Quantum Theory is a Quasi-stochastic Process Theory

John van de Wetering

Radboud University
Nijmegen, Netherlands
`wetering@cs.ru.nl`

There is a long history of representing a quantum state using a quasi-probability distribution: a distribution allowing negative values. In this paper we extend such representations to deal with quantum channels. The result is a convex, strongly monoidal, functorial embedding of the category of trace preserving completely positive maps into the category of quasi-stochastic matrices. This establishes quantum theory as a subcategory of quasi-stochastic processes. Such an embedding is induced by a choice of minimal informationally complete POVM's. We show that any two such embeddings are naturally isomorphic. The embedding preserves the dagger structure of the categories if and only if the POVM's are symmetric, giving a new use of SIC-POVM's, objects that are of foundational interest in the QBism community. We also study general convex embeddings of quantum theory and prove a dichotomy that such an embedding is either trivial or faithful.

Introduction

In Feynman's famous 1981 paper on quantum computation [9] he writes "The only difference between a probabilistic classical world and the equations of the quantum world is that somehow or other it appears as if the probabilities would have to go negative". In this paper we wish to make this statement exact.

Of course, much work has already been done in this regard going all the way back to the Wigner quasi-probability distribution in 1932 [24]. The Wigner function allows you to associate a probability distribution over a phase space for a quantum particle, with the only caveat that the probability sometimes has to be negative. The negativity appearing in probabilistic representations of quantum systems is something that lies at the heart of quantum theory: Spekkens has shown [20] that the necessity of negativity in probabilistic representations is equivalent to the contextuality of quantum theory. It is also a necessity for a quantum speedup as states represented positively by the Wigner function can be efficiently simulated [16, 17, 22, 23]. An operational interpretation of negative probabilities is given by Abramsky and Brandenburger [2].

The main contribution of this paper will be to represent *all* of (finite-dimensional) quantum theory as a set of quasi-stochastic processes, not just the states. In particular we will use the language of category theory to establish that the category of quantum processes is a subcategory of the category of quasi-stochastic processes. A central object for studying these representations is the *informationally complete POVM* (IC-POVM), this is a measurement that completely characterizes a quantum state. We are particularly interested in IC-POVMs that are minimal as these form a basis of the statespace. A particular type of such a measurement is a symmetric IC-POVM: this is a very special type of POVM that is of particular interest to the QBism community [10] as it allows the updating of states to be written in a particularly elegant manner. It also allows quantum states to be written down with a minimal amount of negativity [25]. Minimal and symmetric POVMs turn out to be essential to preserving respectively the tensor product and adjoint in the quasi-stochastic representations considered in this paper.

Representing quantum processes by quasi-stochastic matrices is not a new idea. In particular it is used in [4] to argue the similarity of a unitary process and the Born rule, although they stop short of extending

the rule to all quantum channels and of composing them. In [17] they also don't go into detail about the compositional nature of these representations either. As far as the author is aware, this paper is the first to consider the compositional structure of quasi-stochastic representations of quantum theory. An approach that comes close is that of the *duotensor* framework of Hardy [14], in particular his *hopping metric* is similar to the transition matrices T in this paper, but because Hardy's *fiducial effects* don't have to form a POVM the values in the hopping metric don't form a quasi-probability distribution. In [4, 8] it was shown that the negativity in the representations can be overcome by modifying the way probabilities are calculated. In this paper this modification takes the form of a category that has a modified composition rule.

We will represent quantum theory using a category **CPTP** close to that of Selinger's [19] consisting of trace preserving completely positive maps. This category models the dynamics of finite dimensional quantum systems where throwing away systems is allowed. It excludes classical systems which are, for instance, present in the category of finite dimensional C*-algebras. Most of the results carry over to this setting, a small example of which we will give at the end of section 2. The category consisting of quasi-stochastic matrices that compose via matrix multiplication will be denoted **QStoch**. The primary contribution of the paper is the following theorem.

Theorem. Any family of minimal informationally complete (IC) POVM's gives rise to a functor F : **CPTP** $\rightarrow$ **QStoch**. This functor is faithful (injective) and strong monoidal (preserving tensor product), it furthermore preserves the convex structure of **CPTP**. The functors arising from two different families of minimal IC-POVM's are naturally isomorphic. The functor preserves the adjoint of unital channels if and only if the POVM's are generalised symmetric informationally complete (SIC).

Note that the fact that any two representations by minimal POVM's are naturally isomorphic in a way seems to answer the question why there doesn't seem to be a preferred way to represent quantum theory by probabilities: the category doesn't 'see' this difference. This theorem establishes that **CPTP** can be considered a subcategory of **QStoch**. **CPTP** is not a *full* subcategory: for a given system not all quasi-probability distributions correspond to valid quantum states. In particular, since no orthogonal IC-POVM exists, the probability distribution that is associated to a quantum state will always be somewhat mixed. This is qualitatively similar to the concept of Spekkens' *epistricted theories* [21]. As shown in that paper, many of the characteristic features of quantum theory (no-cloning, teleportation, dense coding, entanglement) arise in classical epistemically restricted theories: theories where states of maximal knowledge are not available. Some features of quantum theory however do *not* occur in classical epistemically restricted theories, primarily Bell / noncontextuality inequality violations and a computational speed-up. As shown by Abramsky and Brandenburger [1], any non-signalling correlations (including those that maximally violate the Bell inequality) can be represented by a hidden variable model if one allows negative probabilities. Furthermore it was shown that any tomographically local theory has a complexity bound of **AWPP** [15] which was later shown by the same authors to be achieved by a computational model based on a quasi-probabilistic Turing machine [5]. These results together suggest that the features of quantum theory that don't occur in classical epistricted theories can be explained by the presence of negative probabilities in quantum theory and that in fact the necessity of this negativity is the 'cause' of violating Bell inequalities and achieving a computational speed-up.

We also prove a converse statement to the above theorem. Defining a quasi-stochastic representation of quantum theory to be a convex functor F : **CPTP** $\rightarrow$ **QStoch** we show the following.

Theorem. To each quasi-stochastic representation we can associate a family of quasi-POVM's (POVM's that don't have to consist of positive components) that determine the representation on the states. Exactly one of the following holds for all quasi-stochastic representations.

1. All the components of the associated quasi-POVM's are multiples of the identity, in which case the representation is trivial.

2. All the associated quasi-POVM's are informationally complete, in which case the representation is faithful.

Furthermore, for nontrivial representations it holds that

1. the representation is strong monoidal if and only if the associated quasi-POVM's are *minimal* informationally complete.

2. the representation preserves the dagger if and only if the associated quasi-POVM's are *symmetric* informationally complete.

This gives an interesting new way to look at minimal generalised SIC-POVM's: they are the only POVM's that lead to a quasi-stochastic representation of quantum theory that is both strong monoidal (preserving the tensor product) and that preserves the dagger (the adjoint). The fact that each representation comes from a family of quasi-POVM's mirrors a very similar statements about frames in [8].

Any representation is either trivial or faithful and if it is strong monoidal it also has to be minimal. It might even be the case that along the lines of the proof of Theorem 19 that any nontrivial representation has to be minimal regardless, although we do not show this. If this is in fact the case then it is a strong argument that minimal (quasi-)IC-POVM's are an essential object in quantum theory: they would be the only objects inducing faithful quasi-stochastic representations.

Considering the above results it is a natural question to ask whether any non-signalling theory allows an embedding into the category of quasi-stochastic processes. This turns out to be true for any causal operational-probabilistic theory [6] which allows coarse-graining and is nondeterministic. This follows easily from the existence of minimal IC measurements in those theories and an adaptation of the proofs in this paper. Such a theory allows a strong monoidal embedding if and only if it allows local tomography.

We can also try to start with **QStoch** and identify suitable properties of subcategories that make them isomorphic to **CPTP**. Work has already been done in this direction by Appleby, Fuchs, Schack and others [4, 3, 11]. They have restricted themselves to looking at state spaces. By considering the entire compositional structure of quantum theory, finding suitable axioms might be easier. In fact, using some simple axioms and the structure of **QStoch** it is possible to derive the modified composition rule $q(j) = \sum_i (\alpha p(i) - \beta) r(j|i)$ used in those papers. This will appear in later work.

We assume some basic familiarity with some concepts of category theory, namely that of a category, functor and natural transformation. All other categorical concepts will be explained when necessary. In section 1 we will establish the standard way of associating stochastic matrices and probability distributions to quantum channels and states as done in for instance [8, 7]. We will extend this to a functor in section 2. The preservation of the tensor product and the adjoint will be studied respectively in sections 3 and 4. Finally, general representations of quantum theory are considered in section 5. The appendix contains some of the longer proofs.

1   Preliminaries

Definition 1. Let $A \in M^{n \times m}(\mathbb{R})$ be a $n \times m$ matrix with real entries. It is called *quasi-stochastic* when the values in each column sum up to 1. It is called *stochastic* when it is quasi-stochastic and all the entries

are positive. A matrix is called *doubly (quasi-)stochastic* when it is (quasi-)stochastic and its transpose is also (quasi-)stochastic. A doubly quasi-stochastic matrix is necessarily square.

Note that if a (doubly) quasi-stochastic matrix has an inverse, this inverse will also be (doubly) quasi-stochastic. The inverse of a stochastic matrix is not necessarily stochastic (that is, some negative components might occur).

Definition 2. Let $M_n = M^{n \times n}(\mathbb{C})$. We call $\rho \in M_n$ a *state* when $\rho \geq 0$ and $\text{tr}(\rho) = 1$. We call $E \in M_n$ an *effect* when $0 \leq E \leq 1$. A set of effects $\{E_i\} \subseteq M_n$ is called a *POVM* when $\sum_i E_i = I_n$ where I_n denotes the identity. If the set of E_i are Hermitian and satisfy $\sum_i E_i = I_n$ while no longer being necessarily positive, then we call this set a *quasi-POVM*[1]. If a (quasi-)POVM spans M_n we call it *Informationally Complete* (IC) and if the elements of a IC-POVM are also linearly independent then it forms a basis for M_n and we call it *minimal* informationally complete. Such a set always has n^2 elements.

Measuring a state ρ using a POVM $\{E_i\}$ leads to probabilities $p(i) = \text{tr}(\rho E_i)$ by the Born rule. Since the E_i are positive and they sum up to I_n these probabilities indeed form a proper probability distribution: $\sum_i p(i) = 1$.

If the POVM is minimal IC then we can write each ρ uniquely in terms of the E_i:

$$\rho = \sum_i \alpha_i \frac{E_i}{\text{tr}(E_i)}.$$

We have chosen to normalise E_i to trace 1 so that the α_i sum up to 1.

Now we can find a relation between the α's and the $p(i)$'s:

$$p(i) = \text{tr}(\rho E_i) = \sum_j \alpha_j \text{tr}\left(\frac{E_j}{\text{tr}(E_j)} E_i \right) = \sum_j T(i|j) \alpha_j$$

where T is a matrix defined as

$$T_{ij} = T(i|j) = \text{tr}\left(\frac{E_j}{\text{tr}(E_j)} E_i \right).$$

T is a stochastic matrix: $\sum_i T(i|j) = 1$, and it is doubly stochastic if and only if all the E_i have the same trace, which then necessarily has to be $\frac{1}{n}$ because $n = \text{tr}(I_n) = \sum_{i=1}^{n^2} \text{tr}(E_i)$. We will refer to T as a *transition matrix* for $\{E_i\}$.

We have the relation $p = T\alpha$ where we consider p and α as vectors. T has to be invertible because the E_i form a basis, so $\alpha = T^{-1}p$, which means we can write ρ in terms of its probabilities over E_i:

$$\rho = \sum_i (T^{-1}p)_i \frac{E_i}{\text{tr}(E_i)}.$$

Note that while T is stochastic, T^{-1} will in every case contain some negative elements and so will be just quasi-stochastic [8]. This is not too surprising because if we had a POVM that has a T^{-1} that is stochastic we would have succeeded in finding a non-contextual hidden variable model for quantum theory.

Now suppose we have for M_m a minimal IC-POVM $\{E_i'\}$. And consider a completely positive trace preserving (CPTP) map $\Phi : M_n \to M_m$. Let $\sigma = \Phi(\rho)$. We wish to know the probability distribution of σ over E_i' in terms of the probability distribution of ρ over E_i. Using the expansion of ρ in terms of p:

---

[1] In fact, a quasi-POVM is nothing more than a Hermitian basis where the elements happen to sum up to the identity, we use the term 'quasi-POVM' to be consistent about the use of 'quasi-' to refer to an object that is usually positive, but in this case not.

$$q(i) = \mathrm{tr}\big(\sigma E_i'\big) = \mathrm{tr}\big(\Phi(\rho)E_i'\big) = \sum_j (T^{-1}p)_j\, \mathrm{tr}\left(\Phi\left(\frac{E_j}{\mathrm{tr}(E_j)}\right)E_i'\right) = \sum_j r(i|j)(T^{-1}p)_j$$

where we have introduced a matrix r defined as

$$r_{ij} = r(i|j) = \mathrm{tr}\left(\Phi\left(\frac{E_j}{\mathrm{tr}(E_j)}\right)E_i'\right).$$

The matrix r is again stochastic because Φ is trace preserving. The matrix is doubly stochastic if and only if the E_i and E_j' all have the same trace and Φ is unital.

We see that we have now translated the equation $\sigma = \Phi(\rho)$ into the equation $q = rT^{-1}p$. We can also translate composition of maps. Let $\Psi : M_m \to M_k$ be CPTP and $\{E_i''\}$ be a minimal IC-POVM on M_k. Set

$$T'(i|j) = \mathrm{tr}\left(\frac{E_j'}{\mathrm{tr}\big(E_j'\big)}E_i'\right), \qquad s(i|j) = \mathrm{tr}\left(\Psi\left(\frac{E_j'}{\mathrm{tr}\big(E_j'\big)}\right)E_i''\right).$$

Then setting $\sigma' = (\Psi \circ \Phi)(\rho) = \Psi(\Phi(\rho)) = \Psi(\sigma)$ and letting q' be the probability distribution associated to σ' we can derive in the same way as before that

$$q' = s(T')^{-1}q = s(T')^{-1}rT^{-1}p.$$

Since the map $\Psi \circ \Phi$ is completely determined by what it does on states this also completely determines the matrix that we should associate with it, namely: $s(T')^{-1}r$.

Note that the constructions in this section work equally well with quasi-POVMs, replacing probabilities with quasi-probabilities and stochastic matrices with quasi-stochastic matrices.

2 Functorial embedding from POVM's

We've seen how to translate composition of quantum channels into composition of quasi-stochastic matrices. To make this transition formal we will show that this induces a functor between the relevant categories.

Definition 3. CPTP is the category which has as objects the natural numbers $n > 0$, and its morphisms $\Phi : n \to m$ are CPTP maps $\Phi : M_n \to M_m$. Composition is the usual composition of linear maps. *States* are maps $\tilde{\rho} : 1 \to n$.

Note that the definition of states corresponds to the definition of state we have used above, because if $\tilde{\rho} : M_1 = \mathbb{C} \to M_n$ then it is completely defined by its action on 1, and we see that $\tilde{\rho}(1) = \rho \in M_n$ is positive and trace 1. We will most of the time simply use the actual states and not the morphism using the abuse of notation $\Phi(\rho) := \Phi \circ \tilde{\rho}$.

Definition 4. QStoch is the category which has as objects the natural numbers $n > 0$ and its morphisms $A : n \to m$ are quasi-stochastic matrices $A \in M^{n \times m}(\mathbb{R})$ and composition works by regular matrix multiplication. States $p : 1 \to n$ correspond to quasi-probability distributions.

We saw above that when we associate stochastic matrices s and r to Ψ and Φ, the composition $\Psi \circ \Phi$ has the matrix $sT^{-1}r$ associated to it, where T is determined by the choice of POVM on the intermediate space. We can capture this in a category by modifying the notion of composition: $s * r := sT^{-1}r$.

Definition 5. Let $T = (T^{(i)})_{i=1}^{\infty}$ be a family of invertible quasi-stochastic matrices where $T^{(i)}$ is a $i \times i$ square matrix. Define **QStoch**$_T$ to be the category with the same objects and morphisms as **QStoch** but with composition of $r : n \to m$ and $s : m \to k$ defined as $s * r = s(T^{(m)})^{-1}r$.

The associativity of this composition follows from the associativity of matrix multiplication and the new identity morphism for n is now $T^{(n)}$ so that **QStoch**$_T$ is indeed a category for any choice of invertible quasi-stochastic matrices.

Theorem 6. Fix for every finite dimension $n > 0$ a minimal IC-POVM $\{E_i^{(n)}\}$ and let $T = (T^{(i)})$ be the family of matrices where $T^{(i)} = T_n$ is the transition matrix for $\{E_j^{(n)}\}$ as defined above when $i = n^2$ and otherwise let it be the identity matrix. Then $Q : \textbf{CPTP} \to \textbf{QStoch}_T$ defined by $Q(n) = n^2$ and

$$Q(\Phi : n \to m)_{ij} = \operatorname{tr}\left(\Phi\left(\frac{E_j^{(n)}}{\operatorname{tr}\left(E_j^{(n)}\right)} \right) E_i^{(m)} \right)$$

is a faithful functor.

Proof. We have defined the composition in the category **QStoch**$_T$ precisely so that Q preserves composition: $Q(\Psi \circ \Phi) = Q(\Psi) * Q(\Phi)$, and it is easy enough to realise that $Q(id_n) = T_n$ which acts as the identity in **QStoch**$_T$, so that Q is indeed a functor.

Note that $M_1 = \mathbb{C}$ has a unique choice for a minimal IC-POVM, namely $\{1\}$. We then see that $Q(\tilde{\rho} : 1 \to n)_{i1} = \operatorname{tr}\left(\tilde{\rho}(1)E_i^{(n)} \right) = \operatorname{tr}\left(\rho E_i^{(n)} \right)$ which is just the Born rule as expected.

Since the E_i are informationally complete, all the states are sent to different probability distributions. Now suppose $Q(\Phi) = Q(\Psi)$. Then we also get $Q(\Phi \circ \rho) = Q(\Phi) * Q(\rho) = Q(\Psi) * Q(\rho) = Q(\Psi \circ \rho)$ so that we must have $\Phi \circ \rho = \Psi \circ \rho$ for all states ρ. Since a map is completely defined by its action on states we must then have $\Phi = \Psi$, so that Q is indeed faithful. $\qquad\square$

Note that this functor maps all states and maps to distributions and matrices with positive entries. All the negativity is hidden in the modified composition. This is analogous to the *modified probability calculus* of [8] and the *urgleichung* of [4]. We can bring this negativity more to the forefront using the following result.

Theorem 7. Let $T = (T^i)_{i=1}^{\infty}$ be a family of invertible quasi-stochastic matrices of size $i \times i$. The functor $F_T : \textbf{QStoch}_T \to \textbf{QStoch}$ defined by $F_T(n) = n$ and $F_T(A : n \to m) = A(T^{(n)})^{-1}$ is an isomorphism of categories.

Proof. The identity morphism of n in **QStoch**$_T$ is $T^{(n)}$, so that $F_T(id_n) = F_T(T^{(n)}) = T^{(n)}(T^{(n)})^{-1} = I_n$. That it preserves composition follows easily from the definition of composition in **QStoch**$_T$, so it is indeed a functor. That it is an isomorphism of categories follows because it has an inverse functor $F_T^{-1}(A) = AT^{(n)}$. $\qquad\square$

Now we see that the composition $F_T \circ Q$ gives for any family of minimal IC-POVM's an embedding of quantum theory into the category of quasi-stochastic maps. States are all mapped to proper probability distributions (no negative components), while effects do contain negative components. Instead of $F_T(A) = A(T^{(n)})^{-1}$ we could have also defined the isomorphism $F_T'(A) = (T^{(m)})^{-1}A$. In that case effects would be mapped to proper probabilistic effects and states would instead contain negative components.

There are a lot of ways to represent quantum theory as a quasi-stochastic theory, but one of the problems is that it is hard to find a reason to prefer one over the other. Using the functorial embedding we can make it clear why there doesn't seem to be a preferred one:

Theorem 8. Let $\{E_i^{(n)}\}$ and $\{F_i^{(n)}\}$ be families of minimal IC-POVM's, and let the T^1 and T^2 be the respective families of matrices as defined in Theorem 6, and let $Q_i : \textbf{CPTP} \to \textbf{QStoch}_{T_i}$ be the respective functors. Then there exists a natural isomorphism $\eta : F_{T^1} \circ Q_1 \Rightarrow F_{T^2} \circ Q_2$.

Proof. See appendix A. $\qquad\square$

Any minimal embedding is naturally isomorphic to another, which means that as far as the categories are concerned there really is only one embedding. Note that the category **QStoch** doesn't "see" negativity, so these embeddings can still be very different in terms of which maps are represented with negative components.

We've now shown that all quantum channels and all states can be represented in a coherent manner in terms of quasi-probabilities, but this is not really *all* of quantum theory. We should also consider as done in [4, 8] the probabilities arising from measuring a state using a POVM. So lets consider a measurement with K different outcomes where the probabilities for a state ρ are given as $P(k|\rho) = \text{tr}(A_k\rho)$ where $\{A_k\}_{k=1}^{K}$ denote a POVM on M_n. Expanding ρ in terms of its stochastic representation: $P(k|\rho) = \sum_i (T_n^{-1}Q(\rho))_i \text{tr}(A_k E_i / \text{tr}(E_i))$. It is helpful to consider the POVM $\{A_k\}$ as a map $\hat{A} : M_n \to \mathbb{C}^K$ in the category of C*-algebras where $\hat{A}(B)_k = \text{tr}(A_k B)$. Abusing notation a bit we can then write

$$Q(\hat{A})(k|i) = \langle e_k, \hat{A}\left(\frac{E_i}{\text{tr}(E_i)}\right)\rangle = \text{tr}\left(A_k \frac{E_i}{\text{tr}(E_i)}\right)$$

where we let Q map $\mathbb{C}^K$ to k and we take the standard basis $\{e_k\}$ as 'the POVM' for $\mathbb{C}^K$. This makes sense when we view $\mathbb{C}^K$ as the diagonal matrices in M_K in which case the standard basis components e_k correspond to the diagonal projections to the kth component. It is easy to check that $Q(\hat{A})$ is indeed a stochastic matrix and we see that $P(k|\rho) = \sum_i Q(\hat{A})(k|i)(T_n^{-1}Q(\rho))_i$ so that

$$P(\cdot|\rho) = Q(\hat{A})T_n^{-1}Q(\rho) = Q(\hat{A}) * Q(\rho).$$

This looks exactly the same as applying a CPTP map to a state. In this view measuring a POVM corresponds to a 'quantum-classical' channel that takes as input a quantum state and outputs a classical state (a probability distribution).

3   Preserving the tensor product

An important part of quantum theory is the possibility of parallel composition: the tensor product. This can be captured by the fact that the category of quantum processes **CPTP** is a (strict) *monoidal category*[2]:

Definition 9. A category $\mathbb{A}$ is called *strict monoidal* when it has a bifunctor $\otimes : \mathbb{A} \times \mathbb{A} \to \mathbb{A}$, and an identity object I, such that for all objects $A, B, C \in \mathbb{A}$: $I \otimes A = A \otimes I = A$ and $A \otimes (B \otimes C) = (A \otimes B) \otimes C$.

The functoriality condition boils down to $id_A \otimes id_B = id_{A \otimes B}$ and $(f_1 \otimes f_2) \circ (g_1 \otimes g_2) = (f_1 \circ g_1) \otimes (f_2 \circ g_2)$ for all compatible morphisms.

It is easy to verify that the linear algebraic tensor product turns **CPTP** and **QStoch** into monoidal categories, where on objects it acts as $n \otimes m = nm$.

The relevant notion of a morphism between monoidal categories is that of a strong monoidal functor.

---

[2]We will work exclusively with strict monoidal categories in this paper, so we will ignore the coherence isomorphisms that usually appear.

Definition 10. A functor between two strict monoidal categories $F : \mathbb{A} \to \mathbb{B}$ is called *strong monoidal* when the two functors $F_1(A,B) = F(A) \otimes F(B)$ and $F_2(A,B) = F(A \otimes B)$ are naturally isomorphic via $\alpha : F_1 \Rightarrow F_2$ such that the components $\alpha_{A,B} : F(A) \otimes F(B) \to F(A \otimes B)$ satisfy the coherence condition $\alpha_{A \otimes B,C} \circ (\alpha_{A,B} \otimes id_C) = \alpha_{A,B \otimes C} \circ (id_A \otimes \alpha_{B,C})$. The naturality condition means that $\alpha_{B_1,B_2} \circ (F(f_1) \otimes F(f_2)) = F(f_1 \otimes f_2) \circ \alpha_{A_1,A_2}$ for all morphisms $f_i : A_i \to B_i$.

The functor is called *strict monoidal* when all the α are identities and it is called *lax monoidal* when the α are not necessarily isomorphisms.

There are multiple choices for a tensor product in $\mathbf{QStoch}_T$. We will choose the tensor product such that the functor $F' : \mathbf{QStoch}_T \to \mathbf{QStoch}$ defined by $F'(A : n \to m) = T_m^{-1}A$ is strict monoidal. Denote the tensor product in $\mathbf{QStoch}_T$ by $\otimes'$, then we should have $F'(A \otimes' B) = F'(A) \otimes F'(B)$. Writing this out we get

$$T_{m_1 m_2}^{-1}(A \otimes' B) = (T_{m_1}^{-1} \otimes T_{m_2}^{-1})(A \otimes B)$$

so that $\otimes'$ should be defined as $A \otimes' B := T_{m_1 m_2}(T_{m_1}^{-1} \otimes T_{m_2}^{-1})(A \otimes B)$. It is instructive to check that this indeed turns $\mathbf{QStoch}_T$ into a monoidal category:

$$
\begin{aligned}
(A_1 \otimes' A_2) * (B_1 \otimes' B_2) &= T_{k_1 k_2}(T_{k_1}^{-1} \otimes T_{k_2}^{-1})(A_1 \otimes A_2) T_{m_1 m_2}^{-1} T_{m_1 m_2}(T_{m_1}^{-1} \otimes T_{m_2}^{-1})(B_1 \otimes B_2) \\
&= T_{k_1 k_2}(T_{k_1}^{-1} \otimes T_{k_2}^{-1})(A_1 T_{m_1}^{-1} B_1) \otimes (A_2 T_{m_2}^{-1} B_2) \\
&= (A_1 * B_1) \otimes' (A_2 * B_2)
\end{aligned}
$$

We can now check that the functor $F : \mathbf{QStoch}_T \to \mathbf{QStoch}$ is strong monoidal. We have $F(A \otimes' B) = T_{m_1 m_2}(T_{m_1}^{-1} \otimes T_{m_2}^{-1})(A \otimes B) T_{n_1 n_2}^{-1}$ and $F(A) \otimes F(B) = (A \otimes B)(T_{n_1}^{-1} \otimes T_{n_2}^{-1})$. This can be rewritten to

$$T_{m_1 m_2}(T_{m_1}^{-1} \otimes T_{m_2}^{-1})(F(A) \otimes F(B)) = F(A \otimes' B) T_{n_1 n_2}(T_{n_1}^{-1} \otimes T_{n_2}^{-1})$$

which means our coherence isomorphisms are $\alpha_{n_1,n_2} = T_{m_1 m_2}(T_{m_1}^{-1} \otimes T_{m_2}^{-1})$. It then follows by straightforward matrix multiplication that these satisfy the coherence conditions.

Note that we could have chosen the tensor product in $\mathbf{QStoch}_T$ such that F would be strict monoidal and F' would be strong monoidal. The reason we have chosen this tensor product is that it makes the following easier.

Theorem 11. The functor $Q : \mathbf{CPTP} \to \mathbf{QStoch}_T$ as defined in Theorem 6 for a family of minimal IC-POVM's is strong monoidal with coherence isomorphisms $\alpha_{n_1,n_2} = S_{n_1,n_2}$, where

$$S_{n_1,n_2}(j|i_1 i_2) = \mathrm{tr}\left(\frac{E_{i_1}^{n_1}}{\mathrm{tr}(E_{i_1}^{n_1})} \otimes \frac{E_{i_2}^{n_2}}{\mathrm{tr}(E_{i_2}^{n_2})} E_j^{n_1 n_2} \right)$$

Proof. See appendix B.   □

The composition of two strongly monoidal functors is again strongly monoidal, so $F \circ Q$ gives a strong monoidal functor of $\mathbf{CPTP}$ into $\mathbf{QStoch}$. The functor is *strong* monoidal precisely because the S_{n_1,n_2} are invertible and therefore are isomorphisms. These S are invertible because the POVM's are minimal. If they weren't minimal we wouldn't get a strong monoidal functor, but at most a *lax* monoidal functor.

4   Preserving the adjoint

Another structure that exists on quantum channels is the *adjoint*. If we have a CP map $\Phi : M_n \to M_m$, its adjoint is the dual of the map with respect to the Hilbert-Schmidt inner product. That is, a map $\Phi^\dagger : M_m \to M_n$ (note that the direction of the map has changed) such that

$$\langle \Phi(A), B \rangle_{HS} = \mathrm{Tr}\big(\Phi(A)B^\dagger\big) = \mathrm{Tr}\big(A(\Phi^\dagger(B))^\dagger\big) = \langle A, \Phi^\dagger(B) \rangle_{HS}.$$

If $\Phi(A) = UAU^\dagger$, then $\Phi^\dagger(A) = U^\dagger A U$, e.g., for unitary conjugation the adjoint is simply the Hermitian adjoint of the unitary. Hamiltonian evolution in time is represented by the unitary $\exp(itH)$ where the Hermitian adjoint is $\exp(-itH)$. The adjoint can therefore be interpreted as reversing the time direction. It is also how one transfers between the Schrödinger and Heisenberg picture of quantum mechanics. The adjoint has the property that it distributes over composition, while changing the order of the morphisms: $(\Psi \circ \Phi)^\dagger = \Phi^\dagger \circ \Psi^\dagger$. This behaviour can be defined in a more general way:

Definition 12. A *dagger category* $\mathbb{A}$ is a category which has a contravariant functor $\dagger : \mathbb{A} \to \mathbb{A}$ that is the identity on objects and is its own inverse. Concretely this means that $id_A^\dagger = id_A$ and for all morphisms $(f^\dagger)^\dagger = f$, while composition satisfies $(g \circ f)^\dagger = f^\dagger \circ g^\dagger$. We will call a category a *partial dagger category*, when the dagger is only defined for a subset of the morphisms. A *dagger functor* is a functor between (partial) dagger categories $F : \mathbb{A} \to \mathbb{B}$ that preserves the dagger (when it exists) in the following way: $F \circ \dagger_\mathbb{A} = \dagger_\mathbb{B} \circ F$, e.g. $F(f^{\dagger_\mathbb{A}}) = (F(f))^{\dagger_\mathbb{B}}$.

The category **CPTP** is not a dagger category as the adjoint of a trace preserving map Φ will only be trace preserving when Φ is unital. We could consider the category of trace preserving unital maps, but this is a bit restrictive in that it doesn't have morphisms between different objects. Therefore we will stick with **CPTP** as a partial dagger category and only define the dual of a map when it is unital.

QStoch is also only a partial dagger category with the transpose taking the role of the dagger. The transpose of a quasi-stochastic matrix is only quasi-stochastic when it is doubly quasi-stochastic (this is the exact analogue of trace preserving unital maps). We could again 'fix' this by only considering doubly quasi-stochastic maps, but instead we will only define the dagger when a map is doubly quasi-stochastic.

QStoch$_T$ doesn't always have a partial dagger structure. The obvious choice is the transpose: $(A * B)^T = (AT^{-1}B)^T = B^T(T^T)^{-1}A^T$. We need $(A * B)^T = B^T * A^T$ so that equality here only holds when $T = T^T$. Looking at the definition of the transition matrix T this is the case when all the components of the POVM have equal trace. In that case we define the partial dagger on **QStoch**$_T$ to be the transpose as defined on doubly quasi-stochastic matrices.

Lemma 13. Let $\{E_i^{(n)}\}$ be a family of minimal IC-POVM's where all the components in a single POVM have equal trace, Then the functor Q as in Theorem 6 is a dagger functor.

Proof. The dagger in **CPTP** is only defined when $\Phi : M_n \to M_n$ is CPTP unital, which necessitates that the input and output dimension are equal. Let $\{E_i\}$ be the POVM for n (dropping the superscript). We have $\mathrm{tr}(E_i) = \frac{1}{n}$ for all i, because the POVM is minimal and all the traces are equal. So now

$$Q(\Phi^\dagger)_{ij} = \mathrm{tr}\left(\Phi^\dagger \left(\frac{E_j}{\mathrm{tr}(E_j)} \right) E_i \right) = \frac{1}{n} \mathrm{tr}\big(E_i \Phi^\dagger(E_j)\big) = \frac{1}{n} \mathrm{tr}\big(E_i(\Phi^\dagger(E_j))^\dagger\big) = \frac{1}{n} \mathrm{tr}\left(\Phi(E_i)E_j^\dagger \right) = Q(\Phi)_{ji}$$

where $E_i^\dagger = E_i$, because they are by definition Hermitian. Indeed $Q(\Phi^\dagger) = Q(\Phi)^T$ as required. $\qquad\square$

Somewhat surprisingly this won't in general extend to a dagger functor from **CPTP** to **QStoch** by appending Q with $F : \mathbf{QStoch}_T \to \mathbf{QStoch}$. Let A be a $n \times n$ doubly quasi-stochastic matrix and assume that T is a symmetric matrix. Recall that $F(A) = AT_n^{-1}$ which gives $F(A^T) = A^T T_n^{-1}$ while $F(A)^T = (AT_n^{-1})^T = T_n^{-1} A^T$. Now, of course in general $A^T T_n^{-1} \neq T_n^{-1} A^T$ so that this functor is *not* a dagger functor. In fact, it will only be a dagger functor if T commutes with all quasi-stochastic matrices.

Lemma 14. Let J_n denote the $n \times n$ matrix with every component equal to 1: $(J_n)_{ij} = 1$. Suppose the $n \times n$ doubly quasi-stochastic matrix T commutes with all $n \times n$ doubly quasi-stochastic matrices, then there are real constants α, β such that $T = \alpha I_n + \beta J_n$ where $\alpha + n\beta = 1$.

Proof. All matrices will be $n \times n$ square matrices. We note that a matrix A is quasi-stochastic if and only if $JA = J$. It is furthermore doubly quasi-stochastic if and only if $AJ = JA = J$. J is obviously a rank 1 matrix, which has a single non-zero eigenvector: $\mathbf{1} = (1, \dots, 1)$, such that $J\mathbf{1} = n\mathbf{1}$. We therefore see that $P = \frac{1}{n}J$ is a 1-dimensional projection, and that the linear subspace spanned by the doubly quasi-stochastic matrices is exactly $\{A; [A, P] = 0\}$. This commutation essentially fixes one of the eigenvalues of the matrices, but apart from that doesn't require anything extra. This space is therefore isomorphic to $\mathbb{R} \times M_{n-1}$. Now any doubly quasi-stochastic matrix A can be written as $PAP + (I_n - P)A(I_n - P)$. Therefore for T to commute with a doubly quasi-stochastic matrix A we need $(I_n - P)T(I_n - P)$ to commute with $(I_n - P)A(I_n - P)$ (since it automatically commutes on the other part). But $(I_n - P)A(I_n - P)$ is just an arbitrary $(n-1) \times (n-1)$ matrix. We know that the only matrices in the space M_k that commute with all the matrices are multiples of the identity. We therefore have $(I_n - P)T(I_n - P) = \alpha(I_n - P)$ (noting that $I_n - P$ acts as the identity on this space). And so in fact

$$T = \alpha(I_n - P) + \beta P = \alpha I_n + (\beta - \alpha)P = \alpha I_n + \frac{1}{n}(\beta - \alpha)J$$

which with a proper relabelling becomes $T = \alpha I_n + \beta J$. T is automatically symmetric, and it is quasi-stochastic exactly when $\alpha + n\beta = 1$. $\square$

So for F to be a dagger functor, the transition matrices have to be particularly simple: $T_n = \alpha_n I_n + \beta_n J_n$. Let $\{E_i\}$ be a minimal IC-POVM for M_n. Suppose its transition matrix has that form, then

$$T_{ij} = \mathrm{tr}\left(\frac{E_j}{\mathrm{tr}(E_j)} E_i \right) = \alpha \delta_{ij} + \beta.$$

We note that the right-hand side is symmetric, so that the left-hand side has to be as well, which means that $\mathrm{tr}(E_j) = \frac{1}{n}$ by minimality. This POVM has a very specific symmetry property.

Definition 15. We say that a POVM $\{E_i\}$ is *generalised symmetric informationally complete* (generalised SIC) for M_n when it is minimal informationally complete and

$$\mathrm{tr}(E_i E_j) = \alpha \delta_{ij} + \beta$$

for some constants α and β. It is called SIC (not generalised) when all the components are rank-1: $E_i = \frac{1}{n}\Pi_i$ for some projections Π_i. For such a SIC the constants need to be $\alpha = \frac{1}{n^2}$ and $\beta = \frac{n-1}{n^4}$.

The existence and behaviour of such POVM's has been subject of intense study especially in the case of the rank 1 SIC's where the existence in arbitrary dimension is still not proven[4, 18]. The generalised SIC's have been classified[12] and there exist many of them for each dimension.

With the previous lemma and our observations about the functor preserving the dagger we get the following theorem.

Theorem 16. Let Q be the functor from Theorem 6. $F \circ Q : \textbf{CPTP} \to \textbf{QStoch}$ preserves the dagger if and only if all POVM's are generalised SIC.

Proof. Q always preserves the dagger, so $F \circ Q$ only preserves the dagger when F does. We've already seen that F preserves the dagger if and only if all the transition matrices have the special symmetric form corresponding to generalised SIC-POVMs. $\square$

Recalling that the adjoint of a map can be interpreted as its time reversal, this gives an interpretation of SIC-POVMs as being the only POVMs 'preserving the arrow of time' in the sense that the image of the time-reversal of a map is the time-reversal of its image.

5 General functorial embeddings

So far we have only considered minimal IC-POVM's: A POVM the components of which form a basis for the underlying matrix space. However we can also consider POVM's consisting of more components. In that case the components won't be linearly independent so that a state can be represented in multiple ways as a combination of the POVM elements:

$$\rho = \sum_i \alpha_i \frac{E_i}{\text{tr}(E_i)} = \sum_i \alpha_i' \frac{E_i}{\text{tr}(E_i)}.$$

This means that

$$\text{tr}(\rho E_i) = \sum_j \alpha_j \text{tr}\left(\frac{E_j}{\text{tr}(E_j)} E_i\right) = \sum_j T(i|j)\alpha_j = \sum_j T(i|j)\alpha_j'$$

or as vectors: $p = T\alpha = T\alpha'$. This is the case because T is no longer a full rank matrix so that it has a nontrivial kernel. Now we can consider the generalised inverse of T defined as having the same kernel as T but otherwise acting as an inverse which gives a canonical choice of α: $\alpha = T^{-1}p$.

We can also still transfer composition: $\Phi(\rho) \mapsto Q(\Phi)T^{-1}Q(\rho)$. This is well-defined in the sense that $\ker(T) \subset \ker(Q(\Phi))$, but we can no longer define a category $\textbf{QStoch}_T$ as there is no longer an identity morphism: $A * T \neq A$, because $T^{-1}T \neq I_n$. The composite functor $F \circ Q$ where we map $\Phi \mapsto Q(\Phi)T^{-1}$ also doesn't work any more as we still don't map the identity to the identity. However, we could consider that there are some smart choices which do produce a valid functor. For instance there is a priori no reason that we can't let $id \mapsto I$. To preserve the convex structure we should then also change where the other maps are sent to.

Suppose we have such a functor, how much of what we previously constructed still goes through? Looking at Theorem 8 we see that we no longer get a natural isomorphism as the matrices involved are no longer invertible. We also see that the proof of Theorem 11 explicitly requires the minimality of the representation since otherwise the morphisms involved wouldn't be isomorphisms. The most a nonminimal representation can therefore be is lax monoidal. There are no obvious barriers to a nonminimal representation preserving the adjoint.

So far we have studied embeddings of **CPTP** into **QStoch** coming from POVMs. Let's consider a bit more general view.

Definition 17. We call $F : \textbf{CPTP} \to \textbf{QStoch}$ a *quasi-stochastic representation* of **CPTP** when F is a functor that preserves the convex structure of the categories: $F(t\Phi_1 + (1-t)\Phi_2) = tF(\Phi_1) + (1-t)F(\Phi_2)$ and maps 1 to 1 (so that states are sent to states). We call it a *positive representation* when it sends all states to nonnegative distributions. A representation is called *strong monoidal / faithful / dagger*

preserving when the functor is. We call a representation *trivial* when for all $n \in \mathbb{N}$ $F(\rho) = F(\sigma)$ for all states $\rho, \sigma \in M_n$.

Lemma 18. Let $F : \textbf{CPTP} \rightarrow \textbf{QStoch}$ be a quasi-stochastic representation. Then we can find for every $n \in \mathbb{N}$ a quasi-POVM $\{E_i^{(n)}\}$ such that for a state $\rho \in M_n$ we have $F(\rho)_i = \text{tr}\left(\rho E_i^{(n)}\right)$.

Proof. Let $n \in \mathbb{N}$. We see that F restricts to a convex map $f^{(n)} : \text{DO}(n) \rightarrow \text{QProb}(F(n))$ where $\text{DO}(n) \subset M_n$ represents the set of density operators and $\text{QProb}(k_n)$ is the set of quasi-probability distributions on $F(n)$ points. We can split this map up into its components $f_i^{(n)} : \text{DO}(n) \rightarrow \mathbb{R}$. These functions are again convex and using standard methods we can extend these first to linear maps on Hermitian matrices and then to linear maps to all matrices: $f_i^{(n)} : M_n \rightarrow \mathbb{C}$. This is exactly an element of the dual space of M_n and since this is finite dimensional it is isomorphic to M_n which means there exists a $E_i^{(n)} \in M_n$ such that $f^{(n)}(A) = \langle A, E_i^{(n)} \rangle_{HS} = \text{tr}\left(A(E_i^{(n)})^\dagger\right)$. Because this function is real valued on the Hermitian matrices, $E_i^{(n)}$ has to be Hermitian and because we have $1 = \sum_i f_i^{(n)}(\rho) = \text{tr}\left(\rho \sum_i E_i^{(n)}\right)$ we get $\sum_i E_i^{(n)} = I_n$ proving that $\{E_i^{(n)}\}$ is indeed a quasi-POVM. $\square$

We will refer to these quasi-POVM's as the representation's associated quasi-POVM's. It is easy to see that the representation is positive if and only if the associated quasi-POVM's are true POVM's (consisting of only positive components).

Note: If we start out with a family of POVM's and construct the regular functor $Q : \textbf{CPTP} \rightarrow \textbf{QStoch}_T$ and then append this with the functor $F' : \textbf{QStoch}_T \rightarrow \textbf{QStoch}$ given by $F'(A) = T^{-1}A$ then $F' \circ Q$ will not be a positive representation, while $F \circ Q$ where $F(A) = AT^{-1}$ will be. $F' \circ Q$ and $F \circ Q$ are naturally isomorphic, so positivity of the representation is not a 'categorical' property in the sense that it isn't preserved by natural isomorphism.

It is not hard to see that a representation is faithful if and only if the associated quasi-POVM's are informationally complete. On the other hand, suppose a representation is trivial, then for a element of the quasi-POVM $E \in M_n$ we must have $\text{tr}(\rho E) = \text{tr}(\sigma E)$ for all states $\rho, \sigma \in M_n$. Calling this value α and realising that $\alpha = \text{tr}(\rho \alpha I_n)$ we see that $\text{tr}(\rho(E - \alpha I_n)) = 0$ for all states ρ so that $E = \alpha I_n$. Perhaps somewhat surprisingly these are the only options for the associated POVM's:

Theorem 19. Let $F : \textbf{CPTP} \rightarrow \textbf{QStoch}$ be a quasi-stochastic representation of **CPTP**. This representation is either faithful or trivial.

Proof. See appendix C. $\square$

This means that we immediately get the following corollary using Theorems 11 and 16.

Corollary 20. Let $F : \textbf{CPTP} \rightarrow \textbf{QStoch}$ be a positive nontrivial representation.

- F is strong monoidal if and only if the associated POVM's are minimal IC.

- F preserves the dagger if and only if the associated POVM's are generally symmetric IC.

The above theorem and corollary show that any nontrivial representation must be induced by a family of informationally complete (quasi-)POVMs. It is not clear whether these representations must also be minimal. It could very well be that the proof of Theorem 19 can be adapted to show that any representation must necessarily be minimal to preserve functoriality. If this is the case then we automatically get that the object n in **CPTP** is mapped to n^2 in **QStoch** providing a natural way of viewing the difference between the amount of perfectly distinguishable states (n) versus the dimension of the state space (n^2). This specific relation between these values is identified as something particular to quantum theory in for instance [6, 3, 13]. Real-valued (or quaternion valued) quantum theory would have a different value here.

References

[1] Samson Abramsky & Adam Brandenburger (2011): *The sheaf-theoretic structure of non-locality and contextuality*. New Journal of Physics 13(11), p. 113036, doi:10.1088/1367-2630/13/11/113036.

[2] Samson Abramsky & Adam Brandenburger (2014): *An operational interpretation of negative probabilities and no-signalling models*. In: Horizons of the Mind. A Tribute to Prakash Panangaden, Springer, pp. 59–75, doi:10.1103/PhysRev.40.749.

[3] DM Appleby, Åsa Ericsson & Christopher A Fuchs (2011): *Properties of QBist state spaces*. Foundations of Physics 41(3), pp. 564–579, doi:10.1007/s10701-010-9458-7.

[4] Marcus Appleby, Christopher A Fuchs, Blake C Stacey & Huangjun Zhu (2017): *Introducing the Qplex: a novel arena for quantum theory*. The European Physical Journal D 71(7), p. 197, doi:10.1140/epjd/e2017-80024-y.

[5] Jonathan Barrett, Niel de Beaudrap, Matty J Hoban & Ciarán M Lee (2017): *The computational landscape of general physical theories*. arXiv preprint arXiv:1702.08483.

[6] Giulio Chiribella, Giacomo Mauro D'Ariano & Paolo Perinotti (2011): *Informational derivation of quantum theory*. Physical Review A 84(1), p. 012311, doi:10.1103/PhysRevLett.103.170502.

[7] Christopher Ferrie (2011): *Quasi-probability representations of quantum theory with applications to quantum information science*. Reports on Progress in Physics 74(11), p. 116001, doi:10.1088/0034-4885/74/11/116001.

[8] Christopher Ferrie & Joseph Emerson (2008): *Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations*. Journal of Physics A: Mathematical and Theoretical 41(35), p. 352001, doi:10.1088/1751-8113/41/35/352001.

[9] Richard P Feynman (1982): *Simulating physics with computers*. International journal of theoretical physics 21(6), pp. 467–488, doi:10.1007/BF02650179.

[10] Christopher A Fuchs (2017): *Notwithstanding Bohr, the Reasons for QBism*. arXiv preprint arXiv:1705.03483.

[11] Christopher A Fuchs & Rüdiger Schack (2013): *Quantum-bayesian coherence*. Reviews of Modern Physics 85(4), p. 1693, doi:10.1016/j.shpsb.2005.05.005.

[12] Gilad Gour & Amir Kalev (2014): *Construction of all general symmetric informationally complete measurements*. Journal of Physics A: Mathematical and Theoretical 47(33), p. 335302, doi:10.1088/1751-8113/47/33/335302.

[13] Lucien Hardy (2001): *Quantum theory from five reasonable axioms*. arXiv preprint quant-ph/0101012.

[14] Lucien Hardy (2013): *A formalism-local framework for general probabilistic theories, including quantum theory*. Mathematical Structures in Computer Science 23(02), pp. 399–440, doi:10.1007/978-3-642-70029-3.

[15] Ciarán M Lee & Jonathan Barrett (2015): *Computation in generalised probabilisitic theories*. New Journal of Physics 17(8), p. 083001, doi:10.1088/1367-2630/17/8/083001.

[16] A Mari & Jens Eisert (2012): *Positive Wigner functions render classical simulation of quantum computation efficient*. Physical review letters 109(23), p. 230503, doi:10.1103/PhysRevA.79.062302.

[17] Hakop Pashayan, Joel J Wallman & Stephen D Bartlett (2015): *Estimating outcome probabilities of quantum circuits using quasiprobabilities*. Physical review letters 115(7), p. 070501, doi:10.1088/1464-4266/6/10/003.

[18] Andrew James Scott & Markus Grassl (2010): *Symmetric informationally complete positive-operator-valued measures: A new computer study*. Journal of Mathematical Physics 51(4), p. 042203, doi:10.1006/jsco.1996.0125.

[19] Peter Selinger (2007): *Dagger compact closed categories and completely positive maps*. Electronic Notes in Theoretical Computer Science 170, pp. 139–163, doi:10.1016/j.entcs.2006.12.018.

[20] Robert W Spekkens (2008): *Negativity and contextuality are equivalent notions of nonclassicality*. Physical review letters 101(2), p. 020401, doi:10.1023/B:FOOP.0000019581.00318.a5.

[21] Robert W Spekkens (2016): *Quasi-quantization: classical statistical theories with an epistemic restriction.* In: *Quantum Theory: Informational Foundations and Foils,* Springer, pp. 83–135, doi:10.1007/978-94-017-7303-4˙4.

[22] Victor Veitch, Christopher Ferrie, David Gross & Joseph Emerson (2012): *Negative quasi-probability as a resource for quantum computation.* New Journal of Physics 14(11), p. 113011, doi:10.1088/1367-2630/14/11/113011.

[23] Victor Veitch, Nathan Wiebe, Christopher Ferrie & Joseph Emerson (2013): *Efficient simulation scheme for a class of quantum optics experiments with non-negative Wigner representation.* New Journal of Physics 15(1), p. 013037, doi:10.1088/1367-2630/15/1/013037.

[24] Eugene Wigner (1932): *On the quantum correction for thermodynamic equilibrium.* Physical review 40(5), p. 749, doi:10.1103/PhysRev.40.749.

[25] Huangjun Zhu (2016): *Quasiprobability representations of quantum mechanics with minimal negativity.* Physical Review Letters 117(12), p. 120404, doi:10.1016/j.aop.2015.08.005.

A Proof of Theorem 8

Theorem. Let $\{E_i^{(n)}\}$ and $\{F_i^{(n)}\}$ be families of minimal IC-POVM's, and let the T^1 and T^2 be the respective families of matrices as defined in Theorem 6, and let $Q_i : \textbf{CPTP} \to \textbf{QStoch}_{T_i}$ be the respective functors. Then there exists a natural isomorphism $\eta : F_{T^1} \circ Q_1 \Rightarrow F_{T^2} \circ Q_2$.

Proof. We need a family of transition matrices from the POVM $\{E_i^{(n)}\}$ to $\{F_i^{(n)}\}$. So first note that there is a matrix $\alpha^{(n)}$ such that

$$\frac{E_j^{(n)}}{\text{tr}\left(E_j^{(n)}\right)} = \sum_k \alpha^{(n)}(k|j) \frac{F_k^{(n)}}{\text{tr}\left(F_k^{(n)}\right)}.$$

Then define the matrix $S^{(n)}$ by

$$S^{(n)}(i|j) = \text{tr}\left(\frac{E_j^{(n)}}{\text{tr}\left(E_j^{(n)}\right)} F_i^{(n)}\right) = \sum_k \alpha^{(n)}(k|j) \text{tr}\left(\frac{F_k^{(n)}}{\text{tr}\left(F_k^{(n)}\right)} F_i^{(n)}\right) = \sum_k \alpha^{(n)}(k|j) T_n^2(i|k)$$

which means that $S^{(n)} = T_n^2 \alpha^{(n)}$ or equivalently (using the fact that the POVM's are minimal so that their transition matrices are invertible) that $\alpha^{(n)} = \left(T_n^2\right)^{-1} S^{(n)}$.

Now given a CPTP map $\Phi : M_n \to M_m$ we can write

$$Q_1(\Phi)(j|i) = \text{tr}\left(\Phi\left(\frac{E_i^{(n)}}{\text{tr}\left(E_i^{(n)}\right)}\right) E_j^{(m)}\right) = \text{tr}\left(E_j^{(m)}\right) \text{tr}\left(\Phi\left(\frac{E_i^{(n)}}{\text{tr}\left(E_i^{(n)}\right)}\right) \frac{E_j^{(m)}}{\text{tr}\left(E_j^{(m)}\right)}\right)$$

$$= \text{tr}\left(E_j^{(m)}\right) \sum_{i',j'} \alpha^{(n)}(i'|i)\alpha^{(m)}(j'|j) \frac{1}{\text{tr}\left(F_{j'}^{(m)}\right)} \text{tr}\left(\Phi\left(\frac{F_{i'}^{(n)}}{\text{tr}\left(F_{i'}^{(n)}\right)}\right) F_{j'}^{(m)}\right)$$

$$= \sum_{i',j'} \alpha^{(n)}(i'|i)\alpha^{(m)}(j'|j) \frac{\text{tr}\left(E_j^{(m)}\right)}{\text{tr}\left(F_{j'}^{(m)}\right)} Q_2(\Phi)(j'|i').$$

This can be simplified by defining

$$\beta^{(m)}(i|j) := \alpha^{(m)}(j|i)\frac{\text{tr}\left(E_i^{(m)}\right)}{\text{tr}\left(F_j^{(m)}\right)}.$$

Note that the indexes of α and β are reversed. This is intentional as this allows us to write

$$Q_1(\Phi)(j|i) = \sum_{j'} \beta^{(m)}(j|j')\sum_{i'} Q_2(\Phi)(j'|i')\alpha^{(n)}(i'|i) = \sum_{j'} \beta^{(m)}(j|j')\left(Q_2(\Phi)\alpha^{(n)}\right)(j'|i)$$

$$= \left(\beta^{(m)}Q_2(\Phi)\alpha^{(n)}\right)(j|i).$$

What is this β? If we take $\Phi = id_n$ we get $Q_i(\Phi) = T_n^i$. By using $\alpha^{(n)} = \left(T_n^2\right)^{-1}S^{(n)}$ we then get

$$T_n^1 = Q_1(\Phi) = \beta^{(n)}Q_2(\Phi)\alpha^{(n)} = \beta^{(n)}T_n^2\left(T_n^2\right)^{-1}S^{(n)} = \beta^{(n)}S^{(n)}.$$

$S^{(n)}$ has an inverse since it acts as a basis transformation between the two POVM's. This allows us to write $\beta^{(n)} = \left(S^{(n)}\right)^{-1}T_n^1$ which means we get the following expression:

$$Q_1(\Phi) = \left(S^{(m)}\right)^{-1}T_m^1 Q_2(\Phi)\left(T_n^2\right)^{-1}S^{(n)}.$$

Now by multiplying this equation with $\left(T_n^1\right)^{-1}$ on the right and using $(F_i \circ Q_i)(\Phi) = Q_i(\Phi)\left(T_n^i\right)^{-1}$ this translates to

$$(F_1 \circ Q_1)(\Phi) = \left(S^{(m)}\right)^{-1}T_m^1(F_2 \circ Q_2)(\Phi)S^{(n)}\left(T_n^1\right)^{-1}.$$

Define $\eta_n := S^{(n)}\left(T_n^1\right)^{-1}$. The equation can now be recast as

$$\eta_m(F_1 \circ Q_1)(\Phi) = (F_2 \circ Q_2)(\Phi)\eta_n$$

which means the collection of η's forms a natural transformation. For all n, $S^{(n)}$ and $T^{(n)}$ are both invertible, so that $\eta^{(n)}$ is as well. Hence η is a natural isomorphism. $\qquad\square$

B Proof of Theorem 11

Theorem. The functor $Q : \textbf{CPTP} \to \textbf{QStoch}_T$ as defined in Theorem 6 for a family of minimal IC-POVM's is strong monoidal with coherence isomorphisms $\alpha_{n_1,n_2} = S_{n_1,n_2}$ where

$$S_{n_1,n_2}(j|i_1 i_2) = \text{tr}\left(\frac{E_{i_1}^{n_1}}{\text{tr}\left(E_{i_1}^{n_1}\right)} \otimes \frac{E_{i_2}^{n_2}}{\text{tr}\left(E_{i_2}^{n_2}\right)}E_j^{n_1 n_2}\right)$$

Proof. Let ρ_i be states in M_{n_i}, then we have $\rho_i = \sum_j (T_{n_i}^{-1}Q(\rho_i))_j\frac{E_j^{n_i}}{\text{tr}(E_j^{n_i})}$ which means we can write

$$Q(\rho_1 \otimes \rho_2)_j = \text{tr}\left((\rho_1 \otimes \rho_2)E_j^{n_1 n_2}\right) = \sum_{i_1,i_2}(T_{n_1}^{-1}Q(\rho_1))_{i_1}(T_{n_2}^{-1}Q(\rho_2))_{i_2}\text{tr}\left(\frac{E_{i_1}^{n_1}}{\text{tr}\left(E_{i_1}^{n_1}\right)} \otimes \frac{E_{i_2}^{n_2}}{\text{tr}\left(E_{i_2}^{n_2}\right)}E_j^{n_1 n_2}\right)$$

$$= \sum_{i_1,i_2}S_{n_1,n_2}(j|i_1,i_2)(T_{n_1}^{-1}Q(\rho_1))_{i_1}(T_{n_2}^{-1}Q(\rho_2))_{i_2}.$$

This can now be written as $Q(\rho_1 \otimes \rho_2) = S_{n_1,n_2}(T_{n_1}^{-1} \otimes T_{n_2}^{-1})(Q(\rho_1) \otimes Q(\rho_2))$ and substituting the definition of the tensor product of **QStoch**$_T$ $\otimes'$ it becomes

$$Q(\rho_1 \otimes \rho_2) = S_{n_1,n_2}T_{n_1n_2}^{-1}(Q(\rho_1) \otimes' Q(\rho_2)) = S_{n_1,n_2} * (Q(\rho_1) \otimes' Q(\rho_2)).$$

Now fix CPTP maps $\Phi_i : M_{n_i} \to M_{m_i}$ and write

$$\begin{aligned}
Q(\Phi_1 \otimes \Phi_2) * Q(\rho_1 \otimes \rho_2) &= Q(\Phi_1(\rho_1) \otimes \Phi_2(\rho_2)) = S_{m_1,m_2} * (Q(\Phi_1(\rho_1)) \otimes' Q(\Phi_2(\rho_2))) \\
&= S_{m_1,m_2} * (Q(\Phi_1) \otimes' Q(\Phi_2)) * (Q(\rho_1) \otimes' Q(\rho_2)) \\
&= S_{m_1,m_2} * (Q(\Phi_1) \otimes' Q(\Phi_2))S_{n_1,n_2}^{-1}T_{n_1n_2} * Q(\rho_1 \otimes \rho_2).
\end{aligned}$$

Because this has to hold for all states ρ_i we can drop that term and we get the equality $Q(\Phi_1 \otimes \Phi_2) = S_{m_1,m_2} * (Q(\Phi_1) \otimes' Q(\Phi_2))S_{n_1,n_2}^{-1}T_{n_1n_2}$. Now by bringing some terms to the other side and using the definition of $*$ again we get the desired naturality equation:

$$S_{m_1,m_2} * (Q(\Phi_1) \otimes' Q(\Phi_2)) = Q(\Phi_1 \otimes \Phi_2) * S_{n_1,n_2}.$$

We still need to check that the coherence condition holds:

$$S_{n_1n_2,n_3} * (S_{n_1,n_2} \otimes' id_{n_3}) = S_{n_1,n_2n_3} * (id_{n_1} \otimes' S_{n_2,n_3}) \iff$$
$$S_{n_1n_2,n_3}T_{n_1n_2n_3}^{-1}T_{n_1n_2n_3}(T_{n_1n_2}^{-1} \otimes T_{n_3}^{-1})(S_{n_1,n_2} \otimes T_{n_3}) = S_{n_1,n_2n_3}T_{n_1n_2n_3}^{-1}T_{n_1n_2n_3}(T_{n_1}^{-1} \otimes T_{n_2n_3}^{-1})(T_{n_1} \otimes S_{n_2,n_3}) \iff$$
$$S_{n_1n_2,n_3}(T_{n_1n_2}^{-1}S_{n_1,n_2} \otimes I_{n_3}) = S_{n_1,n_2n_3}(I_{n_1} \otimes T_{n_2n_3}^{-1}S_{n_2,n_3}).$$

To prove this equality we need to write out $S_{n_1n_2,n_3}$ and to do that we first need the following: define β such that

$$\frac{E_{i_1}^{n_1n_2}}{\text{tr}(E_{i_1}^{n_1n_2})} = \sum_{k_1,k_2} \beta(k_1,k_2|i_1)\frac{E_{k_1}^{n_1}}{\text{tr}(E_{k_1}^{n_1})} \otimes \frac{E_{k_2}^{n_2}}{\text{tr}(E_{k_2}^{n_2})}$$

which then gives

$$\begin{aligned}
T_{n_1n_2}(l|i_1) = \text{tr}\left(\frac{E_{i_1}^{n_1n_2}}{\text{tr}(E_{i_1}^{n_1n_2})}E_l^{n_1n_2}\right) &= \sum_{k_1,k_2} \beta(k_1,k_2|i_1)\,\text{tr}\left(\frac{E_{k_1}^{n_1}}{\text{tr}(E_{k_1}^{n_1})} \otimes \frac{E_{k_2}^{n_2}}{\text{tr}(E_{k_2}^{n_2})}E_l^{n_1n_2}\right) \\
&= \sum_{k_1k_2} \beta(k_1,k_2|i_1)S_{n_1,n_2}(l|k_1,k_2) = (S_{n_1,n_2}\beta)(l|i_1),
\end{aligned}$$

so that $\beta = S_{n_1,n_2}^{-1}T_{n_1n_2}$. Now we can write

$$\begin{aligned}
S_{n_1n_2,n_3}(j|i_1,i_2) &= \text{tr}\left(\frac{E_{i_1}^{n_1n_2}}{\text{tr}(E_{i_1}^{n_1n_2})} \otimes \frac{E_{i_2}^{n_3}}{\text{tr}(E_{i_2}^{n_3})}E_j^{n_1n_2n_3}\right) \\
&= \sum_{k_1,k_2} \beta(k_1,k_2|i_1)\,\text{tr}\left(\frac{E_{k_1}^{n_1}}{\text{tr}(E_{k_1}^{n_1})} \otimes \frac{E_{k_2}^{n_2}}{\text{tr}(E_{k_2}^{n_2})} \otimes \frac{E_{i_2}^{n_3}}{\text{tr}(E_{i_2}^{n_3})}E_j^{n_1n_2n_3}\right).
\end{aligned}$$

Defining the quantity in the trace to be $S_{n_1,n_2,n_3}(j|k_1,k_2,i_2)$ (note the comma's) this becomes

$$S_{n_1n_2,n_3}(j|i_1,i_2) = \sum_{k_1,k_2} S_{n_1,n_2,n_3}(j|k_1,k_2,i_2)\beta(k_1,k_2|i_1) = \sum_{k_1,k_2,k_3} S_{n_1,n_2,n_3}(j|k_1,k_2,k_3)\beta(k_1,k_2|i_1)\delta_{k_3,i_2}$$

so that we get the equality $S_{n_1n_2,n_3} = S_{n_1,n_2,n_3}(\beta \otimes I_{n_3}) = S_{n_1,n_2,n_3}(S_{n_1,n_2}^{-1}T_{n_1n_2} \otimes I_{n_3})$.

Filling this in the left-hand side of the coherence equation:

$$S_{n_1n_2,n_3}(T_{n_1n_2}^{-1}S_{n_1,n_2} \otimes I_{n_3}) = S_{n_1,n_2,n_3}(S_{n_1,n_2}^{-1}T_{n_1n_2} \otimes I_{n_3})(T_{n_1n_2}^{-1}S_{n_1,n_2} \otimes I_{n_3}) = S_{n_1,n_2,n_3}.$$

By doing a similar rewriting exercise for S_{n_1,n_2n_3} we get the same expression on the right-hand side, which proves the coherence equation. $\qquad\square$

C Proof of Theorem 19

We first prove a short lemma about subspaces of matrices.

Lemma. Let $L \subset M_n$ be a subspace containing the identity and at least one Hermitian matrix with at least two distinct eigenvalues, then $V = \mathrm{span}\left(\bigcup_{U \in U(n)} ULU^\dagger\right) = M_n$, where the union is taken over all unitary matrices.

Proof. We need to show that V contains all matrices. Since it contains all unitary conjugations of any matrix, it suffices to show that it contains all diagonal matrices. Let E be the matrix in L with at least two distinct eigenvalues. We can diagonalise $E = UD_1U^\dagger$, so D_1 is in V. Write $D_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ taking $\lambda_1 \neq \lambda_2$. Since I_n is in V we also have $D_2 = 1/(\lambda_2 - \lambda_1)(D_1 - \lambda_1 I_n)$ in V. We see that $D_2 = \mathrm{diag}(0, 1, \lambda_3', \ldots, \lambda_n')$. Now we can apply the unitary conjugation to D_2 that interchanges the first and second coordinate and subtract it from D_1 giving $D_3 = D_2 - PD_2P = \mathrm{diag}(-1, 1, 0, \ldots, 0)$. By considering conjugation with permutation matrices we can get the -1 and 1 at arbitrary spots along the diagonal. The linear span of these operators is the set of diagonal matrices with zero trace. Because we also have the identity we can then create arbitrary diagonal matrices. $\qquad\square$

Theorem. Let $F : \mathbf{CPTP} \to \mathbf{QStoch}$ be a quasi-stochastic representation of $\mathbf{CPTP}$. This representation is either faithful or trivial.

Proof. Let $\{E_i^{(n)}\}$ denote the nth quasi-POVM associated to F. Let $V = \mathrm{span}\{E_i^{(n)}\}$ and let $V^\perp = \{A \in M_n; \langle A, E_i^{(n)}\rangle_{HS} = 0 \ \forall i\}$. We note that $M_n = V \oplus V^\perp$. F is faithful on n if and only if $V^\perp = \{0\}$, since otherwise we can find a state $\rho = \rho_1 + \rho_2$ where $\rho_1 \in V$, $\rho_2 \in V^\perp$, $\rho_2 \neq 0$ such that $F(\rho) = F(\rho_1)$.

Because of functoriality we must have $F(\Phi(\rho)) = F(\Phi)F(\rho) = F(\Phi)F(\rho_1) = F(\Phi(\rho_1))$ for all CPTP maps Φ. If we can find a Φ such that $\Phi(\rho_2) \notin V^\perp$ then this leads to a contradiction because we would have $F(\Phi(\rho)) \neq F(\Phi(\rho_1))$. Therefore for the POVM to fit in a valid representation we must have $V^\perp$ closed under application of an arbitrary CPTP map. In particular, it has to be closed under unitary conjugation: when $\mathrm{tr}\left(\rho_2 E_i^{(n)}\right) = 0$ for all i we must also have $\mathrm{tr}\left(U\rho_2 U^\dagger E_i^{(n)}\right) = \mathrm{tr}\left(\rho_2 U^\dagger E_i^{(n)} U\right) = 0$ for all unitaries U. This means we should have $\mathrm{tr}(\rho_2 A) = 0$ for all $A \in W = \mathrm{span}\left(\bigcup_{U \in U(n)} UVU^\dagger\right)$.

Now we distinguish two cases. Either all the $E_i^{(n)}$ are multiples of the identity in which case $F(\rho)_i = \mathrm{tr}\left(\rho E_i^{(n)}\right) = \mathrm{tr}(\rho \alpha_i I_n) = \alpha_i$ where $E_i^{(n)} = \alpha_i I_n$, so that the representation is trivial, or there is a $E_i^{(n)}$ that isn't a multiple of the identity in which case it has at least two distinct eigenvalues. In the latter case the space W satisfies the conditions for the previous lemma which gives $W = M_n$. But then $\mathrm{tr}(\rho_2 A) = 0$ for all $A \in M_n$ so that necessarily $\rho_2 = 0$, which shows that $V^\perp = \{0\}$. The representation is then indeed faithful for n.

Let us now suppose that F isn't faithful. Then there is an n such that two states in M_n are mapped to the same distribution. We then know that all states are mapped to the same distribution for this n

by the argument above. Let $m \geq n$. Pick two orthogonal pure states $|v\rangle \langle v|, |w\rangle \langle w| \in M_m$ and let ρ and σ be orthogonal pure states in M_n, then there exists a partial isometry Φ such that $\Phi(\rho) = |v\rangle \langle v|$ and $\Phi(\sigma) = |w\rangle \langle w|$. By functoriality we get $F(|v\rangle \langle v|) = F(\Phi(\rho)) = F(\Phi)F(\rho) = F(\Phi)F(\sigma) = F(\Phi(\sigma)) = F(|w\rangle \langle w|)$. Since v and w were arbitrary, all pure states must be mapped to the same distribution and by convexity this holds for all states. This means that the representation is trivial for all $m \geq n$. If we now consider a CPTP surjection $\Phi : M_n \rightarrow M_2$ we can use the same argument to show that the representation is trivial for $m = 2$ which shows that indeed the entire representation is trivial. $\qquad \square$