FAST TRACK COMMUNICATION • OPEN ACCESS

Composition rules for quantum processes: a no-go theorem

To cite this article: Philippe Allard Guérin *et al* 2019 *New J. Phys.* **21** 012001

View the article online for updates and enhancements.

IOP ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

New Journal of Physics

The open access journal at the forefront of physics

CrossMark

FAST TRACK COMMUNICATION

Composition rules for quantum processes: a no-go theorem

Philippe Allard Guérin[1,2,3] , Marius Krumm[1,2], Costantino Budroni[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), Austrian Academy of Sciences, Boltzmanngasse 3, A-1090 Vienna, Austria
3 Author to whom any correspondence should be addressed.

E-mail: philippe.guerin@univie.ac.at

Abstract

A quantum process encodes the causal structure that relates quantum operations performed in local laboratories. The process matrix formalism includes as special cases quantum mechanics on a fixed background space-time, but also allows for more general causal structures. Motivated by the interpretation of processes as a resource for quantum information processing shared by two (or more) parties, with advantages recently demonstrated both for computation and communication tasks, we investigate the notion of composition of processes. We show that under very basic assumptions such a composition rule does not exist. While the availability of multiple independent copies of a resource, e.g. quantum states or channels, is the starting point for defining information-theoretic notions such as entropy (both in classical and quantum Shannon theory), our no-go result means that a Shannon theory of general quantum processes will not possess a natural rule for the composition of resources.

1. Introduction

Experimental tests with elementary quantum systems, most notably Bell tests, radically challenge the very notions of physical reality and cause-effect relations [1, 2]. Notwithstanding such fundamental novel effects, quantum mechanics still assumes a definite causal order of events. Namely, given two events, i.e. two operations performed locally in two quantum laboratories, say A and B, we always assume that they are either time-like separated, hence, A cannot signal to B or vice versa, or they are space-like separated, hence, they cannot signal in either direction.

Motivated by the problem of quantum gravity, operational formalisms have been proposed for computing the joint probabilities for the outcome of local experiments, without the assumption of a fixed space–time background [3–8]. *Process matrices* [6] are introduced as the most general class of multilinear mappings of local quantum operations into probability distributions. The process matrix formalism provides a unified description of causally ordered quantum mechanics (quantum states and quantum channels), but also includes experimentally relevant non-causal processes such as the quantum switch [7, 9–14]. Furthermore, the formalism predicts novel and potentially observable phenomena, such as the violation of so-called *causal inequalities* [6, 14–17].

Moreover, it has been proven that such processes are able to provide advantages for quantum information processing tasks, both for computation and communication [7, 18–24]. One would, then, expect that a theory of information can be developed also for processes. Such a theory would deal with, e.g., rates of information compression and communication, i.e. a process-analog of the classical and quantum Shannon theory. A fundamental assumption in classical and quantum Shannon theory [25, 26] is the availability of multiple independent copies of a resource (for example a classical source of random variables, a quantum state, or a channel), which is at the basis of the definition of information-theoretic entropy, i.e. Shannon or von Neumann entropy. To be more concrete, in the example of Schumacher's compression [25, 27], the optimal data compression of n samples of an independent and identically distributed quantum source ρ into $nS(\rho) + \delta$ qubits (with $\delta \to 0$ for $n \to \infty$), and the subsequent transmission, can be achieved only if the sender can act globally on multiple copies of the quantum state in which the information is encoded.

A natural question then arises, namely, whether a process matrix can be understood as a resource available in multiple (possibly identical) copies to experimenters, similarly to the example of Schumacher's compression above. Answering this question will provide us with deeper insight into the nature of process matrices. For instance, if we consider an experimental realization of a process, e.g. consisting of a sequence of optical elements as in photonic experiments [10, 11], one can easily imagine that it is possible to create two identical copies of the setup, and share them among the two parties. Alternatively, if one imagines that a process matrix does not only represent an experimental setup, but also the space-time structure [28–30], then it is harder to imagine how two 'copies of spacetime' may be shared between the two parties. More generally, such a composition rule should not be only about identical copies, but it should also allows us to combine different processes.

It is important, at this point, to distinguish two different scenarios and their corresponding composition rules. On the one hand, one may simply ask what is the rule for composing different processes *independently*, with the requirement that experimenters act locally on each copy of the process; this rule is given by the tensor product. On the other hand, going back to the example of Schumacher's compression protocol, one may require that a single experimenter (or many experimenters for multipartite systems) has access to multiple copies of a process, in order to perform a protocol that involves global operations. We will see that the latter notion is incompatible with the definition of a process.

For quantum states, quantum channels, or for any collection of processes with the same definite causal order [31, 32], the parallel composition can be described by the tensor product. However, it is known that a parallel composition of process matrices via the tensor product can fail [33], as the resulting process matrix contains causal 'double-loops' [6], which give rise to the 'grandfather paradox', or equivalently, to unnormalised probabilites.

In this work, we show that under weak assumptions (bilinearity, every output is a valid process matrix, reduction to the usual tensor product for definite causal structure) there exists no composition that allows the experimenters to have access to multiple shared processes. This result means that many information theoretic protocols relying on many copies of a resource have no straightforward generalization to process matrices.

2. Preliminary notions

The most general operation that can be performed on a quantum system is represented by a quantum instrument, namely, a collection $\{\mathcal{M}_a\}_a$ of completely positive trace-nonincreasing maps that sum up to a trace-preserving map $\mathcal{M} = \sum_a \mathcal{M}_a$. An operation represented by the instrument $\{\mathcal{M}_a\}_a$ will give an output a with probability $P(a) = \mathrm{tr}[\mathcal{M}_a(\rho)]$ and transformation of the state $\rho \mapsto \mathcal{M}_a(\rho)/P(a)$. We admit the possibility of an input x, and label the corresponding operations as $\{\mathcal{M}_{a|x}\}_{a,x}$. Such maps can be represented as matrices via the Choi-Jamiołkowski isomorphism [34, 35]

$$\mathcal{M}_{a|x} \mapsto M_{a|x} = \sum_{ij} |i\rangle\langle j|^{A_\mathrm{I}} \otimes \mathcal{M}_{a|x}(|i\rangle\langle j|)^{A_\mathrm{O}}. \tag{1}$$

We will call $M_{a|x}$ the Choi matrix of $\mathcal{M}_{a|x}$ [4]. Consider a set of local operations, i.e. Choi matrices, $\{M_{a|x}^A\}_{a,x}$ and $\{M_{b|y}^B\}_{b,y}$, associated with Alice's and Bob's laboratories, where A denotes Alice's input–output space $\mathcal{H}_{A_\mathrm{I}} \otimes \mathcal{H}_{A_\mathrm{O}}$, and similarly for B. A process W is understood as the most general linear mapping of such operations into probabilities, which can be represented using the trace inner product as

$$p(ab|xy) = \mathrm{tr}[(M_{a|x}^A \otimes M_{b|y}^B)W^T], \tag{2}$$

where T denotes the transposition in the computational basis. A visual representation of this probability rule is given in figure 1. In order to obtain valid probabilities, i.e. non-negative numbers summing up to one, for arbitrary operations $\{M_{a|x}^A\}_{a,x}$, $\{M_{b|y}^B\}_{b,y}$ (including operations that involve shared entangled ancillary systems), it can be proven [9] that the following constraints must be satisfied

$$W \geqslant 0, \tag{3}$$

$$\mathrm{tr}\,W = d_O = d_{A_\mathrm{O}} d_{B_\mathrm{O}}, \tag{4}$$

$$_{B_\mathrm{I}B_\mathrm{O}}W = {}_{A_\mathrm{O}B_\mathrm{I}B_\mathrm{O}}W, \tag{5}$$

$$_{A_\mathrm{I}A_\mathrm{O}}W = {}_{A_\mathrm{I}A_\mathrm{O}B_\mathrm{O}}W, \tag{6}$$

$$W = {}_{B_\mathrm{O}}W + {}_{A_\mathrm{O}}W - {}_{A_\mathrm{O}B_\mathrm{O}}W, \tag{7}$$

---

[4] Alternatively, one could define $\mathcal{M}_{a|x}$ with a global transposition t, taken w.r.t. the $\{|ij\rangle\}_{ij}$ basis, as in [9]. This allows one to write the process matrix associated to a quantum state ρ shared between the parties simply as $\rho_\mathrm{I} \otimes \mathbb{1}_\mathrm{O}$, instead of $\rho_\mathrm{I}^t \otimes \mathbb{1}_\mathrm{O}$. We will not use this convention here.
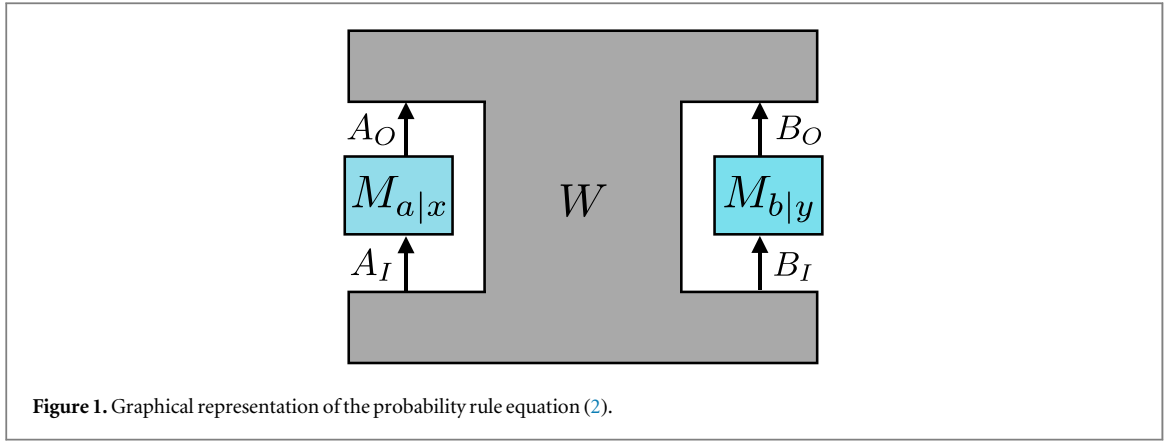
Figure 1. Graphical representation of the probability rule equation (2).

where $_X W := \frac{\mathbb{1}^X}{d_X} \otimes \mathrm{tr}_X W$. The linear constraints in equations (4)–(7) can be written in a more compact form as

$$L_V(W) = W, \tag{8}$$

where L_V is the projector onto the subspace of operators in $\mathcal{L}(\mathcal{H}_{AB})$ that satisfy equations (5)–(7). We will denote such a linear subspace as $L_V(\mathcal{L}(\mathcal{H}_{AB}))$. This projector enforces the normalization of probabilites, and can be interpreted as preventing the paradoxes that would occur in processes with 'causal loops' [6]. It is also convenient to define $\mathcal{W} \subset \mathcal{L}(\mathcal{H}_{AB})$ as the set of matrices that satisfy the conditions in equations (3)–(7), and similarly $\mathcal{W}'$ for the spaces $A'B' := \mathcal{H}_{A_{I'}} \otimes \mathcal{H}_{A_{O'}} \otimes \mathcal{H}_{B_{I'}} \otimes \mathcal{H}_{B_{O'}}$. If $_{B_O} W = W$, one can show that Bob cannot signal to Alice, i.e. $p(a|x, y) = p(a|x, y')$ for all a, x, y, y', we denote it as $A \preceq B$ and we say that the process is causally ordered [9]. Similarly, the case $_{A_O} W = W$ corresponds to the opposite causal order and it is denoted as $B \preceq A$. If $_{A_O B_O} W = W$ we have at the same time $A \preceq B$ and $B \preceq A$, then W represents a bipartite quantum state and we have no-signaling in both directions.

Similarly, in the case of N parties $A^1, \ldots, A^N$, linear constraints can be written in the compact form [9]

$$L_{V_N}(W) := \left[1 - \prod_{i=1}^{N} (1 - A_O^i + A_I^i A_O^i) + \prod_{i=1}^{N} A_I^i A_O^i \right] W = W, \tag{9}$$

where the index i runs through the different parties. Notice that if $W = W^1 \otimes W^2$, then the set $\{1, \ldots, N\}$ can be split as $\chi_1 \cup \chi_2$, with $\chi_1 \cap \chi_2 = \varnothing$, where χ_k indexes the parties appearing in W^k. Then

$$L_{V_N}(W_1 \otimes W_2) = W_1 \otimes W_2 \iff$$

$$\left[1 - \prod_{i \in \chi_1} (1 - A_O^i + A_I^i A_O^i) + \prod_{i \in \chi_1} A_I^i A_O^i \right] W_1 = W_1 \text{ and } \left[1 - \prod_{i \in \chi_2} (1 - A_O^i + A_I^i A_O^i) + \prod_{i \in \chi_2} A_I^i A_O^i \right] W_2 = W_2. \tag{10}$$

2.1. Examples

The process matrix formalism allows one to treat quantum states, quantum channels, and even situations where the causal order is indefinite, in a unified way. For example, the process matrix associated to a quantum state ρ can be described as a single party process matrix, as $W = \rho^{A_I} \otimes \mathbb{1}^{A_O}$. The process matrix associated to N spatially separated copies of the state is a N-partite process $W = \prod_{i=1}^{N} \rho^{A_I^i} \otimes \mathbb{1}^{A_O^i}$, where each of the A_I^i and A_O^i are isomorphic. However, one could also consider the same W as a global single party process, with input Hilbert space $A_I = \prod_i A_I^i$, and output Hilbert space $A_O = \prod_i A_O^i$.

A quantum channel $\mathcal{C} \colon \mathcal{L}(\mathcal{H}_{A_O}) \to \mathcal{L}(\mathcal{H}_{B_I})$, connecting the output Hilbert space of Alice to Bob's input Hilbert space, can be described in process matrix language as $W = C^{A_O B_I}$, where C is the Choi matrix of the channel $\mathcal{C}$, as defined by equation (1). The process matrix describing N parallel uses of the channel $\mathcal{C}$ is simply $W = \prod_{i=1}^{N} C^{A_O^i B_I^i}$. Again, this process can be considered as a $2N$-partite process, or as a bipartite process with $A_O := \prod_i A_O^i$ and $B_I = \prod_i B_I^i$.

3. Composition rules

From the above considerations, it seems that one could simply take the tensor product as a composition rule to obtain multipartite processes representing multiple independent copies of a resource. In fact, equation (10)
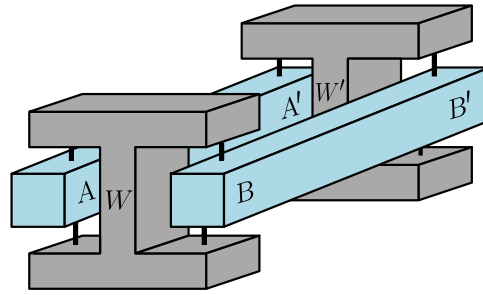
Figure 2. The tensor product composition rule $\mu(W, W') = W \otimes W'$. Here AA' is a composite party that can perform general quantum operations $\mathcal{L}(\mathcal{H}_{A_I A_I'}) \to \mathcal{L}(\mathcal{H}_{A_O A_O'})$, and similarly for BB'; the corresponding probabilities are given by equation (2). We shall show that this composition rule does not satisfy all requirements that we demand on such a rule.
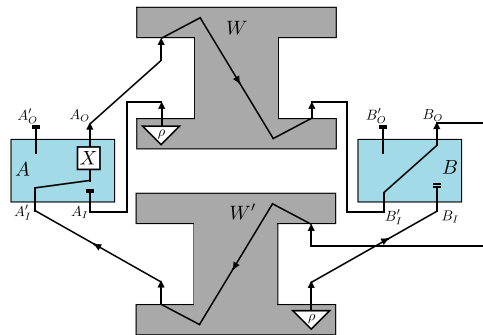
Figure 3. The tensor product composition rule $\mu(W, W') = W \otimes W'$ does not produce valid processes for all choices of W and W'. Here the process W corresponds to Alice receiving a state ρ, with an identity channel connecting her output system to Bob's input; W' is the same thing with the order of the parties reversed. The specific choice of local maps (X being the Pauli-X unitary gate) have zero probability under the 'generalized Born rule' equation (2), instead of one, as it should be for deterministic operations.

implies that whenever the linear constraints are satisfied for both W_1 and W_2, then the corresponding multipartite constraints will be satisfied for $W_1 \otimes W_2$.

The situation is different, however, if we require W_1 and W_2 to be shared by *the same* parties. To keep the discussion simple, consider only two parties, Alice and Bob, who share two possible processes, $W_1 \in \mathcal{W}$ and $W_2 \in \mathcal{W}'$. We want now to create the composite process $\mu(W^1, W^2)$ such that it is still a bipartite one, i.e. Alice can access both the systems $A_I A_O$ and $A_I' A_O'$, and Bob both $B_I B_O$ and $B_I' B_O'$. If both processes have the same definite order, i.e. $_{A_O} W_1 = W_1$ and $_{A_O'} W_2 = W_2$, or the analogous condition with B_O, B_O', then, we know from standard quantum theory that the right operation for composing such processes is $W_1 \otimes W_2$. This composition rule is represented in figure 2. One can easily prove that whenever the two processes do not have the same definite causal order, then $L_V(W^1 \otimes W^2) \neq W^1 \otimes W^2$, where L_V is taken with respect to the bipartition (AA', BB') [33]. For instance, consider the process

$$W \otimes W, \quad \text{with } W = \frac{1}{2}(W^{A \preceq B} + W^{B \preceq A}) \tag{11}$$

then, it is sufficient to check directly the violation of equation (7) with respect to the bipartition (AA', BB'), namely, $W \otimes W \neq {}_{B_O B_O'}(W \otimes W) + {}_{A_O A_O'}(W \otimes W) - {}_{A_O A_O' B_O B_O'}(W \otimes W)$. This problem is illustrated in figure 3, where two processes W, W' corresponding to channels in different directions can be seen to lead to a 'loop', and to unnormalised probabilities. It is then natural to ask whether the tensor product can be replaced with another composition rule.

One may, however, argue that it is in principle possible to define more general composition rules that take this problem into account. For instance, one could take the tensor product and then 'project' back the corresponding operator onto the space of valid processes, or one could first decompose the process into a linear combination of processes in a definite order, then take the tensor product of each term and then recombine them. There are infinitely many possible recipes to define a composition rule; an abstract prescription for general composition rules is provided in [36, 37]. In the following, we will ask three reasonable and physically motivated requirements and show that there is no way of satisfying all three.

To define our composition rule μ, we may ask the following minimal requirements:

R.1 $\mu(W_1, W_2)$ is a valid process w.r.t. the bipartition (AA', BB'), for $W_1 \in \mathcal{W}$, $W_2 \in \mathcal{W}'$ (validity).

R.2 $\mu(W_1, W_2) = W_1 \otimes W_2$ if $W_1 \in \mathcal{W}$, $W_2 \in \mathcal{W}'$, and W_1, W_2 are in the same order, i.e. ($_{A_\mathrm{O}}W_1 = W_1$ and $_{A_\mathrm{O}'}W_2 = W_2$), or ($_{B_\mathrm{O}}W_1 = W_1$ and $_{B_\mathrm{O}'}W_2 = W_2$) (consistency).

R.3 $\mu(W_1, W_2)$ is convex linear in both arguments (convex linearity);

Requirement R.1 is needed for the composition of two processes to still belong to a bipartite scenario, i.e. where Alice has access to both systems AA', and Bob to BB'. R.2 is a consistency condition, i.e. the case of definite order should coincide with standard quantum theory. R.3 can be derived by requiring that our composition is well-behaved with respect to statistical mixtures, i.e. classical randomness, as explained in appendix A.

It will be interesting to first consider a weaker assumption than R.1, because it will help us to single out the usual mathematical tensor product as a composition rule:

R'.1 $\mu(W_1, W_2) \geqslant 0$ for $W_1 \in \mathcal{W}$, $W_2 \in \mathcal{W}'$ (positivity);

Assume that μ is a composition rule satisfying R'.1 (or R.1), R.2, R.3. Then there is a *unique* real-linear extension μ^L that satisfies $\mu^L(W_1, W_2) = \mu(W_1, W_2)$, for all $W_1 \in \mathcal{W}$, $W_2 \in \mathcal{W}'$. By construction this extension satisfies:

R'.3 $\mu(W_1, W_2)$ is real linear in both arguments (linearity);

For the linear extension, we only demand R.1' (or R.1) for process matrices as inputs, so it will trivially continue to be satisfied. As R.2 itself is a (bi)linear condition, the linear extension will satisfy it even when it is extended to the linear span of process matrices:

R'.2 $\mu(W_1, W_2) = W_1 \otimes W_2$ if $W_1 \in L_V(\mathcal{L}(\mathcal{H}_{AB}))$, $W_2 \in L_V(\mathcal{L}(\mathcal{H}_{A'B'}))$, and W_1, W_2 satisfy ($_{A_\mathrm{O}}W_1 = W_1$ and $_{A_\mathrm{O}'}W_2 = W_2$), or ($_{B_\mathrm{O}}W_1 = W_1$ and $_{B_\mathrm{O}'}W_2 = W_2$) (consistency)

Details can be found in appendix A.

With our axioms, we will be able to prove

Theorem 1. *The only function satisfying R'.1, R'.2, R'.3 is* $\mu(W_1, W_2) := W_1 \otimes W_2$.

Theorem 1 can be applied to the linear extension μ^L, implying that $\mu(W_1, W_2) = W_1 \otimes W_2$, and from that it will follow

Theorem 2. *There exists no function satisfying R.1–R.3.*

In particular, theorem 1 will imply that for the multipartite case the choice of the composition rule is unique. We will prove theorem 1 for the simple case of local systems consisting of n-qubits, i.e. with local dimension 2^n for each one of A_I, A_I', …, B_O, B_O', the general proof can be found in appendix B. Given theorem 1, for the proof of theorem 2 it is sufficient to use the result of [33], or the example in equation (11).

First, we need the following

Lemma 1. *Given A_1, A_2 Hermitian operators such that $A_1 \in L_V(\mathcal{L}(\mathcal{H}_{AB}))$ and $A_2 \in L_V(\mathcal{L}(\mathcal{H}_{A'B'}))$, and let μ be a composition rule satisfying R'.1–3. Then $\mu(A_1, A_2) = \mu(A_1, A_2)^\dagger$ and $\|\mu(A_1, A_2)\| \leqslant \|A_1 \otimes A_2\|$.*

Proof. For A Hermitian, its norm can be written as: $\|A\| = \min\{\lambda \mid -\lambda\mathbb{1} \leqslant A \leqslant \lambda\mathbb{1}\}$. Consider $A_1 \in L_V(\mathcal{L}(\mathcal{H}_{AB}))$ and $A_2 \in L_V(\mathcal{L}(\mathcal{H}_{A'B'}))$ Hermitian and with $\lambda_i = \|A_i\|$ for $i = 1, 2$. We define

$$W_1^\pm = \lambda_1\mathbb{1} \pm A_1, \qquad W_2^\pm = \lambda_2\mathbb{1} \pm A_2, \tag{12}$$

which are valid processes, up to a normalization factor, on the spaces AB and $A'B'$. We then have

$$0 \leqslant \frac{\mu(W_1^+, W_2^+) + \mu(W_1^-, W_2^-)}{2} = \lambda_1\lambda_2\mathbb{1} + \mu(A_1, A_2),$$

$$0 \leqslant \frac{\mu(W_1^+, W_2^-) + \mu(W_1^-, W_2^+)}{2} = \lambda_1\lambda_2\mathbb{1} - \mu(A_1, A_2), \tag{13}$$

which implies $\mu(A_1, A_2) = \mu(A_1, A_2)^\dagger$ and $\|\mu(A_1, A_2)\| \leqslant \lambda_1 \lambda_2$. In the above, we used R′.1 for positivity, then R′.3 to split the different terms, and finally, R.2′ to take the identity out of μ. □

For the following, we need to specialize the form of the operator A_1 and A_2. We define the set of tensor products of either traceless operators or the identity as

$$\mathrm{PTI}_{AB} := \{M = X_{A_I}^1 \otimes X_{A_O}^2 \otimes X_{B_I}^3 \otimes X_{B_O}^4 \,|\, M \in L_V(\mathcal{L}(\mathcal{H}_{AB})), \ X^i \ \text{identity or traceless}\}, \tag{14}$$

and analogously for A′B′. For $M \in \mathrm{PTI}_{AB}$, an operator of the form $\mathbb{1} + M$ is, up to normalization, a causally ordered process. With the above definition, we prove the following

Lemma 2. *Let μ be a composition rule satisfying R′.1–3, and let $M \in \mathrm{PTI}_{AB}$ and $N \in \mathrm{PTI}_{A'B'}$ be Hermitian operators with eigenvalues in the interval $[-1, 1]$. Given an eigenvector $|k\rangle$ of M with eigenvalue $(-1)^k$ and an eigenvector $|j\rangle$ of N with eigenvalue $(-1)^j$, we have*

$$\mu(M, N)|k, j\rangle = (-1)^{k+j}|k, j\rangle \tag{15}$$

Proof. To prove the lemma, it is sufficient to consider the (unnormalized) processes $W_1^k := \mathbb{1} + (-1)^{k+1}M$ and $W_2^j := \mathbb{1} + (-1)^{j+1}N$. By R′.2, $\mu(\mathbb{1}, \mathbb{1}) = \mathbb{1} \otimes \mathbb{1}$ and $\mu(M, \mathbb{1}) = M \otimes \mathbb{1}$, since for $M \in \mathrm{PTI}_{AB}$, either $_{A_O}M = M$ or $_{B_O}M = M$. Then

$$\mu(W_1^k, W_2^j) = \mathbb{1} + (-1)^{k+1}M \otimes \mathbb{1} + (-1)^{j+1}\mathbb{1} \otimes N + (-1)^{k+j}\mu(M, N). \tag{16}$$

by R′.2 and R.3′, and finally, by R′.1

$$0 \leqslant \langle k, j| \ \mu(W_1^k, W_2^j)|k, j\rangle = 1 - 1 - 1 + (-1)^{j+k}\langle k, j| \ \mu(M, N)|k, j\rangle, \tag{17}$$

which implies $\mu(M, N)|k, j\rangle = (-1)^{j+k}|k, j\rangle$, since $\|\mu(M, N)\| \leqslant 1$, by lemma 1. □

A straightforward corollary of lemma 2 is that $\mu(M, N) = M \otimes N$ whenever M, N have eigenvalues only in $\{-1, 1\}$. By linearity, this is enough to prove theorem 1 for all processes defined on n-qubit systems (i.e. local dimension 2^n) since we have a basis of operators, given by tensor products of Pauli matrices and the identity, that satisfy the assumptions. The same reasoning can be extended to arbitrary dimensions, see the details in appendix B.

4. Discussion and conclusions

In this letter, we considered the parallel composition of process matrices. As the tensor product is known to lead to invalid process matrices, we investigated whether there is another map that can describe this parallel composition. We only asked for three weak desiderata: First of all, in contrast to the usual tensor product, it should always result in a valid process matrix. Furthermore, it should reduce to the familiar tensor product in the case of definite causal order. At last, we demanded bilinearity for compatibility with the interpretation of convex mixtures as statistical mixtures. However we have seen that even those reasonable desiderata are incompatible with each other.

Our results imply that an information theory of general quantum processes cannot rely on the assumption that multiple independent processes can be shared between two (or more) parties. In information theory, it is typical to assume that many independent samples of a random source, many independent uses of a channel, etc are available, and that agents can perform global operations on many independent copies of the resource; this will not be possible in an information theory of general quantum processes. Rather, these results suggest that the proper setting for defining information-theoretic quantities such as entropies, capacities, etc, for process matrices is *single-shot* information theory [38–40].

One can infer from the main proof that even the case of two channels with opposing signaling direction will lead to a contradiction, which is perhaps unsurprising in the usual case of quantum mechanics on a fixed background spacetime. Indeed, suppose that an event A is in the causal past of an event B, and that A' is in the causal future of B'. Our desiderata that A and A' correspond to the same party can be interpreted as requiring that the events A, A' occur at the same space-time point p. This could be the case, but then B must be in the future light-cone of p, while B' must be in it's past light-cone. It is thus impossible to satisfy the requirement that B and B' also occur at the same spacetime point.

Therefore any composition rule for process matrices must take care of removing the two-way signaling terms, whose impossibility has a clear interpretation as discussed above. We have shown that there is no linear way of doing so, if we ask for that our composition rule reduces to the usual tensor product in the case of two processes with the same definite causal order.

However, there might exist reasonable nonlinear composition rules, in the cases where processes have a concrete physical interpretation. A meaningful way to define an event for the composite party AA' is by the 'simultaneous' entering of both systems $\mathcal{H}_A$ and $\mathcal{H}_{A'}$ in a localized laboratory, and similarly for BB'. There can be a probability that the systems do not enter the laboratories simultaneously, in which case it is necessary to post-select on the runs of the experiment where this was indeed the case. Since the post-selection probability depends on the two processes that we wish to compose, the map will be nonlinear. An important issue with such a post-selected composition map for information-theoretic applications is that the parallel composition of resources is usually a 'free operation', while in the post-selected case it would have a probability of failure.

Acknowledgments

Appendix A. Linearity and convex linearity

In this appendix, we discuss convex-linearity and the linear extension of convex maps. First, let us argue why convex-linearity is a reasonable physical assumption. In operational approaches to physical theories [41, 42], one studies the probabilities that can be obtained from an abstract set of preparations and measurements. Given two preparations α, and β, there exists another preparation γ that consists of preparing α with classical probability p, and preparing β with probability $(1 - p)$. The probability for any measurement on γ is the weighted sum of the probabilities associated with preparations α and β. If we associate 'states' with preparations, this means that the state space is convex linear. For example, the density matrix formalism can be seen to arise by adding classical uncertainty to the pure state formalism (i.e. kets in a Hilbert space). If one knows that with probability p_j, one prepares $|j\rangle$, then the density matrix is given by $\rho = \sum_j p_j |j\rangle \langle j|$. Another motivation for allowing arbitrary probabilistic mixtures appears in [43, 44], where it is shown that it implies that optimal compression is equivalent to linear compression.

The same interpretation can be used for process matrices: if the process matrices W_j are prepared with probabilities p_j, then all expectations values (and by that all statistics) can be calculated with the effective process matrix $W = \sum_j p_j W_j$. This can be seen by noting that $p(a, b) = \text{Tr}[W M_a^{(A)} \otimes M_b^{(B)}]$ is a linear function in W and applying the law of total probability.

Consistency demands that the composition rule μ remains compatible with this interpretation of convex mixtures: if the first process is W_j with probability p_j and the second process is W_k' with probability p_k', then the effective process matrices determining the statistics are $W = \sum_j p_j W_j$ and $W' = \sum_k p_k' W_k'$. The resulting combined process would be $\mu(\sum_j p_j W_j, \sum_k p_k' W_k')$. However, a different point of view would be to say: with probabilities p_j and p_k' we combined the processes W_j and W_k' to $\mu(W_j, W_k')$. So we prepared $\mu(W_j, W_k')$ with probability $p_j p_k'$. Now, the effective process matrix is described by $\sum_{jk} p_j p_k' \mu(W_j, W_k')$. As both points of view describe the same operational scenario, they have to be consistent:

$$\mu\left(\sum_j p_j W_j, \sum_k p_k' W_k'\right) = \sum_{jk} p_j p_k' \mu(W_j, W_k'). \tag{A1}$$

Next, we explain in further detail how to extend a function satisfying R.1 (or R′.1), R.2 and R.3 to a function satisfying R.1 (or R′.1), R′.2 and R′.3 on the linear span of all the process matrices.

Constructing the (bi)linear extension itself is a standard procedure in quantum information theory and is explained e.g. in [41, 42] for general abstract state spaces. Let S_1, S_2 be two convex sets, and let $f: S_1 \to S_2$ be a convex linear map. Let V_1, V_2 be the real vector spaces obtained respectively from S_1, S_2 by taking their linear span. Then f can be extended in the obvious way to a linear function $f^L: V_1 \to V_2$, defined by $f^L(\lambda a + b) = \lambda f(a) + f(b)$, for all $a, b \in S_1, \lambda \in \mathbb{R}$.

However, we still need to check that the bilinear extension still satisfies our postulates: we do not change R.1 (or R′.1), i.e. we only demand the output to be a process matrix (or positive) if the inputs are process matrices.

Therefore R.1 (or R′.1) trivially continues to hold as the extension does not change the function on inputs that are process matrices.

Less trivial is how to generalize R.2. We will explicitly show that it still holds for the cases we need. Let us assume we have operators $M_1 \in L_V(\mathcal{L}(\mathcal{H}_{AB}))$ and $M_2 \in L_V(\mathcal{L}(\mathcal{H}_{A'B'}))$ with $_{A_O}M_1 = M_1$ and $_{A_O'}M_2 = M_2$ (or alternatively $_{B_O}M_1 = M_1$ and $_{B_O'}M_2 = M_2$). We now show that

$$\mu^L(M_1, M_2) = M_1 \otimes M_2. \tag{A2}$$

By definition, M_1 and M_2 are allowed terms satisfying the projective condition (8). Therefore there exist λ_1, λ_2 such that $\frac{1}{d_I} + \lambda_1 M_1$ and $\frac{1}{d_{I'}} + \lambda_2 M_2$ are valid process matrices. Similarly $\frac{1}{d_I d_{I'}}$ itself is a valid process matrix, with no signaling at all. Using R.2 for the original μ on valid process matrices, we find for the linear extension:

$$\mu^L\left(\frac{1}{d_I} + \lambda_1 M_1, \frac{1}{d_{I'}}\right) = \mu\left(\frac{1}{d_I} + \lambda_1 M_1, \frac{1}{d_{I'}}\right) = \left(\frac{1}{d_I} + \lambda_1 M_1\right) \otimes \frac{1}{d_{I'}} = \frac{1}{d_I d_{I'}}\mathbb{1} \otimes \mathbb{1} + \frac{\lambda_1}{d_{I'}}M_1 \otimes \mathbb{1}$$

$$= \mu\left(\frac{1}{d_I}, \frac{1}{d_{I'}}\right) + \frac{\lambda_1}{d_{I'}}M_1 \otimes \mathbb{1} = \mu^L\left(\frac{1}{d_I}, \frac{1}{d_{I'}}\right) + \frac{\lambda_1}{d_{I'}}M_1 \otimes \mathbb{1}.$$

Therefore by bilinearity we find $\mu^L(M_1, \mathbb{1}) = M_1 \otimes \mathbb{1}$ and similarly $\mu^L(\mathbb{1}, M_2) = \mathbb{1} \otimes M_2$. Similarly, applying R.2 to the process matrices $\frac{1}{d_I} + \lambda_1 M_1$ and $\frac{1}{d_{I'}} + \lambda_2 M_2$, which have the same signaling direction, we obtain

$$\mu^L\left(\frac{1}{d_I} + \lambda_1 M_1, \frac{1}{d_{I'}} + \lambda_2 M_2\right) = \mu\left(\frac{1}{d_I} + \lambda_1 M_1, \frac{1}{d_{I'}} + \lambda_2 M_2\right) = \left(\frac{1}{d_I} + \lambda_1 M_1\right) \otimes \left(\frac{1}{d_{I'}} + \lambda_2 M_2\right). \tag{A3}$$

Collecting our results and using bilinearity on the left hand side of equation (A3) above, we finally see that R′.2 is satisfied:

$$\mu^L(M_1, M_2) = M_1 \otimes M_2. \tag{A4}$$

Appendix B. Proof of theorem 1 in arbitrary dimension

In this appendix, we will extend the proof of theorem 1 to the case of arbitrary dimension. We start with the following

Lemma 3. *Let $M \in \mathrm{PTI}_{AB}$ and $N \in \mathrm{PTI}_{A'B'}$ be Hermitian operators such that $|k\rangle = |k_1\rangle_{A_I} \otimes |k_2\rangle_{A_O} \otimes |k_3\rangle_{B_I} \otimes |k_4\rangle_{B_O}$ is an eigenvector for M, with eigenvalues given, according to the above factorization, by the products $\lambda_k = \lambda_k^{(1)}\lambda_k^{(2)}\lambda_k^{(3)}\lambda_k^{(4)}$, with $\lambda_k^{(i)} \in \{-1, 0, 1\}$ $i = 1, 2, 3, 4$, and, similarly, $|j\rangle = |j_1\rangle_{A_{I'}} \otimes |j_2\rangle_{A_{O'}} \otimes |j_3\rangle_{B_{I'}} \otimes |j_4\rangle_{B_{O'}}$ is an eigenvector of N, with eigenvalue $\eta_j = \eta_j^{(1)}\eta_j^{(2)}\eta_j^{(3)}\eta_j^{(4)}$, with $\eta_j^{(i)} \in \{-1, 0, 1\}$ $i = 1, 2, 3, 4$. We then have*

$$\mu(M, N)|k, j\rangle = \lambda_k \eta_j |k, j\rangle. \tag{B1}$$

Proof. The cases λ_k, $\eta_j = \pm 1$ are included in lemma 2. Let us consider the case $M|k\rangle = 0$ and $N|j\rangle \neq 0$, the case $M|k\rangle = N|j\rangle = 0$ can be obtained in a similar way, by applying the same argument first to M, then to N. Since M is in PTI_{AB}, we can write it as $M = X_{A_I}^1 \otimes X_{A_O}^2 \otimes X_{B_I}^3 \otimes X_{B_O}^4$. Let us now further assume $X^1|k_1\rangle_{A_I} = 0$, and $X^i|k_i\rangle_Y \neq 0$ for $i = 2, 3, 4, Y = A_O, B_I, B_O$, in particular, this implies that $|k_i\rangle$ are eigenvectors for eigenvalues ± 1 for $i = 2, 3, 4$. We can then write:

$$X^1 = (X^1 + |k_1\rangle\langle k_1| - |(k+1)_1\rangle\langle (k+1)_1|) + (|(k+1)_1\rangle\langle (k+1)_1| - |k_1\rangle\langle k_1|) =: X'^1 + X''^1, \tag{B2}$$

where $|(k+1)_1\rangle$ is a vector orthogonal to $|k_1\rangle$. Then X'^1, X''^1 are both traceless and $X'^1|k_1\rangle = |k_1\rangle$, $X''^1|k_1\rangle = -|k_1\rangle$. We then have that $M' := X_{A_I}'^1 \otimes X_{A_O}^2 \otimes X_{B_I}^3 \otimes X_{B_O}^4$ and $M'' := X_{A_I}''^1 \otimes X_{A_O}^2 \otimes X_{B_I}^3 \otimes X_{B_O}^4$ are again in PTI_{AB}. Thus, by lemma 2

$$\mu(M, N)|k, j\rangle = \mu(M' + M'', N)|k, j\rangle = \mu(M', N)|k, j\rangle + \mu(M'', N)|k, j\rangle$$
$$= M' \otimes N|k, j\rangle + M'' \otimes N|k, j\rangle = 0. \tag{B3}$$

If another operator, say X^2, is zero on the corresponding eigenvector, say $|k_2\rangle_{A_O}$, we can again repeat the construction in equation (B2) to construct X'^2, X''^2 with $+1$, -1 eigenvalues and use again linearity and lemma 2. Similarly, the same argument can be extended to all X^i and to N. $\square$

To conclude the proof of theorem 1, it is sufficient to construct a basis of operators containing the identity and where each element, except the identity, is traceless and with eigenvalues in $\{-1, 0, 1\}$. Let $\mathcal{H}$ be a Hilbert space with dimension d, and let $\{|k\rangle\}_{k=1}^d$ be a basis for $\mathcal{H}$. The space of Hermitian operators on $\mathcal{H}$ is a real vector space of dimension d^2. We define the following operators

$$Z_i = |i\rangle\langle i| - |i+1\rangle\langle i+1|, \quad 1 \leqslant i \leqslant d-1, \tag{B4}$$

$$X_{jk} = |j\rangle\langle k| + |k\rangle\langle j|, \quad 1 \leqslant j < k \leqslant d, \tag{B5}$$

$$Y_{jk} = i(|j\rangle\langle k| - |k\rangle\langle j|), \quad 1 \leqslant j < k \leqslant d, \tag{B6}$$

which are traceless, hermitian and with eigenvalues in $\{-1, 0, 1\}$. The X_{jk} and Y_{jk} are also known as part of an operator basis called *Generalized Gell-Mann matrices* [45]. For completeness we now show that, together with $\mathbb{1}$, the above set of matrices form a basis for the space of Hermitian operators on $\mathcal{H}$. It is clear that the $\{X_{jk}\}$ and $\{Y_{jk}\}$ span the space of Hermitian operators whose diagonal is zero in the $|k\rangle$ basis. All that remains to be shown is that $\{\mathbb{1}, Z_i\}$ forms a basis for the space of diagonal real matrices, which we prove by expressing the basis $\{|k\rangle\langle k|\}$ in terms of the new basis $\{\mathbb{1}, Z_i\}$.

Notice that for $1 \leqslant i \leqslant d-1$

$$\sum_{j=i}^{j=d-1} Z_j = |i\rangle\langle i| - |d\rangle\langle d|, \tag{B7}$$

and also that

$$\sum_{j=1}^{d-1} jZ_j = \sum_{j=1}^{d-1} |j\rangle\langle j| - (d-1)|d\rangle\langle d| = \mathbb{1} - d|d\rangle\langle d|. \tag{B8}$$

Combining the above two expressions gives

$$|d\rangle\langle d| = \frac{1}{d}\mathbb{1} - \frac{1}{d}\sum_{j=1}^{d-1} jZ_j, \tag{B9}$$

$$|i\rangle\langle i| = \frac{1}{d}\mathbb{1} + \sum_{j=i}^{d-1} Z_j - \frac{1}{d}\sum_{j=1}^{d-1} jZ_j, \quad 1 \leqslant i < d, \tag{B10}$$

which concludes that $\{\mathbb{1}, Z_i, X_{jk}, Y_{jk}\}$ is a basis for the space of Hermitian operators of $\mathcal{H}$.

We can use the above construction to build a basis for $\mathcal{L}(\mathcal{H}_{A_I} \otimes \mathcal{H}_{A_O} \otimes \mathcal{H}_{B_I} \otimes \mathcal{H}_{B_O})$ consisting of tensor products of local Hermitian operators whose eigenvalues are in $\{-1, 0, 1\}$. We then remove from this basis all the terms that do not satisfy the linear constraints L_V. This gives us a basis for the linear space of valid Ws, which is contained in PTI_{AB}. We will call this basis simply $\{M_i\}_{i \in I}$, and by lemma 3, we have

$$\mu(M_i, M_j) = M_i \otimes M_j. \tag{B11}$$

We can then decompose any pair W, W' as

$$W = \mathbb{1} + \sum_i c_i M_i, \quad W' = \mathbb{1} + \sum_i d_i M_i, \tag{B12}$$

and apply μ, namely

$$\mu(W, W') = \mathbb{1} + \sum_i c_i M_i \otimes \mathbb{1} + \mathbb{1} \otimes \sum_i d_i M_i + \sum_{ij} c_i d_j \mu(M_i, M_j)$$

$$= \mathbb{1} + \sum_i c_i M_i \otimes \mathbb{1} + \mathbb{1} \otimes \sum_i d_i M_i + \sum_{ij} c_i d_j M_i \otimes M_j = W \otimes W', \tag{B13}$$

which concludes the proof of theorem 1.

ORCID iDs

Philippe Allard Guérin ⓘ https://orcid.org/0000-0002-5044-4799
Costantino Budroni ⓘ https://orcid.org/0000-0002-6562-7862

References

[1] Bell J S 1964 On the Einstein–Podolsky–Rosen paradox *Physics* **1** 195
[2] Wood C J and Spekkens R W 2015 The lesson of causal discovery algorithms for quantum correlations: causal explanations of Bell-inequality violations require fine-tuning *New J. Phys.* **17** 033002
[3] Hardy L Probability theories with dynamic causal structure: a new framework for quantum gravity arXiv:gr-qc/0509120
[4] Hardy L 2007 Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure *J. Phys. A: Math. Theor.* **40** 3081
[5] Hardy L 2009 Quantum gravity computers: on the theory of computation with indefinite causal structure *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony* (Dordrecht: Springer) pp 379–401
[6] Oreshkov O, Costa F and Brukner Č 2012 Quantum correlations with no causal order *Nat. Commun.* **3** 1092

[7] Chiribella G, D'Ariano G M, Perinotti P and Valiron B 2013 Quantum computations without definite causal structure *Phys. Rev.* A **88** 022318

[8] Oeckl R A local and operational framework for the foundations of physics arXiv:1610.09052

[9] Araújo M, Branciard C, Costa F, Feix A, Giarmatzi C and Brukner Č 2015 Witnessing causal nonseparability *New J. Phys.* **17** 102001

[10] 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 Experimental superposition of orders of quantum gates *Nat. Commun.* **6** 7913

[11] Rubino G, Rozema L A, Feix A, Araújo M, Zeuner J M, Procopio L M, Brukner Č and Walther P 2017 Experimental verification of an indefinite causal order *Sci. Adv.* **3** e1602589

[12] Goswami K, Giarmatzi C, Kewming M, Costa F, Branciard C, Romero J and White A G 2018 Indefinite causal order in a quantum switch *Phys. Rev. Lett.* **121** 090503

[13] Oreshkov O On the whereabouts of the local operations in physical realizations of quantum processes with indefinite causal order arXiv:1801.07594

[14] Oreshkov O and Giarmatzi C 2016 Causal and causally separable processes *New J. Phys.* **18** 093020

[15] Branciard C, Araújo M, Feix A, Costa F and Brukner Č 2016 The simplest causal inequalities and their violation *New J. Phys.* **18** 013008

[16] Abbott A A, Giarmatzi C, Costa F and Branciard C 2016 Multipartite causal correlations: polytopes and inequalities *Phys. Rev.* A **94** 032131

[17] Miklin N, Abbott A A, Branciard C, Chaves R and Budroni C 2017 The entropic approach to causal correlations *New J. Phys.* **19** 113041

[18] Araújo M, Costa F and Brukner Č 2014 Computational advantage from quantum-controlled ordering of gates *Phys. Rev. Lett.* **113** 250402

[19] Feix A, Araújo M and Brukner Č 2015 Quantum superposition of the order of parties as a communication resource *Phys. Rev.* A **92** 052326

[20] Guérin P A, Feix A, Araújo M and Brukner Č 2016 Exponential communication complexity advantage from quantum superposition of the direction of communication *Phys. Rev. Lett.* **117** 100502

[21] Baumeler Ä and Wolf S 2017 Non-causal computation *Entropy* **19** 326

[22] Araújo M, Guérin P A and Baumeler Ä 2017 Quantum computation with indefinite causal structures *Phys. Rev.* A **96** 052315

[23] Baumeler Ä and Wolf S 2018 Computational tameness of classical non-causal models *Proc. R. Soc.* A **474** 20170698

[24] Ebler D, Salek S and Chiribella G 2018 Enhanced communication with the assistance of indefinite causal order *Phys. Rev. Lett.* **120** 120502

[25] Nielsen M A and Chuang I L 2011 *Quantum Computation and Quantum Information: X Anniversary Edition* 10th edn (New York : Cambridge University Press)

[26] Wilde M M 2017 *Quantum Information Theory* 2nd edn (Cambridge: Cambridge University Press)

[27] Schumacher B 1995 Quantum coding *Phys. Rev.* A **51** 2738–47

[28] Zych M 2017 Quantum systems under gravitational time dilation *Springer Theses* Springer International Publishing

[29] Feix A and Brukner Č 2017 Quantum superpositions of common-cause and direct-cause causal structures *New J. Phys.* **19** 123028

[30] Zych M, Costa F, Pikovski I and Brukner C Bell's theorem for temporal order arXiv:1708.00248

[31] Chiribella G, D'Ariano G M and Perinotti P 2009 Theoretical framework for quantum networks *Phys. Rev.* A **80** 022339

[32] Costa F and Shrapnel S 2016 Quantum causal modelling *New J. Phys.* **18** 063032

[33] Jia D and Sakharwade N 2018 Tensor products of process matrices with indefinite causal structure *Phys. Rev.* A **97** 032110

[34] Jamiołkowski A 1972 Linear transformations which preserve trace and positive semidefiniteness of operators *Rep. Math. Phys.* **3** 275–8

[35] Choi M-D 1975 Completely positive linear maps on complex matrices *Linear Algebra Appl.* **10** 285–90

[36] Perinotti P 2017 Causal structures and the classification of higher order quantum computations *Time in Physics (Tutorials, Schools, and Workshops in the Mathematical Sciences)* ed R Renner and S Stupar 1st edn (Basel: Birkhäuser) (https://doi.org/10.1007/978-3-319-68655-4_7)

[37] Bisio A and Perinotti P Axiomatic theory of higher-order quantum computation arXiv:1806.09554

[38] Renner R, Wolf S and Wullschleger J 2006 The single-serving channel capacity *2006 IEEE Int. Symp. on Information Theory.* IEEE arXiv:cs/0608018 [cs.IT]

[39] Konig R, Renner R and Schaffner C 2009 The operational meaning of min- and max-entropy *IEEE Trans. Inf. Theory* **55** 4337–47

[40] Chiribella G and Ebler D 2016 Optimal quantum networks and one-shot entropies *New J. Phys.* **18** 093053

[41] Hardy L Quantum theory from five reasonable axioms arxiv:quant-ph/0101012

[42] Barrett J 2007 Information processing in generalized probabilistic theories *Phys. Rev.* A **75** 032304

[43] Hardy L Foliable operational structures for general probabilistic theories arXiv:0912.4740

[44] Hardy L Reformulating and reconstructing quantum theory arXiv:1104.2066

[45] Bertlmann R A and Krammer P 2008 Bloch vectors for qudits *J. Phys. A: Math. Theor.* **41** 235303