $S^{\text{act.}}$ thus defined, one finds $\text{Tr}[S^{\text{act.}} \cdot W^{\text{act.}}] = -(\frac{4}{\sqrt{3}} - 2) < 0$, which proves that $W^{\text{act.}}$ is indeed causally nonseparable according to our definition 5—or equivalently, to Araújo *et al*'s definition 2, or 'extensibly causally nonseparable' according to definition 4—even though, as proven in section 3.3, it is causally separable according to OG's definition 3.

Since the causal witness $S^{\text{act.}}$ above was obtained with the SDP optimisation technique described in section 4.4, it allows us to determine the robustness of $W^{\text{act.}}$ to white noise. From equation (39) we thus find that its random robustness is $r^* = -\text{Tr}[S^{\text{act.}} \cdot W^{\text{act.}}] = \frac{4}{\sqrt{3}} - 2 \simeq 0.31$.

'Activation' of causal nonseparability with $W^{\text{act.}}$

It is instructive to see explicitly how causal nonseparability can be 'activated' by attaching an entangled ancillary state to $W^{\text{act.}}$.

Recall that $W^{\text{act.}}$ is compatible with party C acting first. As shown in section 3.3, it is such that for any CP map (or POVM element) $M_{\tilde{c}}$ applied by C, the conditional bipartite process matrix $(W^{\text{act.}})_{|M_{\tilde{c}}}$ is causally separable. This is precisely why $W^{\text{act.}}$ is considered to be causally separable according to OG's definition 3.

Consider now attaching an ancillary maximally entangled state $\rho = |\Phi^+\rangle\langle\Phi^+|^{A_{I'}/C_{I'}}$, shared by A and C with dimensions $d_{A_{I'}} = d_{C_{I'}} = d_{C_I}$, and letting C project his two incoming systems onto $|\Phi^+\rangle\langle\Phi^+|^{C_I/C_{I'}}$. The resulting conditional process matrix $(W^{\text{act.}} \otimes \rho)_{|M_C = |\Phi^+\rangle\langle\Phi^+|}$ shared by A and B is then (up to normalisation) $(W^{\text{act.}})^{C_I \to A_{I'}}$, i.e. it is formally represented by the same matrix as $W^{\text{act.}}$, equation (10), with party C's incoming system now given to party A (see lemma B1 of appendix B). One can verify that $(W^{\text{act.}} \otimes \rho)_{|M_C = |\Phi^+\rangle\langle\Phi^+|}$ thus obtained is causally nonseparable by constructing a (bipartite) causal witness using, for instance, the explicit characterisation of equation (G3) below, in a similar way to what we did for $W^{\text{act.}}$ above.

Note, however, that this argument is not sufficient to conclude that $W^{\text{act.}}$ is (extensibly) causally nonseparable: one indeed needs to prove that *for any* possible decomposition of the form $W^{\text{act.}} \otimes \rho = W_{(A)}^{\rho} + W_{(B)}^{\rho} + W_{(C)}^{\rho}$ with each $W_{(X)}^{\rho}$ compatible with party X acting first, there exist CP maps M_A, M_B, or M_C, that make either $(W_{(A)}^{\rho})_{|M_A}$, $(W_{(B)}^{\rho})_{|M_B}$ or $(W_{(C)}^{\rho})_{|M_C}$ causally nonseparable[22]. Our construction of a

[22] Indeed, a process matrix compatible with C first (in short, of the form $W = W_{(C)}$), and such that for some CP map M_C the conditional process matrix $W_{|M_C}$ is causally nonseparable, may still be causally separable if it also has another, causally separable, decomposition of the form $W = W_{(A)}{}' + W_{(B)}{}' + W_{(C)}{}'$. An example is for instance $W = W_0 \otimes |\Phi^+\rangle\langle\Phi^+|^{A_{I'}/C_{I'}}$ with (written again in the order $C_I A_I B_I A_O B_O$)

$$W_0 = \frac{1}{8}\Big[(\mathbb{1}\mathbb{1}\mathbb{1} + \mathbb{1}\hat{z}\hat{z} + \hat{z}\mathbb{1}\hat{z} + \hat{z}\hat{z}\mathbb{1})\mathbb{1}\mathbb{1} + \frac{1}{\sqrt{2}}(\hat{x}\hat{x}\hat{x} - \hat{x}\hat{y}\hat{y} - \hat{y}\hat{x}\hat{y} - \hat{y}\hat{y}\hat{x})\hat{x}\mathbb{1} + \frac{1}{\sqrt{2}}(\hat{y}\hat{y}\hat{y} - \hat{y}\hat{x}\hat{x} - \hat{x}\hat{y}\hat{x} - \hat{x}\hat{x}\hat{y})\mathbb{1}\hat{y}\Big]. \quad \text{(C3)}$$

One can check that $W \in \mathcal{L}^{C \prec \{A,B\}}$ and that with $M_C = |\Phi^+\rangle\langle\Phi^+|^{C_I/C_{I'}}$, the bipartite conditional process matrix $W_{|M_C}$ is causally nonseparable—even though W is also compatible with the fixed order $A \prec B \prec C$ (and is hence causally separable). (Here the ancillary entangled state attached to W_0 and the CP map M_C allow party C to 'teleport' their incoming system in W_0 to A; the same observation holds if C teleports his system to B instead.) A similar observation can be made at the level of correlations: a tripartite correlation $P(a, b, c|x, y, z)$ compatible with C first and such that the bipartite conditional correlation $P_{z,c}(a, b|x, y) \coloneqq P(a, b|x, y, z, c)$ is noncausal for some z, c may in general still be causal. An example with binary inputs and outputs 0, 1 is $P(a, b, c|x, y, z) \coloneqq \frac{1}{2}\delta_{b,x}\delta_{c,a\oplus y}$, which is indeed compatible with C first (as $P(c|x, y, z) = \frac{1}{2}$ does not depend on x, y) and is such that conditioned on C's output $c = 0$, the resulting conditional bipartite correlation $P_{z,c=0}(a, b|x, y) = \delta_{b,x}\delta_{a,y}$ shared by A and B is noncausal (it violates the 'Guess Your Neighbour's Input' inequality [34] maximally). Nevertheless, P is clearly also compatible with the fixed order $A \prec B \prec C$, and is hence causal. Note that the argument given by OG in [17] to show activation of causal nonseparability consisted precisely in proving that, after attaching an ancillary state, the correlations generated by a given tripartite process matrix were compatible with C first and such that the bipartite conditional correlation $P_{z,c}(a, b|x, y)$ was noncausal. In that case, however, C was performing a deterministic operation (i.e. c could only take a single fixed value), so this argument was enough, in their case, to prove that the tripartite correlation under consideration was indeed noncausal [22, 46].

causal witness for $W^{\text{act.}}$ confirms nonetheless that this must indeed be the case, which allows us conclude, using OG's terminology, that the entangled ancillary state ρ introduced here indeed 'activates' the causal nonseparability of $W^{\text{act.}}$.

Appendix D. Equivalence between Oreshkov and Giarmatzi's extensible causal (non) separability and our definition of causal (non)separability

In this appendix we prove that OG's definition 4 of extensible causal (non)separability and our definition 5 of multipartite causal (non)separability are equivalent.

Proof. Let W be an N-partite process matrix that is causally separable as per our definition 5. The conditional $(N-1)$-partite process matrices $(W^{\rho}_{(k)})_{|M_k} = \text{Tr}_k[M_k \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W^{\rho}_{(k)}]$ in definition 5 are again causally separable (as per our definition), and thus fulfil in particular definition 3. Therefore, $W \otimes \rho$ is causally separable as per OG's definition 3 (OG-CS) for any $A_{I'}^{\mathcal{N}}$ and any ancillary quantum state $\rho \in A_{I'}^{\mathcal{N}}$. That is, W is extensibly causally separable as per OG's definition 4 (OG-ECS).

The proof of the converse is more involved. The idea is to consider, for an N-partite OG-ECS process matrix W and two ancillary quantum states ρ' and ρ'', the extended process matrices $W \otimes \rho'$ and $W \otimes \rho' \otimes \rho''$, which are both OG-CS. By comparing the corresponding decompositions we will show that the conditional $(N-1)$-partite process matrices obtained from the decomposition of $W \otimes \rho'$ are not only OG-CS, but also OG-ECS. From there, one can conclude by induction that W then also satisfies our definition 5.

The difficulty here is that the causally separable decomposition of $W \otimes \rho$ (for $\rho = \rho'$ or $\rho = \rho' \otimes \rho''$ in our case here) depends, *a priori*, on ρ. The following proposition states, however, that there exists a decomposition of W that provides a unique causally separable decomposition of $W \otimes \rho$ for any ρ.

Proposition D1. *Any* N*-partite extensibly causally separable (OG-ECS) process matrix W, as per definition 4, can be decomposed as*

$$W = \sum_{k\in\mathcal{N}} q_k \, W_{(k)}, \tag{D1}$$

with $q_k \geqslant 0$, $\sum_k q_k = 1$, and where for each k, $W_{(k)}$ is a process matrix compatible with party A_k acting first, and is such that for any extension $A_{I'}^{\mathcal{N}}$, any ancillary quantum state $\rho \in A_{I'}^{\mathcal{N}}$ and any possible CP map $M_k \in A^k_{II'O}$ applied by party A_k, the conditional $(N-1)$-partite process matrix $(W_{(k)} \otimes \rho)_{|M_k} \coloneqq \text{Tr}_k[M_k \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W_{(k)} \otimes \rho]$ is causally separable (OG-CS) as per definition 3.

Proof. Consider an N-partite OG-ECS process matrix W. By definition 4, for any extension $A_{I'}^{\mathcal{N}}$ and any state $\rho \in A_{I'}^{\mathcal{N}}$, the extended N-partite process matrix $W \otimes \rho$ must have a decomposition of the form

$$W \otimes \rho = \sum_{k\in\mathcal{N}} q_k \, W^{\rho}_{(k)}, \tag{D2}$$

where for each k, $W^{\rho}_{(k)}$ is a process matrix compatible with party A_k acting first, and such that whatever that party does, the resulting conditional process matrix for the other $(N-1)$ parties is OG-CS.

By an argument similar to that of proposition B1, it is easy to see that $W^{\rho}_{(k)}$ can without loss of generality be taken to be of the form $W^{\rho}_{(k)} = W_{(k)} \otimes \rho$. We emphasise again that the convex decomposition of $W = \sum_{k\in\mathcal{N}} q_k \, W_{(k)}$ that then follows from equation (D2) could *a priori* depend on the ancillary state ρ. We will however now show that for all extensions and ancillary quantum states one can choose the same decomposition of W.

First, note that for any finite set of extensions $\{A_{I_1}^{\mathcal{N}}, A_{I_2}^{\mathcal{N}}, \ldots, A_{I_n}^{\mathcal{N}}\}$ and ancillary quantum states $\{\rho_1 \in A_{I_1}^{\mathcal{N}}, \rho_2 \in A_{I_2}^{\mathcal{N}}, \ldots, \rho_n \in A_{I_n}^{\mathcal{N}}\}$ under consideration, one can indeed choose the same decomposition—consider the ancillary state $\rho_1 \otimes \cdots \otimes \rho_n \in A_{I_1}^{\mathcal{N}} \otimes \cdots \otimes A_{I_n}^{\mathcal{N}}$, and the corresponding decomposition

$$W \otimes \rho_1 \otimes \cdots \otimes \rho_n = \sum_{k\in\mathcal{N}} q_k \, W_{(k)} \otimes \rho_1 \otimes \cdots \otimes \rho_n, \tag{D3}$$

with each $W_{(k)} \otimes \rho_1 \otimes \cdots \otimes \rho_n$—and therefore, each $W_{(k)}$—a process matrix compatible with party A_k acting first, and such that for any operation M_k applied by A_k the resulting conditional process matrix $(W_{(k)} \otimes \rho_1 \otimes \cdots \otimes \rho_n)_{|M_k}$ for the other $(N-1)$ parties is OG-CS. Proposition B2 now implies that these conditional process matrices remain causally separable when tracing out all but one ancillary states in the tensor product. Therefore, the decomposition of W obtained from equation (D3) can be chosen for any of the individual $\rho_j \in A_{I_j}^{\mathcal{N}}$.

Next, one uses the following result from basic topology:

Theorem 2.36 in. [47] *If* $\{K_\alpha\}$ *is a collection of compact subsets of a metric space* X *such that the intersection of every finite subcollection of* $\{K_\alpha\}$ *is nonempty, then* $\bigcap K_\alpha$ *is nonempty.*

Here, let the index set be the set of all possible ancillary quantum states (of any dimension), and the set K_ρ, indexed by some quantum state ρ, be the set of possible causally separable decompositions of W corresponding to the ancillary state ρ. The finite intersection property follows from the observation above—for any finite set of ancillary states $\{\rho_1, \ldots, \rho_n\}$, there exists a common decomposition, that is, the intersection $K_{\rho_1} \cap \cdots \cap K_{\rho_n}$ is nonempty. As the conditions of the above theorem are satisfied[23], it guarantees that the intersection $\bigcap K_\rho$ over all quantum states ρ is nonempty. That is, there exists indeed a convex decomposition of W,

$$W = \sum_{k \in \mathcal{N}} q_k \, W_{(k)}, \tag{D5}$$

with $q_k \geqslant 0$, $\sum_k q_k = 1$, and where for each k, $W_{(k)}$ is a process matrix compatible with party A_k acting first, and is such that for any extension $A_{I'}^{\mathcal{N}}$, any ancillary quantum state $\rho \in A_{I'}^{\mathcal{N}}$ and any possible CP map $M_k \in A_{II'O}^{k}$ applied by party A_k, the conditional $(N-1)$-partite process matrix $(W_{(k)} \otimes \rho)_{|M_k}$ is OG-CS. □

One can now prove by induction that any OG-ECS process matrix is causally separable according to our definition 5. In the single-partite case ($N = 1$), the claim is trivial. Suppose, for $N \geqslant 2$, that the claim holds true in the $(N-1)$-partite case. Let then W be an N-partite OG-ECS process matrix. According to proposition B8, W has a decomposition of the form (D1), such that for any k, for any arbitrary extensions $A_{I'}^{\mathcal{N}}$, $A_{I''}^{\mathcal{N}\setminus k}$, any ancillary quantum states $\rho' \in A_{I'}^{\mathcal{N}}$, $\rho'' \in A_{I''}^{\mathcal{N}\setminus k}$, and any CP map $M_k \in A_{II'O}^{k}$ applied by party A_k, the conditional $(N-1)$-partite process matrices $(W_{(k)} \otimes \rho')_{|M_k}$ and $(W_{(k)} \otimes \rho' \otimes \rho'')_{|M_k} = (W_{(k)} \otimes \rho')_{|M_k} \otimes \rho''$ are OG-CS. That is, for any $A_{I'}^{\mathcal{N}}$, ρ' and M_k, $(W_{(k)} \otimes \rho')_{|M_k}$ is OG-ECS—and therefore, by the induction hypothesis, it is causally separable according to our definition 5. Summing up, we thus have, for any $A_{I'}^{\mathcal{N}}$ and any $\rho' \in A_{I'}^{\mathcal{N}}$, a decomposition of the form $W \otimes \rho' = \sum_k q_k W_{(k)} \otimes \rho'$ such that for any $M_k \in A_{II'O}^{k}$, $(W_{(k)} \otimes \rho')_{|M_k}$ is causally separable. This means that W itself is causally separable as per our definition 5, which concludes the proof. □

Appendix E. Causally separable process matrices can only generate causal correlations

Here we show explicitly that causally separable process matrices, according to our definition 5, can only generate so-called 'causal correlations' (even when attaching ancillary entangled states).

Let us first recall the definition of N-partite causal correlations given in [22] (which is equivalent to that first introduced in [17]):

Definition E1 (*N*-partite causal correlations). For $N = 1$, any correlation $P(a_1|x_1)$ is causal. For $N \geqslant 2$, an N-partite correlation $P(\vec{a}|\vec{x})$ is said to be *causal* if and only if it can be decomposed as

$$P(\vec{a}|\vec{x}) = \sum_{k \in \mathcal{N}} q_k \, P_k(a_k|x_k) \, P_{k,x_k,a_k}(\vec{a}_{\mathcal{N}\setminus k}|\vec{x}_{\mathcal{N}\setminus k}), \tag{E1}$$

with $q_k \geqslant 0$, $\sum_k q_k = 1$, where (for each k) $P_k(a_k|x_k)$ is a single-partite (and hence causal) correlation and (for each k, x_k, a_k) $P_{k,x_k,a_k}(\vec{a}_{\mathcal{N}\setminus k}|\vec{x}_{\mathcal{N}\setminus k})$ is a causal $(N-1)$-partite correlation.

By this definition, for $N = 1$, any correlation—and in particular, any correlation generated by a (trivially) causally separable single-partite process matrix—is causal.

Assume, for $N > 1$, that any correlation generated by a $(N-1)$-partite causally separable process matrix is causal. Consider then an N-partite process matrix $W \in A_{IO}^{\mathcal{N}}$, some ancillary state $\rho \in A_{I'}^{\mathcal{N}}$, and some CP maps $M_{a_k|x_k} \in A_{II'O}^{\mathcal{N}}$ (for any k, x_k, a_k), which all together generate the probability distribution

[23] More precisely, let X be the space of N-tuples of Hermitian matrices $\vec{W} = (W_{(1)}, \ldots, W_{(N)})$, equipped with the standard Euclidean metric, and let the sets K_ρ be defined as

$$K_\rho := \left\{ (W_{(1)}, \ldots, W_{(N)}) \;\middle|\; \left(\sum_{k=1}^{N} W_{(k)} = W\right) \text{ and } (W_{(k)} \in \mathcal{L}^{A_k \prec (\mathcal{N}\setminus A_k)} \;\forall k) \text{ and } (W_{(k)} \geqslant 0 \;\forall k) \text{ and } ((W_{(k)} \otimes \rho)_{|M_k} \text{ is OG} - \text{CS} \;\forall k, M_k) \right\}. \tag{D4}$$

It follows from the positivity of the $W_{(k)}$'s and the normalisation of W that the sets K_ρ are bounded. One can further easily convince oneself that the sets characterised by the four individual conditions in equation (D4) are closed, and thus, as it is the intersection of closed sets, that K_ρ is closed. The sets K_ρ being bounded and closed, it follows from the Heine-Borel theorem that they are compact, as required for theorem 2.36 in [47] to be applicable here.

$$P(\vec{a}|\vec{x}) = \mathrm{Tr}[M_{a_1|x_1} \otimes \cdots \otimes M_{a_N|x_N} \cdot W \otimes \rho], \tag{E2}$$

as in equation (4). Assuming that W is causally separable, according to definition 5 $W \otimes \rho$ can be decomposed as in equation (13), which allows us to write

$$P(\vec{a}|\vec{x}) = \sum_{k\in\mathcal{N}} q_k \,\mathrm{Tr}[M_{a_1|x_1} \otimes \cdots \otimes M_{a_N|x_N} \cdot W^{\rho}_{(k)}]. \tag{E3}$$

Here $W^{\rho}_{(k)}$ is compatible with party A_k acting first, so that for any set of CPTP maps $M_{x_{k'}}$ with $k' \neq k$,

$$\mathrm{Tr}[M_{a_k|x_k} \bigotimes_{k'\in\mathcal{N}\setminus k} M_{x_{k'}} \cdot W^{\rho}_{(k)}] = \mathrm{Tr}[M_{a_k|x_k} \bigotimes_{k'\in\mathcal{N}\setminus k} \frac{\mathbb{1}}{d_{A^{k'}_O}} \cdot W^{\rho}_{(k)}] = \frac{1}{d_{A_O^{\mathcal{N}\setminus k}}} \mathrm{Tr}[(W^{\rho}_{(k)})_{|M_{a_k|x_k}}], \tag{E4}$$

which does not depend on the choice of CPTP maps $M_{x_{k'}}$, and defines a probability distribution $P_k(a_k|x_k)$ for party A_k.

The conditional process matrix $(W^{\rho}_{(k)})_{|M_{a_k|x_k}}$ for parties in $\mathcal{N}\setminus k$ can be renormalised (when nonzero) by defining $(\widetilde{W}^{\rho}_{(k)})_{|M_{a_k|x_k}} := \frac{1}{P_k(a_k\,|\,x_k)}(W^{\rho}_{(k)})_{|M_{a_k|x_k}}$, so that $(\widetilde{W}^{\rho}_{(k)})_{|M_{a_k|x_k}}$ is now a properly normalised process matrix (according to equation (E4) above that defines $P_k(a_k|x_k)$, we indeed have $\mathrm{Tr}[(\widetilde{W}^{\rho}_{(k)})_{|M_{a_k|x_k}}] = d_{A_O^{\mathcal{N}\setminus k}}$, as required by equation (A3)). We can then write equation (E3) as

$$P(\vec{a}|\vec{x}) = \sum_{k\in\mathcal{N}} q_k \,\mathrm{Tr}\left[\bigotimes_{k'\in\mathcal{N}\setminus k} M_{a_{k'}|x_{k'}} \cdot (W^{\rho}_{(k)})_{|M_{a_k|x_k}}\right] = \sum_{k\in\mathcal{N}} q_k \, P_k(a_k|x_k)\, P_{k,x_k,a_k}(\vec{a}_{\mathcal{N}\setminus k}|\vec{x}_{\mathcal{N}\setminus k}) \tag{E5}$$

with

$$P_{k,x_k,a_k}(\vec{a}_{\mathcal{N}\setminus k}|\vec{x}_{\mathcal{N}\setminus k}) := \mathrm{Tr}\left[\bigotimes_{k'\in\mathcal{N}\setminus k} M_{a_{k'}|x_{k'}} \cdot (\widetilde{W}^{\rho}_{(k)})_{|M_{a_k|x_k}}\right]. \tag{E6}$$

Now, by assumption and according to definition 5 $(\widetilde{W}^{\rho}_{(k)})_{|M_{a_k|x_k}}$ must a causally separable process matrix; by the induction hypothesis it can only generate causal correlations, which implies that $P_{k,x_k,a_k}(\vec{a}_{\mathcal{N}\setminus k}|\vec{x}_{\mathcal{N}\setminus k})$ is causal. Equation (E5) thus provides a causal decomposition of $P(\vec{a}|\vec{x})$ as in equation (E1) of definition E1, which proves that the correlation $P(\vec{a}|\vec{x})$ obtained from the N-partite causally separable process matrix W is causal, and which, by induction, concludes the proof.

Appendix F. Relationship between our necessary and sufficient conditions for causal separability

F.1. A necessary but not sufficient condition

In our recursive necessary condition of proposition 4 for general multipartite causal separability, we require the $(N-1)$-partite process matrices $W^{A^k_{IO}\to A^{k'}_{I'}}_{(k)}$ to be causally separable for each $k' \neq k$. In the tripartite case, it is not necessary to impose this explicitly, since considering the teleportation of A_k's systems to some arbitrary $A_{k'}$ yields necessary conditions that already coincide with the sufficient conditions for tripartite causal separability (see the proof in appendix B.1.2). In the general case, however, considering the teleportation to just one or some of the parties yields weaker necessary conditions that may not be sufficient. In this appendix we present an explicit fourpartite example.

We consider the fourpartite scenario where A has a trivial incoming space ($d_{A_I} = 1$) and D has a trivial outgoing space ($d_{D_O} = 1$), and define the following matrix in $A_O \otimes B_{IO} \otimes C_{IO} \otimes D_I$:

$$W^{\mathrm{gap}} := \frac{1}{8}\left[\mathbb{1}^{\otimes 6} + \frac{1}{\sqrt{2}}\,\hat{z}(\mathbb{1}\hat{z}\hat{z}\mathbb{1} + \hat{z}\mathbb{1}\hat{x}\hat{z})\mathbb{1}\right]. \tag{F1}$$

It is easy to verify that W^{gap} satisfies equation (18) for $A_k = A$, i.e. that it is a valid process matrix compatible with party A acting first (note its similarity with the original process matrix of Oreshkov, Costa and Brukner [2]). Furthermore, it satisfies ${}_{[1-B_O]A_OC_{IO}D_I}W^{\mathrm{gap}} = {}_{[1-C_O]A_OD_I}W^{\mathrm{gap}} = {}_{[1-D_O]}W^{\mathrm{gap}} = 0$ (as well as ${}_{[1-C_O]A_OB_{IO}D_I}W^{\mathrm{gap}} = {}_{[1-B_O]A_OD_I}W^{\mathrm{gap}} = 0$). Thus, W^{gap} can be decomposed as in equations (27)–(28) with $Y = D$ and a single term in the decomposition, $W^{\mathrm{gap}} = W_{(A)} = W^{[A\to D]}_{(A,B)} = W^{[A\to D]}_{(A,B,C,D)}$ (or $W^{\mathrm{gap}} = W_{(A)} = W^{[A\to D]}_{(A,C)} = W^{[A\to D]}_{(A,C,B,D)}$) satisfying equation (29). In other words, the tripartite conditional process matrix that we obtain by teleporting A_O to $D_{I'}$ is causally separable (it is compatible with both fixed causal orders $B \prec C \prec D$ and $C \prec B \prec D$).

However, this is not the case when teleporting A_O to $B_{I'}$, or to $C_{I'}$. W^{gap} indeed cannot be decomposed as in equation (28) with $Y = B$ or C, and is thus causally nonseparable. This can be certified by the causal witness (obtained as described in section 4.4, with the characterisation of equation (G11) in appendix G.2)

$$S^{\text{gap}} := \frac{1}{8}[\mathbb{1}^{\otimes 6} - \hat{z}(\mathbb{1}\hat{z}\hat{z}\mathbb{1} + \hat{z}\mathbb{1}\hat{x}\hat{z})\mathbb{1}], \tag{F2}$$

for which we obtain $\text{Tr}[W^{\text{gap}} \cdot S^{\text{gap}}] = 1 - \sqrt{2} < 0$.

This shows that, in the general multipartite case, there is indeed a gap between the necessary conditions obtained by teleporting to just some of the parties and those obtained by teleporting to each of the parties, and that the former are not sufficient.

(Note that in the example above D_I did not play any role, as we always had $\mathbb{1}^{D_I}$ in all terms We could in fact consider the case where D_I is also trivial, $d_{D_I} = 1$. We kept here a nontrivial system D_I to clarify the fact that W^{gap} was defined in a fourpartite scenario, and that party D does play a role in the argument.)

F.2. Numerically investigating the (in)equivalence of our necessary and sufficient conditions

In order to investigate whether the (full version of the) necessary condition in proposition 4 and the sufficient condition in proposition 5 differ in general, we conducted numerical testing to see whether we could find process matrices contained in the cone $\mathcal{W}_+^{\text{sep}}$ but not in $\mathcal{W}_-^{\text{sep}}$ (i.e. the outer and inner approximations of $\mathcal{W}^{\text{sep}}$ arising from the necessary and sufficient conditions, respectively). To this end, we considered the following general approach: we first generated a large number of random process matrices. For each process matrix W, we then solved the primal SDP optimisation problem (37) over the cones $\mathcal{W}_\pm^{\text{sep}}$ to obtain the corresponding random robustnesses $r_\pm^*$. If we were to find $r_+^* \neq r_-^*$ (up to numerical error; note that, since $\mathcal{W}_-^{\text{sep}} \subseteq \mathcal{W}_+^{\text{sep}}$, one always has $r_+^* \leqslant r_-^*$), this would imply the cones differ since one would have $W + r_+^*\mathbb{1}^\circ \in \mathcal{W}_+^{\text{sep}}$ but $W + r_+^*\mathbb{1}^\circ \notin W_-^{\text{sep}}$.

The size of the SDP problems associated with finding the random robustness of a process matrix meant that we could not solve these problems for the 'complete' fourpartite scenario with qubit incoming and outgoing spaces for each party (recall that, for three parties, the conditions are already known to coincide). We therefore considered the restricted scenario in which $d_{A_I} = 1$ while the remaining Hilbert spaces are two-dimensional, so that W is thus (128×128)-dimensional. We note that in any simpler scenario, the necessary and sufficient conditions can be be proven to coincide, making this the simplest case of interest. Indeed, in appendix B.3 we already showed that they coincide if one of the four parties has a trivial outgoing space. If, on the other hand, a second party were to have a trivial incoming space (e.g. $d_{B_I} = 1$), it is not difficult to show they again coincide by writing explicitly the necessary and sufficient conditions of propositions 4 and 5, by using the fact that they simplify to proposition B4 in a tripartite case where (at least) one party has a trivial incoming space, and by using the linearity of the subspaces appearing in the constraints. We leave the explicit proof of this as an exercise for the reader.

To generate random process matrices, one could follow the hit-and-run approach of [43]. Although this approach is guaranteed to sample process matrices uniformly, the high dimensionality of the space of valid process matrices (in this scenario it is 7597-dimensional) renders this approach intractable. Instead, forgoing uniformity, we generated matrices by randomly sampling Hermitian positive semidefinite matrices, projecting them onto the space $\mathcal{L}^\vee$ of valid process matrices before adding white noise (i.e. $\mathbb{1}^\circ$) until the resulting matrix was again positive semidefinite.

We solved the SDP optimisation problems for the necessary and sufficient conditions for approximately 1000 randomly generated process matrices (including several hundreds in which an additional constraint, namely the symmetry of W between permutations of the parties B, C and D, was imposed). These numerical tests failed to provide any potential counterexamples: in all cases we found $r_+^* = r_-^*$ up to numerical precision.

However, since the space of valid process matrices is so high-dimensional and our sampling method non-uniform, we do not believe that our results on this number of samples provide enough evidence to reasonably conjecture that the necessary and sufficient conditions coincide in this scenario.

Appendix G. Construction of witnesses of causal nonseparability through SDP

In this appendix we give some further details relating to the construction of witnesses of causal nonseparability through SDP. Firstly, we discuss the duality of the two SDP problems given in section 4.4, showing that they are indeed dual and that the Strong Duality theorem is satisfied. We then give some additional details on how the characterisations of causal separability can be explicitly translated into SDP constraints in order to find witnesses in practice, giving some explicit examples that both illustrate this and, at the same time, allow the results in section 4.5 to be readily verified.

G.1. Duality of SDP problems

Since both the set of causally separable process matrices $\mathcal{W}^{\text{sep}}$ and its dual $\mathcal{S} = (\mathcal{W}^{\text{sep}})^*$ (or the inner and outer approximations $\mathcal{W}_\pm^{\text{sep}}$ of $\mathcal{W}^{\text{sep}}$ arising from propositions 4 and 5 and their respective duals $\mathcal{S}_\pm = (\mathcal{W}_\pm^{\text{sep}})^*$, see

section 4.4) are convex cones, the problems of minimising the amount of white noise that must be added to make a process matrix causally separable and finding the witness of causal separability with the most negative value for a given process matrix can be formulated as SDP problems as in equations (37) and (38), respectively. For these problems to be efficiently solvable with standard algorithmic techniques for SDP, however, one must show that they have no duality gap (i.e. no difference between the optimal values of an SDP problem and its dual). Here, we will show that equations (37) and (38) are indeed a primal-dual pair, and that the Strong Duality theorem holds [48], implying that that their optimal solutions indeed coincide and can therefore be efficiently obtained. This shows, in particular, that the solution to the SDP problem (38) is the optimal witness with respect to the random robustness.

Reference [14] showed the duality of two variations of the SDP problems (37) and (38) in the bipartite case: rather than consider the robustness to white noise of a process matrix, they considered the robustness of mixing a given W with any valid process matrix. The optimal solutions to the corresponding SDP problems give the generalised robustness of W. Nevertheless, their approach to proving duality, and the applicability of the Strong Duality Theorem, is easily adapted to (and even simpler for) the random robustness, and the bipartite and some restricted tripartite versions of equations (37) and (38) were already given in [15]. The same approach can be used in the more general multipartite case to show that these problems (considering the cones $\mathcal{W}^{\text{sep}}$ or $\mathcal{W}^{\text{sep}}_{\pm}$) satisfy the required properties. Rather than repeating these (somewhat technical and lengthy) arguments, we instead refer the reader to appendix E of [14] and prove explicitly only the main technical lemma needed to generalise their approach.

First, as noted already in [14, 15], it is sufficient just to consider the restriction $\mathcal{S}^{\mathcal{N}} := \mathcal{S} \cap \mathcal{L}^{\mathcal{N}}$ of witnesses in $\mathcal{L}^{\mathcal{N}}$. Indeed, for any $S^{\perp}$ in the orthogonal subspace $(\mathcal{L}^{\mathcal{N}})^{\perp}$ of $\mathcal{L}^{\mathcal{N}}$ and any process matrix W one has $\text{Tr}[S^{\perp} \cdot W] = 0$, and thus for any $S \in \mathcal{S}$ there exists $S' \in \mathcal{S}^{\mathcal{N}}$ such that $\text{Tr}[S \cdot W] = \text{Tr}[S' \cdot W]$ for all $W \in \mathcal{L}^{\mathcal{N}}$. The formulations given in equations (37) and (38) are only formally dual when $\mathcal{W}^{\text{sep}}$ and $\mathcal{S}$ are considered as subsets of the vector space $\mathcal{L}^{\mathcal{N}}$ [or when $\mathcal{S}$ is replaced by $\mathcal{S}^{\mathcal{N}}$ in equation (38)]. However, the fact that the restriction to $\mathcal{S}^{\mathcal{N}}$ does not change the optimal value of the problem ensures that the optimal solutions coincide in the more general formulation.

The primary element of the proof in [14] which needs to be generalised beyond two parties is the need to show that $\mathcal{W}^{\text{sep}}$ has a nonempty interior (within $\mathcal{L}^{\mathcal{N}}$, see their lemma 7; we also need to check that it is pointed, which is trivial). To this end, it is sufficient to show that the white noise process matrix $\mathbb{1}^{\circ}$ is in the interior of $\mathcal{W}^{\text{sep}}$, i.e. that there exists $\varepsilon > 0$ such that for any $W \in \mathcal{L}^{\mathcal{N}}$ with $\|W\|_{\text{HS}} < \varepsilon$ (where $\|\cdot\|_{\text{HS}}$ is the Hilbert–Schmidt norm), one has $\mathbb{1}^{\circ} + W \in \mathcal{W}^{\text{sep}}$.

Recalling from appendix A.4 the characterisation of $\mathcal{L}^{\mathcal{N}}$ in terms of 'allowed' terms in a Hilbert–Schmidt basis decomposition, let us first note that any allowed Hilbert–Schmidt term T_k which contains $\cdots \sigma_{\mu_k}^{A_I^k} \mathbb{1}^{A_O^k} \cdots$ (with $\sigma_{\mu_k}^{A_I^k} \neq \mathbb{1}$) for some $k \in \mathcal{N}$ is compatible with any fixed causal order where party A_k comes last—i.e. that $T_k \in \mathcal{L}^{A_{\pi_k(1)} \prec \cdots \prec A_{\pi_k(N-1)} \prec A_k}$ for any permutation π_k of parties such that $\pi_k(N) = k$ (the same also trivially holds for the allowed identity term $\mathbb{1}^{\mathcal{N}}$). Indeed, ${}_{[1-A_O^k]} T_k = 0$ and ${}_{A_{IO}^k} T_k = 0$, so that equation (19) holds for any such order. It follows that any $W \in \mathcal{L}^{\mathcal{N}}$ can be written as $W = \sum_{k=1}^{N} \Omega_k$, where each $\Omega_k \in \mathcal{L}^{A_{\pi_k(1)} \prec \cdots \prec A_{\pi_k(N-1)} \prec A_k}$ (for some arbitrary π_k for each k); furthermore, the terms Ω_k can be taken to be orthogonal, so that $\|W\|_{\text{HS}}^2 = \sum_{k=1}^{N} \|\Omega_k\|_{\text{HS}}^2$.

Note that the Ω_k's may not, in general, be positive semidefinite. Nevertheless, if we take W such that $\|W\|_{\text{HS}} < \varepsilon := \frac{1}{N d_I}$, with $d_I := \prod_{k \in \mathcal{N}} d_{A_I^k}$, then each $\|\Omega_k\| \leqslant \|\Omega_k\|_{\text{HS}} < \frac{1}{N d_I}$ (where $\|\cdot\|$ is now the spectral norm), so that $\frac{1}{N d_I} \mathbb{1} + \Omega_k \geqslant 0$. For any such W, we thus obtain a decomposition

$$\mathbb{1}^{\circ} + W = \sum_{k=1}^{N} \left(\frac{1}{N d_I} \mathbb{1} + \Omega_k \right) \tag{G1}$$

with $\frac{1}{N d_I} \mathbb{1} + \Omega_k \in \mathcal{P} \cap \mathcal{L}^{A_{\pi_k(1)} \prec \cdots \prec A_{\pi_k(N-1)} \prec A_k}$, which proves that $\mathbb{1}^{\circ} + W$ is the sum of (valid) process matrices compatible with fixed causal orders, and hence is causally separable: $\mathbb{1}^{\circ} + W \in \mathcal{W}^{\text{sep}}$, as desired.

With this verified, the approach of [14] can be applied, with the appropriate modifications for the random robustness[24], to show that the required duality indeed holds and that the conditions of the Strong Duality theorem are satisfied.

G.2. Explicit SDP constraints and example constructions

In order to characterise $\mathcal{S}$ more explicitly for a given scenario, as well as to solve both the primal and dual SDP problem using convex optimisation algorithms [49], it is helpful to write $\mathcal{W}^{\text{sep}}$ explicitly as intersections and Minkowski sums of convex cones corresponding to individual constraints on causally separable process

[24] Namely, one can change equations (E.3)–(E.7) in [14] to $E = \mathcal{L}^{\mathcal{N}}$, $\mathcal{K} = \mathcal{W}^{\text{sep}}$, $\mathcal{L} = \{r\mathbb{1}^{\circ} | r \in \mathbb{R}\}$, $b = W$ and $c = \mathbb{1} / \prod_{k \in \mathcal{N}} d_{A_O^k}$ (using their notations for E, $\mathcal{K}$, $\mathcal{L}$, b, c) and then adapt the proof accordingly.

matrices. The duality relations (35) can then be exploited to describe $\mathcal{S}$. Here we give some examples to illustrate this procedure.

The simplest example is the bipartite scenario. From the definition in equation (3) we see that $\mathcal{W}^{\text{sep}} = \mathcal{W}^{A\prec B} + \mathcal{W}^{B\prec A}$, where $\mathcal{W}^{A\prec B} = \mathcal{P} \cap \mathcal{L}^{A\prec B}$ and similarly for $\mathcal{W}^{B\prec A}$. Using equation (19) to write $\mathcal{L}^{A\prec B}$ and $\mathcal{L}^{B\prec A}$ in terms of spaces defined by individual linear constraints, or directly referring to proposition 2, we see that

$$\mathcal{W}^{\text{sep}} = \mathcal{P} \cap \mathcal{L}_{[1-A_O]B_{IO}} \cap \mathcal{L}_{[1-B_O]} + \mathcal{P} \cap \mathcal{L}_{[1-B_O]A_{IO}} \cap \mathcal{L}_{[1-A_O]}, \tag{G2}$$

with $\mathcal{L}_{[1-A_O]B_{IO}} \coloneqq \{W \in A_{IO}B_{IO}|_{[1-A_O]B_{IO}} W = 0\}$, $\mathcal{L}_{[1-B_O]} \coloneqq \{W \in A_{IO}B_{IO}|_{[1-B_O]} W = 0\}$, etc. It follows that

$$\mathcal{S} = (\mathcal{W}^{\text{sep}})^* = (\mathcal{P} + \mathcal{L}^{\perp}_{[1-A_O]B_{IO}} + \mathcal{L}^{\perp}_{[1-B_O]}) \cap (\mathcal{P} + \mathcal{L}^{\perp}_{[1-B_O]A_{IO}} + \mathcal{L}^{\perp}_{[1-A_O]}), \tag{G3}$$

where we used the fact that $\mathcal{P}$ is self-dual, and where $\mathcal{L}^{\perp}_{[1-A_O]B_{IO}} = \{S \in A_{IO}B_{IO}|_{[1-A_O]B_{IO}} S = S\}$ is the orthogonal subspace of $\mathcal{L}_{[1-A_O]B_{IO}}$, $\mathcal{L}^{\perp}_{[1-B_O]} = \{S \in A_{IO}B_{IO}|_{[1-B_O]} S = S\}$ is the orthogonal subspace of $\mathcal{L}_{[1-B_O]}$, etc.

Note that a slightly different, but equivalent, characterisation was given for the bipartite scenario in [14, 15]. Although their formulation is slightly simpler, we choose to give the above form as it shows more clearly the procedure of obtaining explicit SDP characterisations from the characterisations of causally separable process matrices given in the main text, and it generalises more directly to the multipartite scenario.

The next simplest case is the tripartite scenario with $d_{C_O} = 1$. In this case, causally separable process matrices are characterised by proposition B3, from which it follows that

$$\mathcal{W}^{\text{sep}} = \mathcal{P} \cap \mathcal{L}_{[1-A_O]B_{IO}C_I} \cap \mathcal{L}_{[1-B_O]C_I} + \mathcal{P} \cap \mathcal{L}_{[1-B_O]A_{IO}C_I} \cap \mathcal{L}_{[1-A_O]C_I}, \tag{G4}$$

with similar notations for $\mathcal{L}_{[1-A_O]B_{IO}C_I}$, $\mathcal{L}_{[1-B_O]C_I}$, etc as before. Similarly to the bipartite case, this leads to

$$\mathcal{S} = (\mathcal{P} + \mathcal{L}^{\perp}_{[1-A_O]B_{IO}C_I} + \mathcal{L}^{\perp}_{[1-B_O]C_I}) \cap (\mathcal{P} + \mathcal{L}^{\perp}_{[1-B_O]A_{IO}C_I} + \mathcal{L}^{\perp}_{[1-A_O]C_I}). \tag{G5}$$

We note again that two slightly different, but once again equivalent, characterisations were given in [14, 15] for this particular tripartite case.

In the tripartite scenario with $d_{A_I} = 1$ instead (as, e.g. in the example of 'activation of causal nonseparability' given by Oreshkov and Giarmatzi [17]), proposition B4 leads to

$$\mathcal{W}^{\text{sep}} = \mathcal{L}_{[1-A_O]B_{IO}C_{IO}} \cap [\mathcal{P} \cap \mathcal{L}_{[1-B_O]C_{IO}} \cap \mathcal{L}_{[1-C_O]} + \mathcal{P} \cap \mathcal{L}_{[1-C_O]B_{IO}} \cap \mathcal{L}_{[1-B_O]}]. \tag{G6}$$

It follows that

$$\mathcal{S} = \mathcal{L}^{\perp}_{[1-A_O]B_{IO}C_{IO}} + (\mathcal{P} + \mathcal{L}^{\perp}_{[1-B_O]C_{IO}} + \mathcal{L}^{\perp}_{[1-C_O]}) \cap (\mathcal{P} + \mathcal{L}^{\perp}_{[1-C_O]B_{IO}} + \mathcal{L}^{\perp}_{[1-B_O]}). \tag{G7}$$

In the general tripartite case, the characterisation of proposition 3 shows that we can write $\mathcal{W}^{\text{sep}}$ as

$$\begin{aligned}\mathcal{W}^{\text{sep}} = {} & \mathcal{L}_{[1-A_O]B_{IO}C_{IO}} \cap (\mathcal{P} \cap \mathcal{L}_{[1-B_O]C_{IO}} \cap \mathcal{L}_{[1-C_O]} + \mathcal{P} \cap \mathcal{L}_{[1-C_O]B_{IO}} \cap \mathcal{L}_{[1-B_O]}) \\ & + \mathcal{L}_{[1-B_O]A_{IO}C_{IO}} \cap (\mathcal{P} \cap \mathcal{L}_{[1-A_O]C_{IO}} \cap \mathcal{L}_{[1-C_O]} + \mathcal{P} \cap \mathcal{L}_{[1-C_O]A_{IO}} \cap \mathcal{L}_{[1-A_O]}) \\ & + \mathcal{L}_{[1-C_O]A_{IO}B_{IO}} \cap (\mathcal{P} \cap \mathcal{L}_{[1-A_O]B_{IO}} \cap \mathcal{L}_{[1-B_O]} + \mathcal{P} \cap \mathcal{L}_{[1-B_O]A_{IO}} \cap \mathcal{L}_{[1-A_O]}),\end{aligned} \tag{G8}$$

from which it follows that cone of witnesses is

$$\begin{aligned}\mathcal{S} = {} & (\mathcal{L}^{\perp}_{[1-A_O]B_{IO}C_{IO}} + (\mathcal{P} + \mathcal{L}^{\perp}_{[1-B_O]C_{IO}} + \mathcal{L}^{\perp}_{[1-C_O]}) \cap (\mathcal{P} + \mathcal{L}^{\perp}_{[1-C_O]B_{IO}} + \mathcal{L}^{\perp}_{[1-B_O]})) \\ & \cap (\mathcal{L}^{\perp}_{[1-B_O]A_{IO}C_{IO}} + (\mathcal{P} + \mathcal{L}^{\perp}_{[1-A_O]C_{IO}} + \mathcal{L}^{\perp}_{[1-C_O]}) \cap (\mathcal{P} + \mathcal{L}^{\perp}_{[1-C_O]A_{IO}} + \mathcal{L}^{\perp}_{[1-A_O]})) \\ & \cap (\mathcal{L}^{\perp}_{[1-C_O]A_{IO}B_{IO}} + (\mathcal{P} + \mathcal{L}^{\perp}_{[1-A_O]B_{IO}} + \mathcal{L}^{\perp}_{[1-B_O]}) \cap (\mathcal{P} + \mathcal{L}^{\perp}_{[1-B_O]A_{IO}} + \mathcal{L}^{\perp}_{[1-A_O]})).\end{aligned} \tag{G9}$$

Finally, in the fourpartite scenario with $d_{A_I} = d_{D_O} = 1$ (as, e.g. in the version of the quantum switch in equation (40)), proposition B7 leads to

$$\mathcal{W}^{\text{sep}} = \mathcal{L}_{[1-A_O]B_{IO}C_{IO}D_I} \cap [\mathcal{P} \cap \mathcal{L}_{[1-B_O]C_{IO}D_I} \cap \mathcal{L}_{[1-C_O]D_I} + \mathcal{P} \cap \mathcal{L}_{[1-C_O]B_{IO}D_I} \cap \mathcal{L}_{[1-B_O]D_I}] \tag{G10}$$

and

$$\mathcal{S} = \mathcal{L}^{\perp}_{[1-A_O]B_{IO}C_{IO}D_I} + (\mathcal{P} + \mathcal{L}^{\perp}_{[1-B_O]C_{IO}D_I} + \mathcal{L}^{\perp}_{[1-C_O]D_I}) \cap (\mathcal{P} + \mathcal{L}^{\perp}_{[1-C_O]B_{IO}D_I} + \mathcal{L}^{\perp}_{[1-B_O]D_I}). \tag{G11}$$

Once again, for more general scenarios it remains an open question whether the necessary and sufficient conditions of propositions 4 and 5 coincide. Nonetheless, the same approach here can be applied to our necessary condition, which defines the cone $\mathcal{W}^{\text{sep}}_{+}$ that is an outer approximation of $\mathcal{W}^{\text{sep}}$, to characterise a subset $\mathcal{S}_+$ of causal witnesses. Solving the dual SDP problem (38) over this set allows one to find valid witnesses of causal nonseparability for a given process matrix W, even though (without proof that the necessary and sufficient conditions coincide) such a witness may not be optimal amongst the full set of causal witnesses $\mathcal{S}$.

ORCID iDs

Julian Wechs https://orcid.org/0000-0002-0395-6791
Alastair A Abbott https://orcid.org/0000-0002-2759-633X
Cyril Branciard https://orcid.org/0000-0001-9460-825X

References

[1] Hardy L 2005 arXiv:gr-qc/0509120
[2] Oreshkov O, Costa F and Brukner Č 2012 *Nat. Commun.* **3** 1092
[3] Zych M, Costa F, Pikovski I and Brukner Č 2017 arXiv:1708.00248
[4] Chiribella G, D'Ariano G M, Perinotti P and Valiron B 2013 *Phys. Rev.* A **88** 022318
[5] Chiribella G 2012 *Phys. Rev.* A **86** 040301
[6] Colnaghi T, D'Ariano G M, Facchini S and Perinotti P 2012 *Phys. Lett.* A **376** 2940
[7] Araújo M, Costa F and Brukner Č 2014 *Phys. Rev. Lett.* **113** 250402
[8] Facchini S and Perdrix S 2015 *Theory and Applications of Models of Computation* ed R Jain *et al* (Berlin: Springer) pp 324–31
[9] Feix A, Araújo M and Brukner Č 2015 *Phys. Rev.* A **92** 052326
[10] Guérin P A, Feix A, Araújo M and Brukner Č 2016 *Phys. Rev. Lett.* **117** 100502
[11] Ebler D, Salek S and Chiribella G 2018 *Phys. Rev. Lett.* **120** 120502
[12] Salek S, Ebler D and Chiribella G 2018 arXiv:1809.06655
[13] Chiribella G, Banik M, Bhattacharya S S, Guha T, Alimuddin M, Roy A, Saha S, Agrawal S and Kar G 2018 arXiv:1810.10457
[14] Araújo M, Branciard C, Costa F, Feix A, Giarmatzi C and Brukner Č 2015 *New J. Phys.* **17** 102001
[15] Branciard C 2016 *Sci. Rep.* **6** 26018
[16] Baumeler Ä and Wolf S 2014 *2014 IEEE Int. Symp. Inf. Theory (ISIT)* pp 526–30
[17] Oreshkov O and Giarmatzi C 2016 *New J. Phys.* **18** 093020
[18] Giacomini F, Castro-Ruiz E and Brukner Č 2016 *New J. Phys.* **18** 113026
[19] Davies E B and Lewis J T 1970 *Commun. Math. Phys.* **17** 239
[20] Jamiołkowski A 1972 *Rep. Math. Phys.* **3** 275
[21] Choi M-D 1975 *Linear Algebr. Appl.* **10** 285
[22] Abbott A A, Giarmatzi C, Costa F and Branciard C 2016 *Phys. Rev.* A **94** 032131
[23] Chiribella G, D'Ariano G M and Perinotti P 2008 *Europhys. Lett.* **83** 30004
[24] Procopio L M, Moqanaki A, Araújo M, Costa F, Alonso Calafell I, Dowd E G, Hamel D R, Rozema L A, Brukner Č and Walther P 2015 *Nat. Commun.* **6** 7913
[25] 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
[26] Rubino G, Rozema L A, Massa F, Araújo M, Zych M, Brukner Č and Walther P 2017 arXiv:1712.06884
[27] Goswami K, Giarmatzi C, Kewming M, Costa F, Branciard C, Romero J and White A G 2018 *Phys. Rev. Lett.* **121** 090503
[28] Wei K *et al* 2018 arXiv:1810.10238
[29] Oreshkov O 2018 arXiv:1801.07594
[30] Gutoski G and Watrous J 2006 *Proc. 39th ACM STOC* pp 565–74
[31] Chiribella G, D'Ariano G M and Perinotti P 2009 *Phys. Rev.* A **80** 022339
[32] Oreshkov O 2016 private communication
[33] Wechs J, Dourdent H, Abbott A A and Branciard C in preparation
[34] Branciard C, Araújo M, Feix A, Costa F and Brukner Č 2016 *New J. Phys.* **18** 013008
[35] Abbott A A, Wechs J, Costa F and Branciard C 2017 *Quantum* **1** 39
[36] Rockafellar R T 1970 *Convex Analysis* (Princeton, NJ: Princeton University Press)
[37] Acín A, Bruß D, Lewenstein M and Sanpera A 2001 *Phys. Rev. Lett.* **87** 040401
[38] Svetlichny G 1987 *Phys. Rev.* D **35** 3066
[39] Gallego R, Würflinger L E, Acín A and Navascués M 2012 *Phys. Rev. Lett.* **109** 070401
[40] Bancal J-D, Barrett J, Gisin N and Pironio S 2013 *Phys. Rev.* A **88** 014102
[41] Jia D and Sakharwade N 2018 *Phys. Rev.* A **97** 032110
[42] Guérin P A, Krumm M, Budroni C and Brukner Č 2018 arXiv:1806.10374
[43] Feix A, Araújo M and Brukner Č 2016 *New. J. Phys.* **18** 083040
[44] Miklin N, Abbott A A, Branciard C, Chaves R and Budroni C 2017 *New J. Phys.* **19** 113041
[45] Milz S, Pollock F A, Le T P, Chiribella G and Modi K 2018 *New J. Phys.* **20** 033033
[46] Oreshkov O 2018 private communication
[47] Rudin W 1976 *Principles of Mathematical Analysis* 3rd edn (New York: McGraw-Hill)
[48] Nesterov Y and Nemirovskii A 1994 *Interior-point Polynomial Algorithms in Convex Programming* (*Stud. Appl. Math.* vol 32) (Philadelphia, PA: SIAM)
[49] Grant M and Boyd S 2014 CVX: matlab software for disciplined convex programming, v2.1