Causally nonseparable processes admitting a causal model

Adrien Feix$^{1,2,3}$, Mateus Araújo$^{1,2}$ and Časlav Brukner$^{1,2}$

$^1$Faculty of Physics, University of Vienna, Boltzmanngasse 5, A-1090 Vienna, Austria
$^2$Institute for Quantum Optics and Quantum Information (IQOQI), Boltzmanngasse 3, A-1090 Vienna, Austria
$^3$Author to whom any correspondence should be addressed.

E-mail: adrien.feix@univie.ac.at

Abstract

A recent framework of quantum theory with no global causal order predicts the existence of ‘causally nonseparable’ processes. Some of these processes produce correlations incompatible with any causal order (they violate so-called ‘causal inequalities’ analogous to Bell inequalities) while others do not (they admit a ‘causal model’ analogous to a local model). Here we show for the first time that bipartite causally nonseparable processes with a causal model exist, and give evidence that they have no clear physical interpretation. We also provide an algorithm to generate processes of this kind and show that they have nonzero measure in the set of all processes. We demonstrate the existence of processes which stop violating causal inequalities but are still causally nonseparable when mixed with a certain amount of ‘white noise’. This is reminiscent of the behavior of Werner states in the context of entanglement and nonlocality. Finally, we provide numerical evidence for the existence of causally nonseparable processes which have a causal model even when extended with an entangled state shared among the parties.

1. Introduction

It is well-known that quantum mechanics is at odds with naive notions of reality and locality as predicted by Bell’s celebrated theorem [1, 2]. One might wonder whether peculiar quantum features could challenge other fundamental notions, like the concept of causality, as well. The process matrix formalism of Oreshkov, Costa and Brukner [3] was developed to explore this question—studying the most general causal structures compatible with local quantum mechanics for two parties $A$ and $B$.

Surprisingly, the formalism predicts causal structures which are ‘causally nonseparable’: they correspond neither to $A$ being before $B$ nor to $B$ being before $A$, nor to a probabilistic mixture thereof. These causal structures can produce correlations incompatible with any definite causal order, violating so-called ‘causal inequalities’ [3, 4].

However, the empirical relevance of these results is still completely unclear. Do they appear in some physical situations or are they merely a mathematical artifact of the process matrix formalism?

For three and more parties, there are causally nonseparable processes whose physical realization is known—one instance is the ‘quantum switch’ [6], where the causal order between two parties $A$ and $B$ is controlled by a quantum system belonging to a third party $C$. Processes of this kind, however, cannot violate causal inequalities [7, 8]; they admit a ‘causal model’, i.e., a causally separable process is capable of reproducing their correlations. Their causal nonseparability can only be certified through device-dependent ‘causal tomography’ or ‘causal witnesses’ [7]. This is analogous to states which are entangled but cannot violate Bell inequalities, i.e., for which a ‘local model’ exists [9].

Since the only causally nonseparable processes known to be physically implementable have a causal model, it is tempting to conjecture that the inability to violate causal inequalities without [10] or with [8] operations

For two parties, the operations implemented to violate causal inequalities have to be quantum [3]; surprisingly, for three or more parties, classical operations are sufficient [5].

© 2016 IOP Publishing Ltd and Deutsche Physikalische Gesellschaft
extended to shared entangled states by all parties singles out the physical causal structures from unphysical ones. Reference [8] contains an example of a tripartite process matrix with a causal model but which does not remain causal under extensions, i.e., is not ‘extensibly causal’, demonstrating the difference between the two notions for more than two parties.

In this paper, we provide an example of a bipartite causally nonseparable process with a causal model. Furthermore, we give numerical evidence that bipartite nonseparable processes exist which do not violate causal inequalities, even when extended with entanglement. No physical interpretation is known for these processes.

The paper is organized as follows: section 2 introduces the process matrix formalism and the definitions of causal nonseparability and causal inequalities. In section 3, we define a class of two-party causally nonseparable processes and construct a causal model for them. This shows that the sets of causally nonseparable and causal inequality violating processes are distinct also in the bipartite case. Since the causally nonseparable processes with a causal model can be interpreted as the mixture of physically implementable process with an unphysical process, this gives evidence that they are not implementable in nature.

In section 4, we provide an algorithm to construct nonseparable processes with a causal model by composing a random causally separable process with a non-completely positive (CP) map on one party. Using a random sample generated by a ‘hit-and-run’ Markov chain [11, 12], we also show that nonseparable processes with a causal model have nonzero measure in the space of all processes.

In section 5, we construct a family of Werner processes', which mimic the behavior of Werner states [9] with respect to nonlocality and entanglement. The Werner processes’ causal nonseparability is more resistant to the introduction of ‘white noise’ than its ability to violate causal inequalities. This shows that the analogy between causal nonseparability and causal inequalities on the one hand, and entanglement and Bell inequalities on the other, extends beyond what was previously known [3, 7, 8, 13].

Finally, in section 6, we examine the behavior of the processes with added shared entanglement between the parties. While some of the processes we constructed do violate causal inequalities when extended in this way (are not ‘extensibly causal’), numerical calculations indicate that the ability to violate causal inequalities disappears when adding a little white noise, at which the causal nonseparability is preserved. We conjecture that there are processes which are extensibly causal, and yet not physically implementable.

2. Causal nonseparability and causal inequalities

Quantum circuits can be thought of as a formalization of causal structures with a definite causal order. They consist of wires, representing quantum systems, which connect boxes, representing quantum operations. While for quantum circuits, the order of the operations is fixed [14], situations where the order of operations is not well-defined are readily represented in the process matrix formalism [3], which can be thought of as a generalization of the quantum circuit formalism. We will briefly introduce the main elements of the formalism; a more detailed introduction to it can be found in [7].

A quantum operation maps a density matrix $\rho_A \in A_I$ to a density matrix $\rho_A \in A_O$ (where $A_I(A_O)$ denotes the space of linear operators on the Hilbert space $\mathcal{H}_I(\mathcal{H}_O)$). The most general operations within the quantum formalism are CP maps $\mathcal{M}_A : A_I \rightarrow A_O$. Using the Choi–Jamiolkowski (CJ) [15, 16] isomorphism, one can represent every CP map as an operator acting on the tensor product of the input and output Hilbert spaces:

$$M_A = ([I] \otimes \mathcal{M}_A)(|I\rangle\langle I|)^T \in A_I \otimes A_O,$$

(1)

where $I$ is the identity map and $|I\rangle\langle I|$ is the space of linear operators on the Hilbert space $\mathcal{H}_I(\mathcal{H}_O)$. The most general operations within the quantum formalism are CP maps $\mathcal{M}_A : A_I \rightarrow A_O$. Using the Choi–Jamiolkowski (CJ) [15, 16] isomorphism, one can represent every CP map as an operator acting on the tensor product of the input and output Hilbert spaces:

$$M_A = ([I] \otimes \mathcal{M}_A)(|I\rangle\langle I|)^T \in A_I \otimes A_O,$$

(1)

where $I$ is the identity map and $|I\rangle\langle I|$ is the space of linear operators on the Hilbert space $\mathcal{H}_I(\mathcal{H}_O)$. The most general operations within the quantum formalism are CP maps $\mathcal{M}_A : A_I \rightarrow A_O$. Using the Choi–Jamiolkowski (CJ) [15, 16] isomorphism, one can represent every CP map as an operator acting on the tensor product of the input and output Hilbert spaces:

$$M_A = ([I] \otimes \mathcal{M}_A)(|I\rangle\langle I|)^T \in A_I \otimes A_O,$$

(1)

where $I$ is the identity map and $|I\rangle\langle I|$ is the space of linear operators on the Hilbert space $\mathcal{H}_I(\mathcal{H}_O)$. The most general operations within the quantum formalism are CP maps $\mathcal{M}_A : A_I \rightarrow A_O$. Using the Choi–Jamiolkowski (CJ) [15, 16] isomorphism, one can represent every CP map as an operator acting on the tensor product of the input and output Hilbert spaces:

$$M_A = ([I] \otimes \mathcal{M}_A)(|I\rangle\langle I|)^T \in A_I \otimes A_O,$$

(1)

where $I$ is the identity map and $|I\rangle\langle I|$ is the space of linear operators on the Hilbert space $\mathcal{H}_I(\mathcal{H}_O)$. The most general operations within the quantum formalism are CP maps $\mathcal{M}_A : A_I \rightarrow A_O$. Using the Choi–Jamiolkowski (CJ) [15, 16] isomorphism, one can represent every CP map as an operator acting on the tensor product of the input and output Hilbert spaces:

$$M_A = ([I] \otimes \mathcal{M}_A)(|I\rangle\langle I|)^T \in A_I \otimes A_O,$$

(1)

where $I$ is the identity map and $|I\rangle\langle I|$ is the space of linear operators on the Hilbert space $\mathcal{H}_I(\mathcal{H}_O)$. The most general operations within the quantum formalism are CP maps $\mathcal{M}_A : A_I \rightarrow A_O$. Using the Choi–Jamiolkowski (CJ) [15, 16] isomorphism, one can represent every CP map as an operator acting on the tensor product of the input and output Hilbert spaces:

$$M_A = ([I] \otimes \mathcal{M}_A)(|I\rangle\langle I|)^T \in A_I \otimes A_O,$$

(1)

where $I$ is the identity map and $|I\rangle\langle I|$ is the space of linear operators on the Hilbert space $\mathcal{H}_I(\mathcal{H}_O)$. The most general operations within the quantum formalism are CP maps $\mathcal{M}_A : A_I \rightarrow A_O$. Using the Choi–Jamiolkowski (CJ) [15, 16] isomorphism, one can represent every CP map as an operator acting on the tensor product of the input and output Hilbert spaces:

$$M_A = ([I] \otimes \mathcal{M}_A)(|I\rangle\langle I|)^T \in A_I \otimes A_O,$$

(1)

where $I$ is the identity map and $|I\rangle\langle I|$ is the space of linear operators on the Hilbert space $\mathcal{H}_I(\mathcal{H}_O)$. The most general operations within the quantum formalism are CP maps $\mathcal{M}_A : A_I \rightarrow A_O$. Using the Choi–Jamiolkowski (CJ) [15, 16] isomorphism, one can represent every CP map as an operator acting on the tensor product of the input and output Hilbert spaces:

$$M_A = ([I] \otimes \mathcal{M}_A)(|I\rangle\langle I|)^T \in A_I \otimes A_O,$$

(1)

where $I$ is the identity map and $|I\rangle\langle I|$ is the space of linear operators on the Hilbert space $\mathcal{H}_I(\mathcal{H}_O)$. The most general operations within the quantum formalism are CP maps $\mathcal{M}_A : A_I \rightarrow A_O$. Using the Choi–Jamiolkowski (CJ) [15, 16] isomorphism, one can represent every CP map as an operator acting on the tensor product of the input and output Hilbert spaces:

$$M_A = ([I] \otimes \mathcal{M}_A)(|I\rangle\langle I|)^T \in A_I \otimes A_O,$$

(1)

where $I$ is the identity map and $|I\rangle\langle I|$ is the space of linear operators on the Hilbert space $\mathcal{H}_I(\mathcal{H}_O)$. The most general operations within the quantum formalism are CP maps $\mathcal{M}_A : A_I \rightarrow A_O$. Using the Choi–Jamiolkowski (CJ) [15, 16] isomorphism, one can represent every CP map as an operator acting on the tensor product of the input and output Hilbert spaces:

$$M_A = ([I] \otimes \mathcal{M}_A)(|I\rangle\langle I|)^T \in A_I \otimes A_O,$$

(1)
where $p_{A < B}$ is a hermitian operator such that $\text{tr}[W_{\text{sep}} S] \geq 0$ for all causally separable processes $W_{\text{sep}}$ [7]. The property $\text{tr}[WS] < 0$ can be verified experimentally by measuring a set of operators for Alice and Bob, certifying that the process $W$ is not causally separable. Note that, in analogy to entanglement witnesses [20], this certification of causal nonseparability relies on a partial tomography of the process and thus requires trust in Alice’s and Bob’s local operations: it is ‘device-dependent’.

It is well-known that the entanglement of a quantum state can be certified device-independently (without requiring trust Alice’s and Bob’s operations) if the probability distribution resulting from a set of measurements violates a Bell inequality [21]. In an analogous way, causal nonseparability can be device-indistinguishably confirmed using causal inequalities [3–5, 7, 8], where the ‘non-causal’ correlations between Alice and Bob alone suffice to show that the process they use is not causally separable, without additional trust in their local operations.

The condition for a probability distribution to be ‘causal’, i.e., not to violate any causal inequality, is simply that it can be decomposed into a convex combination of a probability distribution which is no-signaling from Bob to Alice $^5$ ($p_{A < B}$) and a probability distribution which is no-signaling from Alice to Bob ($p_{B < A}$) [4]:

$$p_{\text{causal}} = q p_{A < B} + (1 - q) p_{B < A}.$$  

Note that the correlations generated by a causally ordered process cannot violate any causal inequality.

For the scenario where Alice (Bob) has an input bit $x$ ($y$) and outputs one bit $a$ ($b$), one causal inequality is a bound on the probability of success of the ‘guess your neighbor’s input’ (GYNI) [4]:

$$p_{\text{GYNI}} = \frac{1}{4} \sum_{x, y} p(a = y, b = x|a, y) \leq \frac{1}{2},$$  

Some valid processes and local strategies which result in correlations violating (9) are described in [4].

The relation between causally nonseparable processes and the violation of causal inequalities is not yet fully understood. On the one hand, there exist causally nonseparable processes that can be physically implemented [22] but have a causal model—they do not violate causal inequalities [7, 8]. An example of such a process is the ‘quantum switch’ [6]. On the other hand, there are processes that can violate causal inequalities, but it is not known if they can be realized in nature, prompting the conjecture that only processes with a causal model are physically implementable [10].

Another natural feature to investigate is the (in)ability for a process to violate causal inequalities, even when extended with an entangled state shared by all parties. We will call processes which do not allow for such a violation ‘extensibly causal’ [8] and come back to this concept in detail in section 6.

In the bipartite case, all previously known nonseparable processes violate causal inequalities and it is not clear if nonseparable processes with a causal model even exist [4]. In the next section, we explicitly provide a class of nonseparable bipartite processes that allow for a causal model.

---

$^5$ When Alice is given the input $x$ and outputs $a$ (Bob is given an input $y$ and outputs $b$), no-signaling from Bob to Alice implies that the marginal probability on Alice’s side does not depend on Bob’s input: $\sum_b p_{A < B}(ab|xy) = \sum_b p_{A < B}(ab|x_1, y_1), \forall x, y, y_1$. 

$^6$ Some correlations with a causal model are known to be quantum switch correlations [6], which can be generated by the quantum switch device [20].
3. Causally nonseparable processes with a causal model

We will consider the following class of processes (all the operators are understood to act on qubits, $d_A = d_{A_0} = d_B = d_{B_0} = 2$):

$$W^{A\rightarrow B} := I^p + \frac{1}{12} (IZZ I + IXX I + IYY I),$$

$$W^{B\rightarrow A} := I^p + \frac{1}{4} (ZIZX),$$

$$W := q W^{A\rightarrow B} + (1 - q + \epsilon) W^{B\rightarrow A} - \epsilon I^p.$$  (10)

Here $I, X, Y, Z$ are the Pauli matrices and the tensor products between the Hilbert spaces $A_I, A_{I_0}, B_I, B_{I_0}$ are implicit, as in the remainder of the paper. The process matrix (10) is positive semidefinite for $\epsilon \leq q - 1 + \sqrt{(1 - q)(q + 3)}$ and causally nonseparable for $\epsilon > 0$. As shown in appendix A, its random robustness is $R_0(W) = \epsilon$. It is maximal for $q = \sqrt{3} - 1 \approx 0.732$, where $\epsilon = \frac{4}{3} - 2 \approx 0.309$.

The proof that the process (10) cannot be used to violate any causal inequalities, for any local strategy$^6$ consists of two steps: (i) we show that the set of correlations compatible with $W$ is the same as the set of correlations achievable with $W^{T\rho}$ (where $T\rho$ denotes the partial transpose of the systems $B_I B_{I_0}$ with respect to the computational basis); (ii) we verify that $W^{T\rho}$ is valid and causally separable, hence cannot violate causal inequalities. Taken together, this establishes that $W$ cannot violate any causal inequalities either and therefore admits a causal model.

The first part of the proof is simple. Using definition (2) and the self-duality of transposition, we rewrite the probability distribution:

$$p(ab|xy) = \text{tr}[W^a_{x^a} \otimes \rho^b_{y^b}] = \text{tr}[W^{T\rho}_{x^a} \otimes (\rho^b)^{T\rho}],$$  (11)

Additionally, for any quantum instrument $[\{\rho^a_k\}]$, the instrument $[\{\rho^b_k\}]$ is also valid, since transposition maps CP maps to CP maps and trace-preserving maps to trace-preserving maps$^7$. This establishes (ii), namely that the correlations achievable with $W^{T\rho}$ are the same as those compatible with $W$—note that this holds for any process, even when $W^{T\rho}$ is not positive semidefinite, and therefore not a valid process matrix. In such a case, the probability distribution will be well-defined for local measurements, but not when extending the process with an entangled state between Alice and Bob, an extension which is physically meaningful and to which we will come back in section 6.

For the class of process matrices given in equation (10), $W^{T\rho}$ is always positive semidefinite. We will now explicitly decompose $W^{T\rho}$ as a convex combination of causally ordered process matrices, proving that it is causally separable (formally, this implies that $R_0(W^{T\rho}) \leq 0$).

First, one should notice the similarity of $W^{A\rightarrow B}$ with the process matrix $D^{A\rightarrow B}_{2,2}$ of a depolarizing channel (with $\frac{2}{3}$ probability of depolarizing and $\frac{1}{3}$ probability of perfectly transmitting the state) from Alice to Bob

$$D^{A\rightarrow B}_{2,2} := I^p + \frac{1}{12} (IZZ I + IXX I - IYY I),$$  (12)

where only the sign of the term IYY I differs compared to $W^{A\rightarrow B}$ of (10). This exactly corresponds to a partial transpose of the systems $B_i B_{i_0}$, such that $W^{A\rightarrow B} = (D^{A\rightarrow B}_{2,2})^{T\rho}$. Using the definition of the depolarizing process

$$D^{A\rightarrow B}_{2,2} = \frac{2}{3} I^p + \frac{1}{3} I^{A\rightarrow B},$$

where $I^{A\rightarrow B} = \mathbb{1}_A | I \rangle \langle I | A_{I_0}$, we can write $W^{T\rho}$ as:

$$W^{T\rho} = \frac{2\epsilon}{3} I^p + \frac{2}{3} I^{A\rightarrow B} + (1 - q + \epsilon) W^{B\rightarrow A} - \epsilon I^p,$$

$$= \frac{2\epsilon}{3} I^{A\rightarrow B} + (1 - q + \epsilon) W^{B\rightarrow A} + \left(\frac{2\epsilon}{3} - \epsilon\right) I^p,$$  (13)

which is a convex decomposition into causally ordered processes as long as $\epsilon \leq \epsilon_c(q) = \frac{2\epsilon}{3}$. Since $\epsilon \leq q - 1 + \sqrt{(1 - q)(q + 3)}$ for the process given in equation (10) to be valid, and

$$q - 1 + \sqrt{(1 - q)(q + 3)} \leq \epsilon_c(q),$$

the whole class of processes defined in equation (10) cannot violate causal inequalities. For a graphical representation of this relationship, see figure 2.

This concludes the proof and provides an explicit causal model for the process $W$ with the instruments $[\{\xi^a_k\}]$ and $[\{\eta^b_k\}]$: the process $W^\rho$ with the instruments $[\xi^a_k]$ and $[\eta^b_k]$. Since $W^{T\rho}$ is causally separable, it can be

$^6$Note that our proof guarantees the existence of a causal model, which means that the correlations belong to every causal polytope, without restriction on the number of inputs and outputs for each party.

$^7$The condition on the CJ representation of a CPTP map is that $\text{tr}_{B_0} M^{B_0} = I^p$ and it implies $\text{tr}_{B_0} (M^{B_0})^T = I^p$. 

---

4
interpreted as a probabilistic mixture of two causally ordered processes. There are infinitely many such decompositions since the term $\mathcal{I} \circ \mathcal{I}$ in equation (13) can be split and added to $\mathcal{I} \mathcal{A}$ and $\mathcal{I} \mathcal{B}$ in any proportion. $W$ can be taken to be $W^\top$ composed with a transpose map on $B$. The process matrix $\frac{1}{2} (\mathcal{I} A \circ \mathcal{I} B)$ becomes positive (meaning that the associated map is CP) when adding at least $\frac{2q}{3}$ of white noise. However, the maximal noise that can be admixed by transferring the $\mathcal{I}$ term in (13) is $\frac{2q}{3} - \epsilon$, which is strictly smaller. Therefore, the causal model for $W$ suggests a natural interpretation for it as a convex combination of an unphysical channel from Alice to Bob (as it is not CP) with a physical channel from Bob to Alice. This provides some evidence—yet not a proof—that the process $W$ is not physically implementable.

4. Random causally nonseparable processes with a causal model

In this section, we develop a method to construct a broad class of causally nonseparable processes with a causal model. Given a random causally separable process, one applies a positive, but not necessarily CP map $Q_B(\cdot)$ on Bob’s side:

$$W_{\text{sep}} = Q_B(W_{\text{sep}}).$$

If the resulting process has negative eigenvalues, it is discarded; otherwise, it is a valid process matrix. Using the same argument as in the preceding section, we know that the resulting process $W_{\text{sep}}$ will have a causal model.
Sometimes—and these are the interesting cases—the process will also be causally nonseparable. This can readily be checked via SDP (7).

We generated causally separable process matrices \( W_{\text{sep}} \) where

\[
\begin{bmatrix}
A & B \\
A & B
\end{bmatrix}
\]

according to an asymptotically uniform distribution using the ‘hit-and-run’ technique (see appendix C for details) and equation (14), using the transposition map \( T_B \) for \(QB\). We found that most (69%) of the resulting matrices were positive and hence valid process matrices. About half of these turn out to be causally nonseparable while—by construction—allowing for a causal model (we denote this set by \( \mathcal{W}_{\text{nsep}} \)). The histogram of the resulting causal nonseparabilities is shown in figure 3. There is therefore a finite probability of generating a nonseparable process with a causal model starting from a random of causally separable process. Since the map \( T_B \) is measure-preserving, the set of causally nonseparable processes which admit a causal model is of the same dimension as the set of valid processes \( \mathcal{W} \) itself (see appendix B for details).

5. ‘Werner’ causally nonseparable processes

We will denote the set of causally nonseparable processes as \( \mathcal{W}_{\text{nsep}} \). It is composed of the set of processes with a causal model (\( \mathcal{W}_{\text{nsep}}^{(c)} \)) and the set of processes that can violate causal inequalities (\( \mathcal{W}_{\text{nsep}}^{(nc)} \)), which are graphically represented in figure 4.

We now construct bipartite processes violating a causal inequality, but which, mixed with some amount of ‘white noise’ \( \mathbb{1} \), turn into causally nonseparable processes with a causal model. This behavior is reminiscent of ‘Werner states’, which violate Bell inequalities until a noise level of up to \( \frac{2}{3} \) but are entangled when mixed with noise up to a level of \( \frac{2}{3} \) [9]. It shows that the analogy between causal inequalities and causal nonseparability, on the one hand, and Bell inequalities and entanglement, on the other hand, applies to the two-party case.

The idea is to use a convex combination of a process in \( \mathcal{W}_{\text{nsep}}^{(c)} \) and a process in \( \mathcal{W}_{\text{nsep}}^{(nc)} \), which is also invariant under partial transposition with respect to \( B \). In this way, one can generate a broad class of ‘Werner causally nonseparable processes’.

We will use the process defined in equation (10) with the maximal causal nonseparability:

\[
W_{\text{opt}} := (\sqrt{3} - 1) W^{A<B} + \frac{1}{\sqrt{3}} W^{B<A} - \left( \frac{4}{\sqrt{3}} - 2 \right) \mathbb{1}
\]  

\[
(15)
\]

---

\[\text{Figure 3. Histogram of the random robustness } R, \text{ of the subset of } 690 (\approx 69\%) \text{ valid processes } W_{\text{sep}} \text{ generated from 1000 uniformly distributed causally separable processes (see appendix C) and applying partial transposition on Bob’s side. The 366 (\approx 53\%) which are causally nonseparable, while admitting a causal model, are represented in red.} \]
together with the process

$$W_{\text{OCB}} = \mathcal{F} + \frac{1}{4\sqrt{2}} (IZZ1 + ZIXZ),$$

which was proposed and shown to violate causal inequalities in [3]. In appendix A, we show that the resulting mixture

$$W_{\text{mix}}(\alpha) = \alpha W_{\text{opt}} + (1 - \alpha) W_{\text{OCB}}$$

has nonseparability $R_{\text{mix}}(\alpha) = R_\ell(W_{\text{mix}}(\alpha)) = \alpha R_\ell(W_{\text{opt}}) + (1 - \alpha) R_\ell(W_{\text{OCB}})$, where $R_\ell(W_{\text{OCB}}) = \sqrt{2} - 1$ is the random robustness of $W_{\text{OCB}}$.

Following the same argument as in the previous section, we can now examine the causal nonseparability of $W_{\text{mix}}$, since, by transferring the partial transpose onto Bob’s CP maps, we know that it can produce exactly the same correlations as $W_{\text{mix}}$. Its nonseparability $R'_{\text{mix}}(\alpha) = R_\ell(W'_{\text{mix}}(\alpha))$ is again the weighted average (see appendix A) of the nonseparabilities of $W'_{\text{opt}}$ and of $W'_{\text{OCB}} = W_{\text{OCB}}$:

$$R'_{\text{mix}}(\alpha) = \alpha R_\ell(W'_{\text{opt}}) + (1 - \alpha) R_\ell(W_{\text{OCB}}) = R_{\text{mix}}(\alpha) + \frac{2\alpha}{3} R_\ell(W'_{\text{opt}}) < R_{\text{mix}}(\alpha),$$

where we used $R_\ell(W'_{\text{opt}}) = \frac{2\sqrt{2} - 1}{3} < 0$ and $R_\ell(W_{\text{OCB}}) = \sqrt{2} - 1$. This means that there is a finite gap between the nonseparability of $W_{\text{mix}}$ and the nonseparability of $W'_{\text{mix}}$.

Using a see-saw algorithm [4], we numerically verified that $W_{\text{mix}}(\alpha)$ indeed violates the causal inequalities of [4] as long as $W'_{\text{mix}}$ is nonseparable ($R'_{\text{mix}}(\alpha) > 0$), i.e., when $\alpha < \frac{3}{4 + \frac{3}{2\sqrt{2} - 2\sqrt{3}}} \approx 0.6987$.

This gap translates into a gap between the level of white noise $\mathcal{F}$ that $W_{\text{mix}}$ can tolerate before admitting a causal model and the level of noise at which it becomes nonseparable. We therefore define the ’Werner process’ as a convex combination of $\mathcal{F}$ and $W_{\text{mix}}(\alpha)$:

$$W_{\text{Wer}}(\gamma, \alpha) = (1 - \gamma) W_{\text{mix}}(\alpha) + \gamma \mathcal{F}.$$  

Using the definition of nonseparability (7), one can verify that the following relations hold (see appendix A):

$$W_{\text{Wer}}(\gamma) < \frac{R_{\text{mix}}(\alpha)}{1 + R_{\text{mix}}(\alpha)}, \quad \alpha \in \mathcal{W}_{\text{nsep}}$$

$$W'_{\text{Wer}}(\gamma) > \frac{R'_{\text{mix}}(\alpha)}{1 + R'_{\text{mix}}(\alpha)}, \quad \alpha \in \mathcal{W}_{\text{sep}}.$$  

As $W_{\text{Wer}}$ can violate causal inequalities only if $W'_{\text{Wer}}$ is causally nonseparable (remember the proof of section 3), we conclude from (20) that

$$W_{\text{Wer}} \left( \frac{R'_{\text{mix}}(\alpha)}{1 + R'_{\text{mix}}(\alpha)} \right) \leq \gamma < \frac{R_{\text{mix}}(\alpha)}{1 + R_{\text{mix}}(\alpha)}, \quad \alpha \in \mathcal{W}_{\text{nsep}}^{(c)}$$

which mimics the behavior of Werner states. See figure 4 for a graphical representation of the location of $W_{\text{Wer}}$ with respect to the different sets of processes.
6. Relationship to extensibly causal processes

In the context of the physical implementability of process matrices, it is natural to consider the extension of a process matrix with an entangled state shared between the parties. A process is ‘extensively causal’ if, even when extended with a shared entangled state, it cannot violate causal inequalities [8].

Extending a physically implementable process with an entangled state shared among the parties should also result in a physically implementable process. On this account, it is important to consider not only whether a process has a causal model, but rather whether it has such a model when extended with an entangled state. In [8], an example of a tripartite process with a causal model but which is not extensively causal was presented, showing that both notions really differ and that the violation of causal inequalities can be ‘activated’ by entanglement—and for which no physical implementation is known.

Note that the proof (section 3) of the existence of causal model for \( W \) does not hold when the process is extended with an entangled state between Alice and Bob. It therefore cannot prove that \( W \) is extensively causal. It crucially relies on the fact that the transpose \( \{ (a_i^b)^T \} \) of a valid instrument for Bob \( \{ a_i^b \} \) is still a valid instrument. However, taking the full transpose on Bob’s instrument would lead to a ‘causal model’ with a partial transpose of the shared entangled state, which can lead to negative probabilities. Conversely, the partial transpose of Bob’s instrument (with no transposition on Bob’s part of the entangled state) is not a valid instrument and does not yield positive probabilities in general.

To numerically study whether \( W_{\text{opt}} \) from equation (15) is extensively causal, we extended it with a maximally entangled state of two ququarts \( |\phi\rangle^A^B^C^D := \frac{1}{4} (|00\rangle + |11\rangle + |22\rangle + |33\rangle) \):

\[
W_{\text{ext}} := W_{\text{opt}} \otimes |\phi\rangle^A^B^C^D.
\]  

We chose a maximally entangled ququart state because we believe that extending \( W_{\text{opt}} \) with a higher dimensional state would not improve its ability to violate causal inequalities.

Using the see-saw algorithm, we optimized \( W_{\text{ext}} \) for a violation of the simplest causal inequalities [4]. We found that \( W_{\text{ext}} \) is able to violate (by about \( 8 \times 10^{-3} \)) the GYNI inequality (9), which proves that \( W_{\text{opt}} \) is not extensively causal. Incidentally, it also shows that in the bipartite case as well, the violation of causal inequalities can be ‘activated’ using entanglement.

If we adopt the view that extensively causal processes are physical, the activation of violation of causal inequalities, this suffices to exclude \( W_{\text{opt}} \) in the same way as the tripartite process with a causal model but which is not extensively causal, given in [8], independently of the argument based on the decomposition of section 3. However, this is not possible anymore when admixing a small amount of white noise: \((1 - \kappa) W_{\text{ext}} + \kappa \mathbb{I}\). We ran the see-saw algorithm for different levels of white noise and found no violation of GYNI for \( \kappa > 3.3 \times 10^{-4} \) (see figure 5 for a graphical representation of the relationship between noise and violation of GYNI). Similarly, neither the other causal inequality from [4] nor the ‘original’ causal inequality from [3] could be violated through see-saw optimization.

This gives reasonable evidence\(^9\) that \( W_{\text{opt}} \) mixed with very little white noise is extensively causal, while still being causally nonseparable (see figure 6 for a graphical representation). The argument for unphysicality of section 3 still applies to it. This leads us to conjecture that some extensively causal processes cannot be physically realized.

7. Conclusions

We studied the classification of causally nonseparable process matrices for two parties and found that composing a class of causally separable processes with a transpose map on one party’s side results in nonseparable processes with causal models, i.e., that cannot violate causal inequalities.

Since the only interpretation we know relies on applying a non-CP map (which is itself is unphysical) to a valid process, the conjecture that processes which do not violate causal inequalities are physically implementable is undermined.

We also provided a simple algorithm to generate nonseparable processes with causal models—starting from a random separable process and composing it with a positive, but not CP map on one party’s side. With a finite probability, this yields a nonseparable process with a causal model and shows that the measure of such processes is nonzero within the space of valid processes. The ‘hit-and-run’ algorithm we used to generate random process matrices might be of independent interest.

\(^9\) It falls short of being a proof because (i) the entangled state added in \( W_{\text{ext}} \) is finite-dimensional; (ii) the see-saw technique is not guaranteed to converge to the global optimum; (iii) only the three known bipartite causal inequalities were tested, inequalities with more settings might still be violated.
We then developed the analogy between entanglement/nonlocality and causal nonseparability/noncausal correlations by providing a process analogous to a Werner state: it starts having a causal model when mixed with a certain amount of white noise, while still being strictly causally nonseparable.

Finally, we studied whether our processes still have a causal model when extended with an entangled state shared between Alice and Bob (whether they are ‘extensibly causal’). The numerical evidence prompted us to conjecture that some of the nonseparable processes we studied are extensibly causal, while not being physically implementable.

An important question remains open: if some processes which have an (extensible) causal model are nonphysical, which other criterion should be used to rule them out? One fairly natural approach is to postulate a ‘purification principle’, according to which physically realizable processes can be recovered as part of a pure process in a larger space [24].

Acknowledgments

We thank Ognian Oreshkov and Fabio Costa for pointing out the significance of ‘extensible causal separability’ in the context of physical implementability, and Ralph Silva for a useful comment on an earlier version of this manuscript. We acknowledge support from the European Commission project RAQUEL (No. 323970); the Austrian Science Fund (FWF) through the Special Research Programme FoQuS, the Doctoral Programme CoQuS and Individual Project (No. 2462) and the John Templeton Foundation.
Appendix A. Analytic proofs for nonseparabilities in the main text

First, we prove that the class of processes, defined by

\[
W^{A \prec B} := W + \frac{1}{12} (1ZZ1 + 1XX1 + 1YY1),
\]

\[
W^{B \prec A} := W + \frac{1}{4} (Z1XZ),
\]

\[
W := q W^{A \prec B} + (1 - q + \epsilon) W^{B \prec A} - \epsilon 1,
\]

has causal separability \( R_t(W) = \epsilon \). To do so, we define a causal witness \( S_W \) for it:

\[
S_W = \epsilon - \frac{1}{4}(1ZZ1 + 1XX1 + 1YY1) - \frac{1}{4}(Z1XZ).
\]

We first verify that \( S_W \) is a causal witness: \( \text{tr}_{\bar{B}} S_W \geq 0 \) and \( \text{tr}_{\bar{B}} S_W \geq 0 \), which is a sufficient condition [7] for \( S \) to have positive trace with any causally separable process. Therefore, \( \text{tr}[W_{\text{sep}} S_W] \geq 0 \) and \( S_W \) is indeed a causal witness.

We compute \( \text{tr}[S_W(W + \lambda 1)] = -\epsilon + \lambda \), which is negative for \( \lambda < \epsilon \) and implies that \( R_t(W) \geq \epsilon \).

From equation (A1), it is clear that \( W + \lambda 1 \) is causally separable for \( \lambda \geq \epsilon \), so \( R_t(W) \leq \epsilon \). This establishes that \( R_t(W) = \epsilon \).

Using the same approach, we can show that the process

\[
W_{\text{mix}}(\alpha) := \alpha W_{\text{opt}} + (1 - \alpha) W_{\text{OCSB}},
\]

where \( W_{\text{opt}} \) is defined in (15) and \( W_{\text{OCSB}} \) in (16), has nonseparability

\[
R_t(W_{\text{mix}}(\alpha)) = \alpha R_t(W_{\text{opt}}) + (1 - \alpha) R_t(W_{\text{OCSB}}) = 1 + \alpha \left( \frac{4}{\sqrt{3}} - 3 \right).
\]

The convexity of random robustness [7] implies that the random robustness of a convex combination is smaller than the convex combination of the random robustnesses, so \( R_t(W_{\text{mix}}) \leq 1 + \alpha \left( \frac{4}{\sqrt{3}} - 3 \right) \). Using the same witness \( S_W \) as before, we compute \( \text{tr}[S_W(W_{\text{mix}} + \lambda 1)] = -1 - \alpha \left( \frac{4}{\sqrt{3}} - 3 \right) + \lambda \), which is strictly negative when \( \lambda > 1 + \alpha \left( \frac{4}{\sqrt{3}} - 3 \right) \) and implies that \( R_t(W_{\text{mix}}) \geq 1 + \alpha \left( \frac{4}{\sqrt{3}} - 3 \right) \). We conclude that \( R_t(W_{\text{mix}}) = 1 + \alpha \left( \frac{4}{\sqrt{3}} - 3 \right) \).

The same proof (with the same witness \( S_W \) given in equation (A2)) can also be used to show that the causal nonseparability of \( W_{\text{mix}}^{(n)} \) is again the convex combination \( R_t(W_{\text{mix}}^{(n)}) = \alpha R_t(W_{\text{opt}}^{(n)}) + (1 - \alpha) R_t(W_{\text{OCSB}}^{(n)}) \).

Appendix B. The dimension of the set causally separable processes

Here we show that the set of causally separable processes \( \mathcal{W}_{\text{sep}} \) has the same dimension as the set of valid processes \( \mathcal{W} \), which establishes that the set of causally separable processes has nonzero measure in the set of valid processes.

We will use the Hilbert–Schmidt decomposition of operators. An arbitrary process \( W \) can be decomposed as \( W = \sum_{ijkl=0} \alpha_{ijk} \sigma^A_{ij} \otimes \sigma^B_{jk} \otimes \sigma^B_{kl} \).

The condition of normalization of probabilities, i.e., \( \text{tr}[W \cdot M_C^{\text{CPTP}} \otimes M_B^{\text{CPTP}}] = 1 \) for all CPTP-representations of CPTP \( W \), implies that some terms of the Hilbert–Schmidt decomposition, corresponding to ‘causal loops’ are excluded. In particular, \( \alpha_{ijk0} = \alpha_{0jkl} = \alpha_{0ijkl} = \alpha_{ijkl} = 0 \) for \( i, j, k, l \geq 1 \) (see the supplementary material of [3]).

Counting all the ‘allowed terms’ in the Hilbert–Schmidt decomposition, we find that the dimension \( d_W \) of \( \mathcal{W} \) is:

\[
d_W = (1 + d_A^2)(d_B^2 - 1)(d_B^2 - 1) + (d_A^2 - 1)(d_B^2 - 1) d_B^2.
\]

For causally ordered processes in \( \mathcal{W}^{A \prec B} \subset \mathcal{W}^{B \prec A} \) compatible with the causal order \( A \prec B \), some additional terms, which allow for signaling from Bob to Alice, are excluded in the Hilbert–Schmidt decomposition, reducing the dimension to

\[
d_W^{A \prec B} = d_A^2 (1 + (d_B^2 - 1) d_B^2) - 1.
\]

This means that the set of causally ordered processes \( \mathcal{W}^{A \prec B} \) has measure zero within the set of all process matrices.
Separable processes are convex combinations of $W^{A\rightarrow B}$ and $W^{B\rightarrow A}$. This means that all the terms allowed in the Hilbert–Schmidt decomposition of a valid process matrix are also allowed in the decomposition of separable processes. Therefore $W$ and $W_{\text{sep}}$ share the same basis and $d_W = d_{W_{\text{sep}}}$.

Appendix C. Generating uniformly distributed processes

We consider the space of process matrices $W$ as being embedded in $\mathbb{R}^{d_W}$. We wish to obtain a uniform sample of $W$ according to the $d_W$-dimensional volume (Lebesgue measure), which also corresponds to the measure generated by the Hilbert–Schmidt metric. We use an adaptation of the 'hit-and-run' Markov chain sampler \cite{11, 12} for this task. The iteration works as follows:

Algorithm 1

1. Select a starting point $W_0$.
2. Choose a traceless matrix $Q_{i+1}$ from a set of $d_W$ orthogonal traceless matrices and generate a random sign variable $s = \pm 1$.
3. Find $\mu$ such that $W_0 + \mu (P^i + s Q_{i+1})$ is on the boundary of the set of valid processes.
4. Generate a random real scalar $\theta \in [0, \mu]$. Take $W_{i+1} = W_i + \theta (P^i + s Q_{i+1})$ and go to step 2.

The set of directions is simply the Hilbert–Schmidt basis of allowed terms; it has dimension $d_W$ (see appendix B). For bipartite processes with $d_A = d_B = d_{A_0} = d_{B_0} = 2$, there are $d_W = 87$ possible directions to choose from.

Finding the intersection with the boundary of the set of positive processes in step 3 turns out to be a SDP

$$\max \mu \ \text{s.t.} \ W_0 + \mu (P^i + s Q_{i+1}) \geq 0.$$  \hspace{2cm} (C1)

However, the SDP which computes $\mu$ at each step of the Markov chain is a bottleneck of the algorithm. Instead, we can skip it and generate $\theta \in [0, 1]$, rejecting and retrying if the resulting process is not positive:

Algorithm 2

1. Select a starting point $W_0$.
2. Choose a traceless matrix $Q_{i+1}$ from a set of $d_W$ orthogonal traceless matrices and generate a random sign variable $s = \pm 1$.
3. Generate a random real scalar $\theta \in [0, 1]$. Take $W_{i+1} = W_i + \theta (P^i + s Q_{i+1})$.
4. If $W_{i+1} \geq 0$, go to step 2, otherwise repeat step 3.

The matrices $P^i + Q_i$ are chosen to be slightly outside the set $W$ by having slightly negative eigenvalues. Therefore, there is always a finite probability of rejection at step 4, which guarantees that the algorithm samples uniformly all the way up to the boundary.

The resulting sample $\{W_i\}_{i=0}^\infty$ is uniform when two conditions hold \cite{12}. First, from every $W_i$, $W'$ the probability to have $W_{i+d_W} = W'$ is nonzero, which is indeed true: in $d_W$ steps, one can reach any $W'$ starting from any $W_i$. Second, the uniform distribution is a stationary distribution of the Markov chain. This is also the

Figure 7. Schematic two-dimensional cut of $W$ in $\mathbb{R}^{d_W}$, the partial transpose $W^T$ here corresponds to a reflection along the horizontal axis $x$. The partial transpose of the set of causally separable processes $W_{\text{sep}}$ consists of three parts: (i) causally separable matrices $\subseteq W_{\text{sep}}$, (ii) non-valid processes $\not\subseteq W$ and (iii) valid, causally noneparable matrices with a causal model $\subseteq W_{\text{sep}}$.
case: for any \( W \), \( W' \) the probability to reach \( W' \) starting from \( W \), in \( d_W \) steps is the same as the probability to reach \( W' \) starting from \( W' \), in \( d_{W'} \) steps.

An upper bound on the convergence of the hit-and-run algorithm for convex sets (which is the case for the set \( \mathcal{W} \)) is known—in particular, the mixing time scales as \( O(d_W^2) = O(d_{W'}^2 \text{polylog } d_W) \), which matches the best known mixing times for other algorithms \[25\]. For \( d_W = 87 \) (which is the case for \( d_{A_1} = d_{B_1} = d_{A_2} = d_{B_2} = 2 \)), we would need around \( 7 \times 10^7 \) samples to achieve the same statistical significance as from a one-dimensional hit-and-run with 100 samples, which we deem sufficient for our purposes.

To sample uniformly distributed causally separable processes, we use the rejection method: after sampling \( 7 \times 10^7 \) process matrices (with a warm-up period of \( 10^6 \) discarded steps), we randomly select 1000 causally separable processes (rejecting the \( \approx 92.5\% \) of nonseparable ones) in the sample.

The map \[15\] preserves the Lebesgue measure, since it corresponds to reflections in \( \mathbb{R}^{d_W} \). Therefore, if upon applying \( \Omega \) to random causally separable matrices, there is a finite probability to obtain a valid, causally nonseparable process, this means that the set of causally nonseparable processes with a separable partial transpose is full dimensional (see figure 7).

References

[1] Bell J S 1964 On the Einstein Podolsky Rosen paradox Physics 1 195–200
[2] Clauser J F, Horne M A, Shimony A and Holt R A 1969 Proposed experiment to test local hidden-variable theories Phys. Rev. Lett. 23 880–4
[3] Oreshkov O, Costa F and Brukner Č 2012 Quantum correlations with no causal order Nat. Commun. 3 1092
[4] Branciard C, Araújo M, Feix A, Costa F and Brukner Č 2016 The simplest causal inequalities and their violation New J. Phys. 18 013008
[5] Baumeler A, Feix A and Wolf S 2014 Maximal incompatibility of locally classical behavior and global causal order in multiparty scenarios Phys. Rev. A 90 042106
[6] Chiribella G, D’Ariano G M, Perinotti P and Valiron B 2013 Quantum computations without definite causal structure Phys. Rev. A 88 022318
[7] Araújo M, Branciard C, Costa F, Feix A, Giarmatzi C and Brukner Č 2015 Witnessing causal nonseparability New J. Phys. 17 102001
[8] Oreshkov O and Giarmatzi C Causal and causally separable processes (arXiv:1506.05449)
[9] Werner R F 1989 Quantum states with Einstein–Podolsky–Rosen correlations admitting a hidden-variable model Phys. Rev. A 40 4277–81
[10] Brukner Č 2014 Quantum causality Nat. Phys. 10 259–63
[11] Smith R I. 1980 Monte Carlo procedures for generating random feasible solutions to mathematical programs A Bulletin of the ORSA/ TIMS Joint National Meeting (Washington, DC) vol 101
[12] Smith R I. 1984 Efficient monte carlo procedures for generating points uniformly distributed over bounded regions Oper. Res. 32 1296–308
[13] Brukner Č 2015 Bounding quantum correlations with indefinite causal order New J. Phys. 17 083034
[14] Chiribella G, D’Ariano G M and Perinotti P 2009 Theoretical framework for quantum networks Phys. Rev. A 80 022339
[15] Choi M–D 1975 Completely positive linear maps on complex matrices Linear Algebra Appl. 10 285–90
[16] Jamiołkowski A 1972 Linear transformations which preserve trace and positive semidefiniteness of operators Rep. Math. Phys. 3 275–8
[17] Gutoski G and Watrous J 2007 Toward a general theory of quantum games Proc. 39th Annual ACM Symposium on Theory of Computing, STOC’07 (New York: ACM) pp 565–74
[18] Chiribella G, D’Ariano G M and Perinotti P 2008 Transforming quantum operations: quantum supermaps Europhys. Lett. 83 30004
[19] Leifer M S and Spekkens R W 2015 Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference Phys. Rev. A 88 052130
[20] Chruściński D and Sarbicki G 2014 Entanglement witnesses: construction, analysis and classification J. Phys. A: Math. Theor. 47 483001
[21] Acín A, Brunner N, Gisin N, Massar S, Pirandola S and Scarani V 2007 Device-independent security of quantum cryptography against collective attacks Phys. Rev. Lett. 98 230501
[22] Procopio L M, Moşanăti A, Araújo M, Costa F, Alonso Calafell I, Dowd E G, Hamel D R, Rozema L A, Brukner Č and Walther P 2015 Experimental superposition of orders of quantum gates Nat. Commun. 6 7913
[23] Davies E B and Lewis J T 1970 An operational approach to quantum probability Commun. Math. Phys. 17 239–60
[24] Araújo M, Feix A, Navascués M and Brukner Č 2016 A purification postulate for quantum mechanics with no causal order (in preparation)
[25] Lovász L 1999 Hit-and–run mixes fast Math. Program. 86 443–61