

A no-go theorem for superpositions of causal orders

Fabio Costa

Centre for Engineered Quantum Systems, School of Mathematics and Physics, The University of Queensland, St Lucia, QLD 4072, Australia

The causal order of events need not be fixed: whether a bus arrives before or after another at a certain stop can depend on other variables—like traffic. Coherent quantum control of causal order is possible too and is a useful resource for several tasks. However, quantum control implies that a controlling system carries the which-order information—if the control is traced out, the order of events remains in a probabilistic mixture. Can the order of two events be in a pure superposition, uncorrelated with any other system? Here we show that this is not possible for a broad class of processes: a pure superposition of any pair of Markovian, unitary processes with equal local dimensions and different causal orders is not a valid process, namely it results in non-normalised probabilities when probed with certain operations. The result imposes constraints on novel resources for quantum information processing and on possible processes in a theory of quantum gravity.

1 Introduction

Quantum superpositions can be viewed as generalisations of classical probabilities: if classically we can be uncertain between two alternatives, assigning to each a probability, quantumly we should be able to consider a superposition of the two, replacing probabilities with complex amplitudes. Feynman’s sum-over-histories approach famously leverages this intuition [1]. Despite this view, one typically considers superpositions of *states*, whereas classical probabilities can be assigned to *any* logical statement. To what extent is it possible to generalise the superposition principle, beyond its original range of applicability? Providing a general, principled answer to

Fabio Costa: f.costa@uq.edu.au

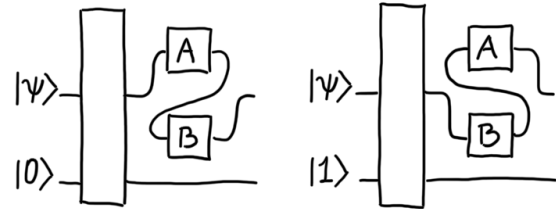


Figure 1: **The quantum switch.** The state of a control system determines the order of two operations on a target system.

this question appears problematic because, unlike classical probabilities, quantum superpositions have no clear interpretation in terms of subjective lack of knowledge.

Of particular interest is the case of causal relations. A main motivation comes from combining the principles of quantum theory and general relativity. In a regime where spacetime itself becomes subject to the laws of quantum mechanics, we expect that causal structure, too, will become nonclassical [2–4]. A second, more practical, motivation is that quantum causal structures can also be realised experimentally (on a classical spacetime background) [5–12] possibly leading to an advantage in solving computational and information-theoretic tasks [13–21].

Despite much recent research, the scenarios considered so far do not support a direct interpretation as superpositions of causal orders. The most discussed example is the so-called “quantum switch” [13], Fig. 1, where a control system determines the order in which a set of operations act on a target. Preparing the control in a superposition produces a nonclassical causal structure for the order of operations. Although sometimes colloquially described as a “superposition of causal orders”, the quantum switch is in fact more appropriately interpreted as ‘entanglement’ between causal relations and the control system. Indeed, if the control is discarded, the quantum switch is indistinguishable from a probabilistic mixture of

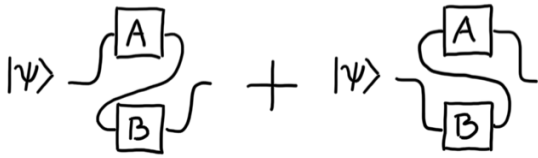


Figure 2: **Pure superposition of causal orders.** Is it possible to superpose the causal order of two events, without the aid of a control system?

causal orders. Only by measuring the control can one verify the nonclassicality of the order of operations. Probabilistic protocols, where one obtains a superposition of orders through postselection, generally succeed with probability less than one, making them not commensurable with deterministic resources.

Is a ‘pure’, deterministic superposition of causal orders possible (Fig. 2), not relying on any additional control system? Perhaps, this would provide a novel resource, enabling tasks not achievable by the quantum switch. It might also suggest types of processes arising in a theory of quantum gravity.

Here, we propose a definition of superpositions of causal orders and present a negative result, ruling out the most natural candidates. We show that for unitary, Markovian processes where all systems have equal dimension superpositions of two different causal orders are not possible, in the sense that they do not constitute valid, deterministic processes. In particular, the result applies to simple sequences of operations on a target system—as in the switch—showing that an additional system is necessary for performing such operations in an indefinite causal order. It remains an open question whether a superposition is possible under weaker conditions.

2 Superpositions of unitaries

Before considering superpositions of different causal orders, it is useful to look at a simpler case: superpositions of unitary transformations connecting an event in the past—a state preparation—to one in the future—a measurement. Given two unitaries U_j , $j = 0, 1$, it is natural to define their superposition as a linear combination $\alpha U_0 + \beta U_1$ for some complex numbers α , β , interpreted as probability amplitudes. However, such a ‘superposition’ is not necessarily a unitary operator. For example, no linear com-

bination of $U_0 = \mathbb{1}$ and $U_1 = \frac{\sigma_x + i\sigma_y}{\sqrt{2}}$ (where σ_x , σ_y , σ_z denote the Pauli matrices) is unitary, for *any* complex amplitudes $\alpha \neq 0$, $\beta \neq 0$.

One could interpret arbitrary linear combinations of unitaries in the following way [22]: Introduce the ‘controlled unitary’ $U_0 \otimes |0\rangle + U_1 \otimes |1\rangle$, which acts on an additional ‘control’ system and where we use the shorthand $[\psi] = |\psi\rangle\langle\psi|$. Starting from a state of the form $|\psi\rangle (\alpha|0\rangle + \beta|1\rangle)$, and postselecting the control on $(|0\rangle + |1\rangle)/\sqrt{2}$, the target system is left into $(\alpha U_0 + \beta U_1)|\psi\rangle/\sqrt{2}$, which one can interpret as resulting from a superposition of unitaries. The problem with this interpretation is that the postselection step requires obtaining one out of a set of possible measurement outcomes, which in general only succeeds with some probability that depends on the initial state. As such a probability can be very small, or even vanish, realising a process through postselection can nullify the associated advantages. Furthermore, postselection can induce apparent signalling, making postselected processes unsuitable for studies of causal relations¹. For these reasons, here we are only interested in processes that can be realised *without* postselection.

We see that, for processes, superposition is not a universal possibility, as it is for states. However, there are unitaries for which superpositions *are* possible. For example, the *Hadamard gate* $H := \frac{\sigma_x + \sigma_z}{\sqrt{2}}$ is a rightful superposition of σ_x and σ_z , not requiring any postselection. H is a ubiquitous resource in quantum information and computation, precisely for its ability to convert basis states into superpositions. Therefore, even if we do not expect to be able to superpose arbitrary processes with different causal orders, we can still ask whether *any* such a superposition is possible.

3 General quantum processes

As we will see, superpositions of causal orders become meaningful for processes connecting more than two events. These can be described conveniently in the so-called “process matrix formalism” [23] (see also other closely related frame-

¹For example, one can place bets on a roulette and consider only the winning spins, effectively predicting the roulette’s outcome in each of the postselected spins. This is unlikely to be a remunerative strategy in any casino.

works [24–28]). A simple example of a multi-event process is one where a sequence of unitaries U_1, U_2, \dots connects a set of events at times t_1, t_2, \dots . At each time t_j , after unitary U_{j-1} and before unitary U_j , it is in principle possible to perform arbitrary measurements or transformations on the system. The particular operation taking place at a given time is what we call an ‘event’, while the set of unitaries connecting the events is the ‘process’. More generally, we consider a scenario where the labels that identify possible events are not necessarily associated with time instants (for example, events might be delocalised in time [29, 30] or take place on a nonclassical background [4]). A single ‘event location’ with label A (also referred to as “region”, “party”, or “laboratory”) is therefore identified with the set of possible operations from an input to an output Hilbert space, \mathcal{H}_I^A and \mathcal{H}_O^A , respectively.

The most general operation [31] is described by a Completely Positive (CP) trace non-increasing map $\mathcal{M}^A : A_I \rightarrow A_O$, where $A_{I,O} \equiv \mathcal{L}(\mathcal{H}_{I,O}^A)$ are the spaces of linear operators over the respective Hilbert spaces. A deterministic operation, i.e., one that happens with probability one, is CP and trace preserving (CPTP): $\text{tr} \circ \mathcal{M}^A = \text{tr}$, where tr and \circ denote operator trace and function composition, respectively. A collection of CP maps $\{\mathcal{M}_a^A\}_a$, where a labels the measurement outcome, is called an *instrument* if $\sum_a \mathcal{M}_a^A$ is CPTP, which implies that there is unit probability that at least one of the outcomes will occur. It is convenient to represent CP maps as operators, using a version of the Choi-Jamiołkowski isomorphism [32, 33], $\mathcal{M}^A \mapsto M^A \in A_I \otimes A_O$:

$$M^A := \left[\sum_{jk} |j\rangle\langle k| \otimes \mathcal{M}^A(|j\rangle\langle k|) \right]^T, \quad (1)$$

where $\{|j\rangle\}_j$ is a basis of \mathcal{H}_I^A and T denotes transposition.

The local validity of quantum mechanics implies that the joint probability for outcomes a_1, \dots, a_n in regions A^1, \dots, A^n is given by [23, 34]

$$P(a_1, \dots, a_n) = \text{tr} \left[W \cdot \left(M_{a_1}^{A^1} \otimes \dots \otimes M_{a_n}^{A^n} \right) \right], \quad (2)$$

where $W \in A^1 \otimes \dots \otimes A^n$ is called the *process matrix* (or simply *process*) and we use the shorthand $A^j \equiv A_I^j \otimes A_O^j$. W encodes all information relevant to the possible events, such as initial state,

transformations, and causal relations. Positivity of probabilities (and the possibility to extend the local operations to additional systems) implies that the process matrix is positive semidefinite, $W \geq 0$, while normalisation requires

$$\text{tr} \left[W \cdot \left(\bar{M}^{A^1} \otimes \dots \otimes \bar{M}^{A^n} \right) \right] = 1 \quad (3)$$

for all CPTP maps $\bar{M}^{A^1}, \dots, \bar{M}^{A^n}$. Crucially, Eq. (3) imposes on W more constraints than the normalisation condition for states, $\text{tr} \rho = 1$. Therefore, whereas all positive semidefinite operators represent states (up to normalisation) the same is not true for process matrices. As we will see, it is this constraint that obstructs general superpositions of processes. More general processes, that do not satisfy the linear constraints implied by Eq. (3), can always be realised probabilistically, through postselection [25, 27, 35, 36]. However, as discussed earlier, we are only interested in processes that can be realised without postselection. Therefore, in the following, we will always assume processes to be deterministic (and condition (3) to be satisfied).

The key property of the formalism is that it encodes signalling correlations: a choice of instrument in a region can change the marginal probability for measurement outcomes in another region. We say that a process is *causally ordered* if there is a permutation σ of n elements such that $A^{\sigma(j)}$ cannot signal to any group of parties $\{A^{\sigma(k)}\}_k$ with $\sigma(k) < \sigma(j)$. Conversely, we say that W is *incompatible* with (the causal order induced by) σ if it is possible to signal from some $A^{\sigma(j)}$ to a group of parties $\{A^{\sigma(k)}\}_k$, all with $\sigma(k) < \sigma(j)$. We say that two processes W_1, W_2 are *differently ordered* if they are causally ordered relative to permutations σ_1, σ_2 , but incompatible with σ_2, σ_1 , respectively. We are interested in superpositions of differently ordered processes.

A related concept is that of *causally nonseparable process*, which formalises the notion of indefinite causal order [23, 37, 38]. A process is causally separable if it is causally ordered or a probabilistic mixture of causally ordered processes. A superposition of differently ordered processes would necessarily be causally nonseparable (because it would be extremal and not causally ordered). However, the existence of locally classical, causally nonseparable processes [39–42] indicates that the opposite is not true, making causal nonsepar-

ability unsuitable to characterise superpositions of causal orders.

4 Superpositions of pure processes

Process matrices are a generalisation of density matrices and, as such, cannot be directly superposed. In analogy to states, we can define superpositions for ‘pure’ processes, namely, for rank-1 process matrices, with $W = [\omega]$ for some *process vector* $|\omega\rangle \in \mathcal{H}_I^{A^1} \otimes \mathcal{H}_O^{A^1} \otimes \dots$. Given two pure process $[\omega_1]$, $[\omega_2]$, we define their superposition through a linear combination $|\omega\rangle = \alpha|\omega_1\rangle + \beta|\omega_2\rangle$, with $\alpha, \beta \in \mathbb{C}$, if $[\omega]$ is a *valid process matrix*. That is to say, if and only if

$$\langle \omega | \left(\bar{M}^{A^1} \otimes \dots \otimes \bar{M}^{A^n} \right) | \omega \rangle = 1 \quad (4)$$

$$\forall \text{ CPTP } \bar{M}^{A^1}, \dots, \bar{M}^{A^n}.$$

If all parties have only input spaces (or, equivalently, if all output spaces are trivial, namely they have dimension 1), the process matrix reduces to a state, and superposition of processes reduces to superposition of states. Another special case is a process that connects a region’s output, A_O , to another region’s input, B_I , of equal dimension. In this case, a pure process represents a unitary transformation (see, e.g., Appendix A of Ref. [43]), for which we use the notation

$$|U\rangle\rangle^{A_O B_I} := \sum_j |j\rangle\rangle^{A_O} \otimes (U|j\rangle\rangle)^{B_I} \quad (5)$$

(where $\{|j\rangle\rangle_j\}$ is the same basis that defines the Choi-Jamiołkowski isomorphism). Because of the linearity of the representation (5), a ‘superposition of unitary processes’ reduces to the ‘superposition of unitaries’ discussed earlier: $\alpha|U_1\rangle\rangle + \beta|U_2\rangle\rangle = |\alpha U_1 + \beta U_2\rangle\rangle$.

More generally, a pure process defines an *isometry* from all the output to all the input spaces: if all parties prepare a pure state, they all receive a pure state, related to the prepared ones through an isometry. In the particular case where the total input and output dimensions are equal, the process defines a unitary transformation from outputs to inputs [44]. We will use the term *unitary process* for such cases².

²Ref. [44] used the term “pure process” for what we call here unitary process.

5 Superpositions of causal orders

We are now in a position to discuss superpositions of differently ordered processes. Let us first analyse the quantum switch, to see what type of superposition it represents.

There are two different versions relevant to the present discussion. The first is a tripartite process, where $A^1 \equiv A$ and $A^2 \equiv B$ act on a target system, while a third party, $A^3 \equiv F$ (for “future”), receives the target system resulting from A and B ’s operations, as well as a control system. In this version, the initial state of the target is set to some fixed state $|\psi\rangle$, while the control is fixed to $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. The resulting process vector is [43]

$$|\omega_{\text{switch}}^{(3)}\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle^{F_c} |\psi\rangle^{A_I} |\mathbb{1}\rangle\rangle^{A_O B_I} |\mathbb{1}\rangle\rangle^{B_O F_t} + |1\rangle^{F_c} |\psi\rangle^{B_I} |\mathbb{1}\rangle\rangle^{B_O A_I} |\mathbb{1}\rangle\rangle^{A_O F_t} \right), \quad (6)$$

where F_c and F_t respectively denote the control and target subsystems of F . Here, F has only input space, while its output is trivial, as no influence from F to A and B is possible.

Expression (6) manifestly represents a superposition of two differently ordered processes, with the target system going first to A and then B or vice versa. However, seen as an isometry, the process maps the two-qubit space ($A_O B_O$) to the larger four-qubit space ($A_I B_I F_c F_t$). This means that the isometry does not describe a ‘pure transformation’, but also the specification of a state. The natural interpretation is that process (6) displays *entanglement* between the control system F_c and the causal order of A and B ; therefore, it does not represent a superposition of causal orders alone. Indeed, if we trace out the control system, the resulting reduced process is causally separable (i.e., a classical mixture of causally ordered ones) [37, 43].

The second version of the switch features an extra laboratory, $A^0 \equiv P$ (for ‘past’), where both the control and target systems can be prepared (corresponding to subsystems P_c , P_t , respectively). As P does not receive any state, its input space can be taken to be trivial. The resulting unitary process is

$$|\omega_{\text{switch}}^{(4)}\rangle = \left(|0\rangle^{P_c} |0\rangle^{F_c} |\mathbb{1}\rangle\rangle^{P_t A_I} |\mathbb{1}\rangle\rangle^{A_O B_I} |\mathbb{1}\rangle\rangle^{B_O F_t} + |1\rangle^{P_c} |1\rangle^{F_c} |\mathbb{1}\rangle\rangle^{P_t B_I} |\mathbb{1}\rangle\rangle^{B_O A_I} |\mathbb{1}\rangle\rangle^{A_O F_t} \right). \quad (7)$$

This process is the sum of two vectors, each responsible for a different causal order. However, the two vectors do not define individually valid processes, as they are not normalised for arbitrary CPTP maps. Therefore, process (7) does not represent a superposition of processes with different causal orders, but rather describes *quantum control* of the order of A and B .

As mentioned earlier, one can obtain a superposition of valid processes through postselection: By preparing P_c in state $|+\rangle$ and postselecting F_c in the same state, process (7) reduces to

$$|\omega_{\text{post}}\rangle = \frac{1}{2} \left(|\mathbb{1}\rangle\langle\mathbb{1}|^{P_t A_I} |\mathbb{1}\rangle\langle\mathbb{1}|^{A_O B_I} |\mathbb{1}\rangle\langle\mathbb{1}|^{B_O F_t} + |\mathbb{1}\rangle\langle\mathbb{1}|^{P_t B_I} |\mathbb{1}\rangle\langle\mathbb{1}|^{B_O A_I} |\mathbb{1}\rangle\langle\mathbb{1}|^{A_O F_t} \right). \quad (8)$$

This, however, is not a valid process vector (nor is it proportional to one), meaning that the probability to get the desired outcome $|+\rangle$ depends on the operations performed at A , B (and it can even vanish, if the operations anticommute). Therefore, $|\omega_{\text{post}}\rangle$ is not the superposition of orders we were looking for either.

The above observations lead us to the following:

Definition 1. A process vector $|\omega\rangle$ for parties P, A^1, \dots, A^n, F (where P and F have trivial input and output space, respectively) is a pure superposition of two causal orders if

- $|\omega\rangle$ represents a unitary process;
- there are two nonvanishing complex numbers α, β and two differently ordered, unitary process vectors $|\omega_1\rangle, |\omega_2\rangle$ for the same set of parties as $|\omega\rangle$ such that

$$|\omega\rangle = \alpha |\omega_1\rangle + \beta |\omega_2\rangle. \quad (9)$$

The presence of a past (P) and future (F) party in the definition follows from the fact that, in a causally ordered process, there must be an initial party, which cannot receive signals from any other, and a final party, which cannot signal to the rest. If the process is unitary, such parties must have trivial input and output space, respectively. As such parties can only be respectively first and last in any causally ordered process, the simplest candidate superposition of orders has to be four-partite, as in the unitary switch (7).

Although we do not know if a superposition of causal orders is possible in general, we can rule

it out for the broad class of *Markovian* processes. These describe evolution from each region to the next, without any memory carried over through an external environment across different steps. A unitary, Markovian process is simply a sequence of unitaries, as in the example discussed in Sec. 3.

A Markovian process matrix is the tensor product of processes describing the individual time evolutions from each region to the next [45–47]. Therefore, a unitary, Markovian process matrix compatible with the causal order σ has the form

$$|\omega\rangle = \bigotimes_{j=0}^n |U_j\rangle\langle\mathbb{1}|^{A_O^{(j)} A_I^{\sigma(j+1)}}, \quad (10)$$

where U_j are unitary matrices, we identify $A_O^0 \equiv P$, $A_I^{n+1} \equiv F$, and we assume that the permutation σ leaves first and last parties unchanged, $\sigma(0) = 0$, $\sigma(n+1) = n+1$.

Note that, for such a process, each party can always signal to any future one (conditioned on the intermediate parties performing appropriate operations). Therefore, a process of this form compatible with σ is necessarily incompatible with any $\sigma' \neq \sigma$.

Our core result is that, for the simplest four-partite case, a linear combination of two unitary, Markovian processes with different orders cannot satisfy the process vector normalisation constraints, Eq. (4). In particular, a CPTP map for the output-only space P is an arbitrary density matrix $\rho^P \geq 0$, with $\text{tr} \rho^P = 1$ (describing a state preparation), while the only CPTP map for the input-only space F is $\mathbb{1}^F$ (describing an arbitrary POVM measurement whose outcome is ignored). For $X = A, B$, a CPTP map is represented by a matrix $\xi^{X_I X_O} \geq 0$ such that $\text{tr}_{X_O} \xi^{X_I X_O} = \mathbb{1}^{X_I}$. The technical statement of our result, proved in the Appendix³, is then as follows:

Lemma 2. For every set of unitaries $\{U_j, V_j\}_{j=1,2,3}$, and complex numbers $\alpha, \beta \neq 0$, it is possible to find matrices $\rho, \xi, \eta \geq 0$, with $\text{tr} \rho = 1$, $\text{tr}_{A_O} \xi^{A_I A_O} = \mathbb{1}^{A_I}$, and

³While completing the current manuscript, the author was made aware of related results in Ref. [48]. In particular, Lemma 2 can be deduced from Corollary 5 in Ref. [48], although the consequence as a no go theorem for superpositions of orders is not discussed there. The argument and the proof in the present manuscript were previously presented at [49].

$\text{tr}_{B_O} \eta^{B_I B_O} = \mathbb{1}^{B_I}$, such that

$$\langle \omega | \left(\rho^P \otimes \xi^{A_I A_O} \otimes \eta^{B_I B_O} \otimes \mathbb{1}^F \right) | \omega \rangle \neq 1, \quad (11)$$

where $|\omega\rangle$ is defined by Eq. (9) and

$$|\omega_1\rangle = |U_1\rangle\rangle^{P A_I} |U_2\rangle\rangle^{A_O B_I} |U_3\rangle\rangle^{B_O F}, \quad (12)$$

$$|\omega_2\rangle = |V_1\rangle\rangle^{P B_I} |V_2\rangle\rangle^{B_O A_I} |V_3\rangle\rangle^{A_O F}. \quad (13)$$

This result rules out the simplest, most intuitive superpositions of causal orders. In particular, it proves that the control system is necessary for the quantum switch to be a valid process. Note that, because each step is unitary, all input and output spaces must have equal dimension⁴. This is not guaranteed for arbitrarily many parties. As it turns out, we need the additional assumption of equal dimensions to prove our general result:

Theorem 3. *Consider a set of parties P, A^1, \dots, A^n, F , where P has trivial input, F has trivial output, and all nontrivial input and output dimensions are equal. For every pair of differently ordered, unitary, Markovian process vectors $|\omega_1\rangle, |\omega_2\rangle$, and complex numbers $\alpha, \beta \neq 0$, the linear combination $|\omega\rangle = \alpha |\omega_1\rangle + \beta |\omega_2\rangle$ is not a valid process vector.*

Proof. We can prove the statement by reducing the general case to the $n = 2$ one. Indeed, let $\sigma_1 \neq \sigma_2$ be the permutations defining the causal orders of $|\omega_1\rangle$ and $|\omega_2\rangle$. As the two permutations are different, we can find $j \neq k$ such that $\sigma_1(j) < \sigma_1(k)$ and $\sigma_2(k) < \sigma_2(j)$. Inserting identity unitaries in all regions except j, k , we obtain the bipartite reduced processes [43] $|\tilde{\omega}_\# \rangle := \left(\bigotimes_{i \neq j, k} \langle \mathbb{1} |^{A_i A_i^o} \right) |\omega_\# \rangle$, with $|\omega_\# \rangle = |\omega\rangle, |\omega_1\rangle, |\omega_2\rangle$. Now, we have $|\tilde{\omega}\rangle = \alpha |\tilde{\omega}_1\rangle + \beta |\tilde{\omega}_2\rangle$, where $|\tilde{\omega}_1\rangle$ and $|\tilde{\omega}_2\rangle$ are unitary, Markovian, and differently ordered. Lemma 2 then implies that $|\tilde{\omega}\rangle$ cannot be a valid process, which in turn implies that $|\omega\rangle$ cannot be a valid process either. \square

6 Conclusions

We have seen that superpositions of processes with different causal orders are strongly constrained, if we require that the order of events

⁴The unitaries in Eq. (12) imply $d^P = d^{A_I}, d^{A_O} = d^{B_I}$, and $d^{B_O} = d^F$, while from Eq. (13) we get $d^P = d^{B_I}, d^{B_O} = d^{A_I}$, and $d^{A_O} = d^F$. Taken together, the two sets of equations imply that all dimensions are the same.

is not correlated with any additional system. In particular, it is not possible to superpose sequences of equal-dimension unitaries connecting events in different orders—arguably the most natural candidate superposition of causal orders. Relaxing the assumptions in the no-go theorem leads to some open possibilities: it is currently unclear whether it is possible to superpose unitary, differently ordered processes that do not satisfy Markovianity or where the local input and output spaces do not have equal dimensions. Another interesting direction is to consider superpositions of more than two processes. Furthermore, it is an intriguing possibility to develop a notion of ‘coherence’ of causal orders that applies to nonpure processes, possibly generalising corresponding resource theories for states [50, 51].

Acknowledgments

The author thanks Mateus Araújo, Cyril Branciard, Časlav Brukner, Matthew Palmer, and Magdalena Zych for useful discussions. This work was partially supported through an Australian Research Council Discovery Early Career Researcher Award (DE170100712). The University of Queensland (UQ) acknowledges the Traditional Owners and their custodianship of the lands on which UQ operates.

References

- [1] R. P. Feynman, “Space-Time Approach to Non-Relativistic Quantum Mechanics,” *Rev. Mod. Phys.* **20**, 367–387 (1948).
- [2] J. Butterfield and C. Isham, *Spacetime and the philosophical challenge of quantum gravity*, p. 33–89. Cambridge University Press. [arXiv:gr-qc/9903072](https://arxiv.org/abs/gr-qc/9903072).
- [3] L. Hardy, “Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure,” *J. Phys. A: Math. Gen.* **40**, 3081–3099 (2007).
- [4] M. Zych, F. Costa, I. Pikovski, and Č. Brukner, “Bell’s theorem for temporal order,” *Nat. Commun.* **10**, 3772 (2019).
- [5] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. A. Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, “Experimental superposition of orders of quantum gates,” *Nat. Commun.* **6**, 7913 (2015).

- [6] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, “Experimental verification of an indefinite causal order,” *Sci. Adv.* **3**, e1602589 (2017).
- [7] G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M. Zych, Č. Brukner, and P. Walther, “Experimental entanglement of temporal order,” *Quantum* **6**, 621 (2022).
- [8] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, “Indefinite Causal Order in a Quantum Switch,” *Phys. Rev. Lett.* **121**, 090503 (2018).
- [9] K. Goswami, Y. Cao, G. A. Paz-Silva, J. Romero, and A. G. White, “Increasing communication capacity via superposition of order,” *Phys. Rev. Research* **2**, 033292 (2020).
- [10] Y. Guo, X.-M. Hu, Z.-B. Hou, H. Cao, J.-M. Cui, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, and G. Chiribella, “Experimental Transmission of Quantum Information Using a Superposition of Causal Orders,” *Phys. Rev. Lett.* **124**, 030502 (2020).
- [11] K. Wei, N. Tischler, S.-R. Zhao, Y.-H. Li, J. M. Arrazola, Y. Liu, W. Zhang, H. Li, L. You, Z. Wang, Y.-A. Chen, B. C. Sanders, Q. Zhang, G. J. Pryde, F. Xu, and J.-W. Pan, “Experimental Quantum Switching for Exponentially Superior Quantum Communication Complexity,” *Phys. Rev. Lett.* **122**, 120504 (2019).
- [12] M. M. Taddei, J. Cariñe, D. Martínez, T. García, N. Guerrero, A. A. Abbott, M. Araújo, C. Branciard, E. S. Gómez, S. P. Walborn, L. Aolita, and G. Lima, “Computational Advantage from the Quantum Superposition of Multiple Temporal Orders of Photonic Gates,” *PRX Quantum* **2**, 010320 (2021).
- [13] G. Chiribella, G. M. D’Ariano, P. Perinotti, and B. Valiron, “Quantum computations without definite causal structure,” *Phys. Rev. A* **88**, 022318 (2013).
- [14] G. Chiribella, “Perfect discrimination of no-signalling channels via quantum superposition of causal structures,” *Phys. Rev. A* **86**, 040301(R) (2012).
- [15] T. Colnaghi, G. M. D’Ariano, S. Facchini, and P. Perinotti, “Quantum computation with programmable connections between gates,” *Phys. Lett. A* **376**, 2940–2943 (2012).
- [16] M. Araújo, F. Costa, and Č. Brukner, “Computational Advantage from Quantum-Controlled Ordering of Gates,” *Phys. Rev. Lett.* **113**, 250402 (2014).
- [17] A. Feix, M. Araújo, and Č. Brukner, “Quantum superposition of the order of parties as a communication resource,” *Phys. Rev. A* **92**, 052326 (2015).
- [18] P. A. Guérin, A. Feix, M. Araújo, and Č. Brukner, “Exponential Communication Complexity Advantage from Quantum Superposition of the Direction of Communication,” *Phys. Rev. Lett.* **117**, 100502 (2016).
- [19] D. Ebler, S. Salek, and G. Chiribella, “Enhanced Communication with the Assistance of Indefinite Causal Order,” *Phys. Rev. Lett.* **120**, 120502 (2018).
- [20] S. Salek, D. Ebler, and G. Chiribella, “Quantum communication in a superposition of causal orders,” [arXiv:1809.06655v2 \[quant-ph\]](https://arxiv.org/abs/1809.06655v2).
- [21] M. K. Gupta and U. Sen, “Transmitting quantum information by superposing causal order of mutually unbiased measurements,” [arXiv:1909.13125v1 \[quant-ph\]](https://arxiv.org/abs/1909.13125v1).
- [22] Y. Aharonov, J. Anandan, S. Popescu, and L. Vaidman, “Superpositions of time evolutions of a quantum system and a quantum time-translation machine,” *Phys. Rev. Lett.* **64**, 2965 (1990).
- [23] O. Oreshkov, F. Costa, and Č. Brukner, “Quantum correlations with no causal order,” *Nat. Commun.* **3**, 1092 (2012).
- [24] R. Oeckl, “A “general boundary” formulation for quantum mechanics and quantum gravity,” *Phys. Lett. B* **575**, 318–324 (2003).
- [25] Y. Aharonov, S. Popescu, J. Tollaksen, and L. Vaidman, “Multiple-time states and multiple-time measurements in quantum mechanics,” *Phys. Rev. A* **79**, 052110 (2009).
- [26] J. Cotler, C.-M. Jian, X.-L. Qi, and F. Wilczek, “Superdensity operators for spacetime quantum mechanics,” *J. High Energ. Phys.* **2018**, 93 (2018).
- [27] R. Silva, Y. Guryanova, A. J. Short, P. Skrzypczyk, N. Brunner, and S. Popescu, “Connecting processes with indefinite causal order and multi-time quantum states,” *New J. Phys.* **19**, 103022 (2017).
- [28] J. Barrett, R. Lorenz, and O. Oreshkov, “Quantum Causal Models,” [arXiv:1906.10726v2 \[quant-ph\]](https://arxiv.org/abs/1906.10726v2).
- [29] P. A. Guérin and Č. Brukner, “Observer-dependent locality of quantum events,” *New J. Phys.* **20**, 103031 (2018).
- [30] O. Oreshkov, “Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics,” *Quantum* **3**, 206 (2019).

- [31] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement*. Cambridge University Press, 2011.
- [32] M.-D. Choi, “Completely positive linear maps on complex matrices,” *Linear Algebra Appl.* **10**, 285–290 (1975).
- [33] A. Jamiołkowski, “Linear transformations which preserve trace and positive semidefiniteness of operators,” *Rep. Math. Phys.* **3**, 275–278 (1972).
- [34] S. Shrapnel, F. Costa, and G. Milburn, “Updating the Born rule,” *New J. Phys.* **20**, 053010 (2018).
- [35] O. Oreshkov and N. J. Cerf, “Operational quantum theory without predefined time,” *New J. Phys.* **18**, 073037 (2016).
- [36] S. Milz, F. A. Pollock, T. P. Le, G. Chiribella, and K. Modi, “Entanglement, non-Markovianity, and causal non-separability,” *New J. Phys.* **20**, 033033 (2018).
- [37] O. Oreshkov and C. Giarmatzi, “Causal and causally separable processes,” *New J. of Phys.* **18**, 093020 (2016).
- [38] J. Wechs, A. A. Abbott, and C. Branciard, “On the definition and characterisation of multipartite causal (non)separability,” *New J. of Phys.* **21**, 013027 (2019).
- [39] Ä. Baumeler, A. Feix, and S. Wolf, “Maximal incompatibility of locally classical behavior and global causal order in multi-party scenarios,” *Phys. Rev. A* **90**, 042106 (2014).
- [40] Ä. Baumeler and S. Wolf, “The space of logically consistent classical processes without causal order,” *New J. of Phys.* **18**, 013036 (2016).
- [41] Ä. Baumeler, F. Costa, T. C. Ralph, S. Wolf, and M. Zych, “Reversible time travel with freedom of choice,” *Class. Quantum Grav.* **36**, 224002 (2019).
- [42] G. Tobar and F. Costa, “Reversible dynamics with closed time-like curves and freedom of choice,” *Classical and Quantum Gravity* **37**, 205011 (2020).
- [43] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, “Witnessing causal nonseparability,” *New J. Phys.* **17**, 102001 (2015).
- [44] M. Araújo, A. Feix, M. Navascués, and Č. Brukner, “A purification postulate for quantum mechanics with indefinite causal order,” *Quantum* **1**, 10 (2017).
- [45] F. Costa and S. Shrapnel, “Quantum causal modelling,” *New J. of Phys.* **18**, 063032 (2016).
- [46] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, “Operational Markov Condition for Quantum Processes,” *Phys. Rev. Lett.* **120**, 040405 (2018).
- [47] C. Giarmatzi and F. Costa, “Witnessing quantum memory in non-Markovian processes,” *Quantum* **5**, 440 (2021).
- [48] W. Yokojima, M. T. Quintino, A. Soeda, and M. Murao, “Consequences of preserving reversibility in quantum superchannels,” *Quantum* **5**, 441 (2021).
- [49] F. Costa, “A no-go theorem for superpositions of causal order.” *Space-time and Information*, Manitoulin Island, Ontario, Canada, 2017.
- [50] T. Theurer, N. Killoran, D. Egloff, and M. B. Plenio, “Resource Theory of Superposition,” *Phys. Rev. Lett.* **119**, 230401 (2017).
- [51] F. Bischof, H. Kampermann, and D. Bruß, “Resource Theory of Coherence Based on Positive-Operator-Valued Measures,” *Phys. Rev. Lett.* **123**, 110402 (2019).

A Proof of Lemma 2

Here we provide the proof of Lemma 2, which we reformulate for convenience:

Lemma. *For every set of unitaries $\{U_j, V_j\}_{j=1,2,3}$, and complex numbers $\alpha, \beta \neq 0$, it is possible to find*

$$\chi := \rho^P \otimes \xi^{A_I A_O} \otimes \eta^{B_I B_O} \otimes \mathbb{1}^F, \quad \text{with} \quad (14)$$

$$\text{tr} \rho^P = 1, \text{tr}_{A_O} \xi^{A_I A_O} = \mathbb{1}^{A_I}, \text{tr}_{B_O} \eta^{B_I B_O} = \mathbb{1}^{B_I},$$

such that

$$\langle \chi \rangle := \langle \omega | \chi | \omega \rangle \neq 1, \quad (15)$$

where $|\omega\rangle = \alpha |\omega_1\rangle + \beta |\omega_2\rangle$ and

$$|\omega_1\rangle = |U_1\rangle\rangle^{P A_I} |U_2\rangle\rangle^{A_O B_I} |U_3\rangle\rangle^{B_O F}, \quad (16)$$

$$|\omega_2\rangle = |V_1\rangle\rangle^{P B_I} |V_2\rangle\rangle^{B_O A_I} |V_3\rangle\rangle^{A_O F}. \quad (17)$$

As $|\omega_1\rangle$ and $|\omega_2\rangle$ are normalised process vectors, we have $\langle \omega_1 | \chi | \omega_1 \rangle = \langle \omega_2 | \chi | \omega_2 \rangle = 1$ for all χ of the form (14), so that

$$\langle \chi \rangle = |\alpha|^2 + |\beta|^2 + 2\text{Re}(\alpha^* \beta \langle \omega_1 | \chi | \omega_2 \rangle). \quad (18)$$

For $|\omega\rangle$ to be a normalised process, we need $\langle \chi \rangle = 1$ for all χ of the form (14), which requires $\text{Re}(\alpha^* \beta \langle \omega_1 | \chi | \omega_2 \rangle)$ to be constant over χ . Writing $\alpha^* \beta = e^{i\phi} |\alpha^* \beta|$, we obtain that the cross term $\text{Re}(e^{i\phi} \langle \omega_1 | \chi | \omega_2 \rangle)$ should be constant for all χ .

In order to simplify the cross term, we restrict the CPTP maps to a pure state preparation $|\psi\rangle$ for P and unitaries R, S for A and B , respectively, so that

$$\chi = [\psi^*]^P \otimes [[R^*]]^{A_I A_O} \otimes [[S^*]]^{B_I B_O} \otimes \mathbb{1}^F, \quad (19)$$

with the notation $[[A]] \equiv |A\rangle\rangle\langle\langle A|$ and where the complex conjugations in $|\psi^*\rangle$, R^* , S^* , result from the transposition in our definition of the Choi representation for local operations, Eq. (1). After some linear algebra, we arrive at

$$\langle \omega_1 | \chi | \omega_2 \rangle = \langle \psi | U_1^\dagger R^\dagger U_2^\dagger S^\dagger U_3^\dagger V_3 R V_2 S V_1 | \psi \rangle. \quad (20)$$

The crucial step of the proof is to simplify this expression through an appropriate choice of local operations. We can do this using the substitutions

$$S \mapsto S' = U_3^\dagger V_3 U_2^\dagger V_1 S V_1^\dagger, \quad (21)$$

$$R \mapsto R' = U_2^\dagger V_1 R V_1^\dagger U_2 V_3^\dagger U_3 V_2^\dagger, \quad (22)$$

under which the cross term takes the form

$$e^{i\phi} \langle \omega_1 | \chi | \omega_2 \rangle = \langle \psi | V T | \psi \rangle, \quad (23)$$

$$V = e^{i\phi} U_1^\dagger V_2 U_3^\dagger V_3 U_2^\dagger V_1, \quad (24)$$

$$T = R^\dagger S^\dagger R S. \quad (25)$$

Now the normalisation of $|\omega\rangle$ reduces to the condition that, given V as in Eq. (24), $\text{Re}(\langle \psi | V T | \psi \rangle)$ has to be constant for all unitaries T of the form (25) and all $|\psi\rangle$. This implies that $V T + T^\dagger V^\dagger$ should be proportional to the identity or, equivalently,

$$V T V + V^\dagger = \lambda T, \quad (26)$$

where λ is a number that depends on α and β :

$$\lambda = \frac{1 - |\alpha|^2 - |\beta|^2}{2|\alpha^*\beta|}. \quad (27)$$

To show that no unitary V can satisfy condition (26) for all R, S , it is sufficient to take block-diagonal unitaries R, S of the form $\sigma \oplus \mathbb{1}$, where σ acts on a two-dimensional subspace spanned by eigenstates of V (so that the restriction of V to that subspace is also unitary). In other words, it is sufficient to show that condition (26) cannot be satisfied by any single-qubit unitary V . We will omit the $\oplus \mathbb{1}$ for notational convenience.

Let us consider first the case where $\lambda = 0$; that is, a superposition of processes with normalised amplitudes $|\alpha|^2 + |\beta|^2 = 1$. Condition (26) reduces to

$$TVT + V^\dagger = 0, \quad (28)$$

which has to hold for all T of the form (25). We have to find a set of unitaries T of that form such that a single unitary V cannot satisfy Eq. (28) for all of them. First, choose $R = S = \mathbb{1}$ to set $T = \mathbb{1}$, so that Eq. (28) gives

$$V + V^\dagger = 0. \quad (29)$$

Substituting this into condition (28), we now have

$$TVT = V. \quad (30)$$

We can see that this is not possible by choosing R_j, S_j such that $i\sigma_j = T_j = R_j^\dagger S_j^\dagger R_j S_j$, with $j = x, y, z$ labelling the three Pauli matrices. (For example, for $j = x$, we can choose $R_x = \sigma_y, S_x = \frac{\mathbb{1} + i\sigma_x}{\sqrt{2}}$.) Now, V has to satisfy

$$\sigma_j V \sigma_j = -V, \quad j = x, y, z. \quad (31)$$

This can only hold for $V = 0$, which is not a unitary.

Let us now go back to the $\lambda \neq 0$ case. As we have seen above, we can choose R and S such that $T = i\sigma_x$. But we can also set $T = -i\sigma_x$ by choosing $R = \sigma_y, S = \frac{\mathbb{1} - i\sigma_x}{\sqrt{2}}$. Thus, we get the two equations

$$-\sigma_x V \sigma_x + V^\dagger = i\lambda \sigma_x, \quad (32)$$

$$-\sigma_x V \sigma_x + V^\dagger = -i\lambda \sigma_x. \quad (33)$$

These can be satisfied simultaneously only for $\lambda = 0$, which brings us back to the previous case, concluding the proof. \square

⁵The general construction is $R_j = \sigma_k$ (with $k \neq j$) and $S_j = \frac{\mathbb{1} + i\sigma_j}{\sqrt{2}}$. This gives $R_j S_j = \frac{\sigma_k - \sum_l \epsilon_{kjl} \sigma_l}{\sqrt{2}}$ and $R_j^\dagger S_j^\dagger = \frac{\sigma_k + \sum_l \epsilon_{kjl} \sigma_l}{\sqrt{2}}$. As $(\sigma_k \sigma_j)^2 = -\mathbb{1}$ for $k \neq j$, we get $R_j^\dagger S_j^\dagger R_j S_j = \frac{1}{2}(\sigma_j - \sigma_k \sigma_j \sigma_k) = i\sigma_j$.