entropy

MDPI

Article

Ruling out Higher-Order Interference from Purity Principles

Howard Barnum [1,2,*], Ciarán M. Lee [3,*], Carlo Maria Scandolo [4,*] and John H. Selby [4,5,*]

[1]   Centre for the Mathematics of Quantum Theory (QMATH), Department of Mathematical Sciences, University of Copenhagen, DK-2100 Copenhagen, Denmark
[2]   Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87131, USA
[3]   Department of Physics, University College London, London WC1E 6BT, UK
[4]   Department of Computer Science, University of Oxford, Oxford OX1 3QD, UK
[5]   Department of Physics, Imperial College London, London SW7 2AZ, UK
[*]   Correspondence: hnbarnum@aol.com (H.B.); ciaran.lee@ucl.ac.uk (C.M.L.); carlomaria.scandolo@cs.ox.ac.uk (C.M.S.); john.selby08@imperial.ac.uk (J.H.S.)

Abstract: As first noted by Rafael Sorkin, there is a limit to quantum interference. The interference pattern formed in a multi-slit experiment is a function of the interference patterns formed between pairs of slits; there are no genuinely new features resulting from considering three slits instead of two. Sorkin has introduced a hierarchy of mathematically conceivable *higher-order* interference behaviours, where classical theory lies at the first level of this hierarchy and quantum theory theory at the second. Informally, the order in this hierarchy corresponds to the number of slits on which the interference pattern has an irreducible dependence. Many authors have wondered why quantum interference is limited to the second level of this hierarchy. Does the existence of higher-order interference violate some natural physical principle that we believe should be fundamental? In the current work we show that such principles can be found which limit interference behaviour to second-order, or "quantum-like", interference, but that do not restrict us to the entire quantum formalism. We work within the operational framework of generalised probabilistic theories, and prove that any theory satisfying Causality, Purity Preservation, Pure Sharpness, and Purification—four principles that formalise the fundamental character of purity in nature—exhibits at most second-order interference. Hence these theories are, at least conceptually, very "close" to quantum theory. Along the way we show that systems in such theories correspond to Euclidean Jordan algebras. Hence, they are self-dual and, moreover, multi-slit experiments in such theories are described by pure projectors.

Keywords: higher-order interference; generalised probabilistic theories; Euclidean Jordan algebras

1. Introduction

Described by Feynman as "impossible, *absolutely* impossible, to explain in any classical way" [1] (volume 1, chapter 37), quantum interference is a distinctive signature of non-classicality. However, as first noted by Rafael Sorkin [2,3], there is a limit to this interference; in contrast to the case of two slits, the interference pattern formed in a three slit experiment *can* be written as a linear combination of two and one slit patterns. Sorkin has introduced a hierarchy of mathematically conceivable *higher-order* interference behaviours, where classical theory lies at the first level of this hierarchy and quantum theory theory at the second. Informally, the order in this hierarchy corresponds to the number of slits on which the interference pattern has an irreducible dependence.

Many authors have wondered why quantum interference is limited to the second level of this hierarchy [2,4–13]. Does the existence of higher-order interference violate some natural physical

principle that we believe should be fundamental [14]? In the current work we show that such natural principles can be found which limit interference behaviour to second-order, or "quantum-like", interference, but that do not restrict us to the entire quantum formalism.

We work in the framework of general probabilistic theories [15–28]. This framework is general enough to accommodate essentially arbitrary operational theories, where an operational theory specifies a set of laboratory devices which can be connected together in different ways, and assigns probabilities to different experimental outcomes. Investigating how the structural and information-theoretic features of a given theory in this framework depend on different physical principles deepens our physical and intuitive understanding of such features. Indeed, many authors [20,22,23,28,29] have derived the entire structure of finite-dimensional quantum theory from simple information-theoretic axioms—reminiscent of Einstein's derivation of special relativity from two simple physical principles. So far, ruling out higher-order interference has required thermodynamic arguments. Indeed, by combining the results and axioms of Refs. [30,31], higher-order interference could be ruled out in theories satisfying the combined axioms. In this paper we show that we can prove this in a more direct way from first principles, using only the axioms of Ref. [30].

Many experimental investigations have searched for divergences from quantum theory by looking for higher-order interference [32–36]. These experiments involved passing a particle through a physical barrier with multiple slits and comparing the interference patterns formed on a screen behind the barrier when different subsets of slits are closed. Given this set-up, one would expect that the physical theory being tested should possess transformations that correspond to the action of blocking certain subsets of slits. Moreover, blocking all but two subsets of slits should not affect states which can pass through either slit. This intuition suggests that these transformations should correspond to projectors.

Many operational probabilistic theories do not possess such a natural mathematical interpretation of multi-slit experiments; indeed many theories do not admit well-defined projectors [9]. Here, we show that there exist natural information-theoretic principles that both imply the existence of the projector structure, and rule out third-, and higher-, order interference. The principles that ensure this structure are Causality, Purity Preservation, Pure Sharpness, and Purification. These formalise intuitive ideas about the fundamental role of purity in nature. More formally, we show that such theories possess a self-dualising inner product, and that there exist pure projectors which represent the opening and closing of slits in a multi-slit experiment. Barnum, Müller and Ududec have shown that in any self-dual theory in which such projectors exist for every face, if projectors map pure states to pure states, then there can be at most second-order interference [4] (Proposition 29). The conjunction of our new results and the principle of Purity Preservation implies the conditions of Barnum et al.'s proposition. Hence sharp theories with purification do not exhibit higher-order interference. In fact we prove a stronger result, that the systems in such theories are Euclidean Jordan algebras which have been studied in quantum foundations [4,13,37].

This paper is organised as follows. In Section 2 we review the basics of the operational probabilistic theory framework. In Section 3 we formally define higher-order interference. In Section 4 we define sharp theories with purification and review relevant known results. In Section 5 we present and prove our new results. Finally, in Section 6, we offer some suggestions on how new experiments might be devised to observe higher-order interference.
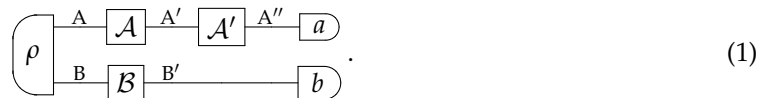
2. Framework

We will describe theories in the framework of operational-probabilistic theories (OPTs) [19,20,24,29,38–40], arising from the marriage of category theory [41–46] with probabilities. The foundation of this framework is the idea that any successful physical theory must provide an account of experimental data. Hence, such theories should have an operational description in terms of such experiments.

The OPT framework is based on the graphical language of circuits, describing experiments that can be performed in a laboratory with physical systems connecting together physical processes, which are denoted as wires and boxes respectively. The systems/wires are labelled with a *type* denoted A,

B, C, ... For example, the type given to a quantum system is the dimension of the Hilbert space describing the system. The processes/boxes are then viewed as transformations with some input and output systems/wires. For instance, in quantum theory these correspond to quantum instruments. We now give a brief introduction to the important concepts in this formalism.

2.1. States, Transformations, and Effects

A fundamental tenant of the OPT framework is composition of systems and physical processes. Given two systems A and B, they can be combined into a composite system, denoted by $A \otimes B$. Physical processes can be composed to build circuits, such as

$$
\begin{array}{c}
\rho \;\rule{0pt}{0pt}\; \overset{A}{\boxed{\mathcal{A}}} \overset{A'}{\boxed{\mathcal{A}'}} \overset{A''}{\boxed{a}} \\
\overset{B}{\boxed{\mathcal{B}}} \overset{B'}{\phantom{\boxed{\mathcal{B}}}} \boxed{b}
\end{array}
\tag{1}
$$

Processes with no inputs (such as ρ in the above diagram) are called *states*, those with no outputs (such as a and b) are called *effects* and, those with both inputs and outputs (such as $\mathcal{A}, \mathcal{A}', \mathcal{B}$) are called transformations. We define:

1. $\mathsf{St}(A)$ as the set of states of system A,
2. $\mathsf{Eff}(A)$ as the set of effects on A,
3. $\mathsf{Transf}(A, B)$ as the set of transformations from A to B, and $\mathsf{Transf}(A)$ as the set of transformations from A to A,
4. $\mathcal{B} \circ \mathcal{A}$ (or $\mathcal{B}\mathcal{A}$, for short) as the sequential composition of two transformations $\mathcal{A}$ and $\mathcal{B}$, with the input of $\mathcal{B}$ matching the output of $\mathcal{A}$,
5. $\mathcal{A} \otimes \mathcal{B}$ as the parallel composition (or tensor product) of the transformations $\mathcal{A}$ and $\mathcal{B}$.

OPTs include a particular system, the trivial system I, representing the lack of input or output for a particular device.

Hence, states (resp. effects) are transformations with the trivial system as input (resp. output). Circuits with no external wires, like the circuit in Equation (1), are called scalars and are associated with probabilities. We will often use the notation $(a|\rho)$ to denote the circuit

$$
(a|\rho) \;:=\; \left(\!\rho\!\right) \overset{A}{\boxed{}} \boxed{a},
$$

and of the notation $(a|\mathcal{C}|\rho)$ to denote the circuit

$$
(a|\mathcal{C}|\rho) \;:=\; \left(\!\rho\!\right) \overset{A}{\boxed{\mathcal{C}}} \overset{B}{\boxed{a}}.
$$

The fact that scalars are probabilities and so are real numbers induces a notion of a sum of transformations, so that the sets $\mathsf{St}(A)$, $\mathsf{Transf}(A, B)$, and $\mathsf{Eff}(A)$ become spanning sets of real vector spaces, denoted by $\mathsf{St}_{\mathbb{R}}(A)$, $\mathsf{Transf}_{\mathbb{R}}(A, B)$, and $\mathsf{Eff}_{\mathbb{R}}(A)$. In this work we will restrict our attention to *finite* systems, i.e., systems for which the vector space spanned by states is finite-dimensional for all systems. Operationally this assumption means that one need not perform an infinite number of distinct experiments to fully characterise a state. Restricting ourselves to non-negative real numbers, we have the convex cone of states and of effects, denoted by $\mathsf{St}_+(A)$ and $\mathsf{Eff}_+(A)$ respectively. We moreover make the assumption that the set of states is close. Operationally this is justified by the fact that up to any experimental error a state space is indistinguishable from its closure.

The composition of states and effects leads naturally to a norm. This is defined, for states ρ as $\|\rho\| := \sup_{a \in \mathsf{Eff}(A)} (a|\rho)$, and similarly for effects a as $\|a\| := \sup_{\rho \in \mathsf{St}(A)} (a|\rho)$. The set of normalised states (resp. effects) of system A is denoted by $\mathsf{St}_1(A)$ (resp. $\mathsf{Eff}_1(A)$).

Transformations are characterised by their action on states of composite systems: if $\mathcal{A}, \mathcal{A}' \in \mathsf{Transf}(A, B)$, we have that $\mathcal{A} = \mathcal{A}'$ if and only if

$$
\left(\!\rho\!\begin{array}{c}\overset{\text{A}}{\underset{\text{S}}{\boxed{\mathcal{A}}}}\overset{\text{B}}{}\end{array}\right) = \left(\!\rho\!\begin{array}{c}\overset{\text{A}}{\underset{\text{S}}{\boxed{\mathcal{A}'}}}\overset{\text{B}}{}\end{array}\right),
\tag{2}
$$

for every system S and every state $\rho \in \mathsf{St}\,(\mathrm{A} \otimes \mathrm{S})$. However it follows that [19] effects (resp. states) are completely defined by their action on states (resp. effects) of a single system.

Equality on states of the single system A is, in general, not enough to discriminate between $\mathcal{A}$ and $\mathcal{A}'$, as is the case for quantum theory over real Hilbert spaces [47]. However, for the scope of the present article, which focuses on single-system properties, we often concern ourselves with equality on single system.

Definition 1. *Two transformations* $\mathcal{A}, \mathcal{A}' \in \mathsf{Transf}\,(\mathrm{A}, \mathrm{B})$ *are* equal on single system, *denoted by* $\mathcal{A} \doteq \mathcal{A}'$, *if* $\mathcal{A}\rho = \mathcal{A}'\rho$ *for all states* $\rho \in \mathsf{St}\,(\mathrm{A})$.

2.2. Tests and Channels

In general, the boxes corresponding to physical processes come equipped with classical pointers. When used in an experiment, the final position of the a given pointer indicates the particular process which occurred for that box in that run. In general, this procedure can be non-deterministic. These non-deterministic processes are described by *tests* [19,39]: a test from A to B is a collection of transformations $\{\mathcal{C}_i\}_{i \in \mathsf{X}}$ from A to B, where X is the set of outcomes. If A (resp. B) is the trivial system, the test is called a *preparation-test* (resp. *observation-test*). If the set of outcomes X has a single element, we say that the test is *deterministic*, because only one transformation can occur. Deterministic transformations will be called *channels*.

A channel $\mathcal{U}$ from A to B is *reversible* if there exists another channel $\mathcal{U}^{-1}$ from B to A such that $\mathcal{U}^{-1}\mathcal{U} = \mathcal{I}_{\mathrm{A}}$ and $\mathcal{U}\mathcal{U}^{-1} = \mathcal{I}_{\mathrm{B}}$, where $\mathcal{I}_{\mathrm{S}}$ is the identity transformation on system S. If there exists a reversible channel transforming A into B, we say that A and B are *operationally equivalent*, denoted as $\mathrm{A} \simeq \mathrm{B}$. The composition of systems is required to be *symmetric*, meaning that $\mathrm{A} \otimes \mathrm{B} \simeq \mathrm{B} \otimes \mathrm{A}$. Physically, this means that for every pair of systems there exists a reversible channel swapping them. A state χ is called *invariant* if $\mathcal{U}\chi = \chi$ for all reversible channels $\mathcal{U}$.

A particularly useful class of observation-tests allows for the following.

Definition 2. *The states* $\{\rho_i\}_{i \in \mathsf{X}}$ *are called* perfectly distinguishable *if there exists an observation-test* $\{a_i\}_{i \in \mathsf{X}}$ *such that* $(a_i | \rho_j) = \delta_{ij}$ *for all* $i, j \in \mathsf{X}$.

Moreover, if there is no *other state* ρ_0 *such that the states* $\{\rho_i\}_{i \in \mathsf{X}} \cup \{\rho_0\}$ *are perfectly distinguishable, the set* $\{\rho_i\}_{i \in \mathsf{X}}$ *is said* maximal.

2.3. Pure Transformations

There are various different ways to define pure transformations, for example in terms of resources [30,48–51] or "side information" [39,52]. Informally pure transformations correspond to an experimenter having maximal control of or information about a process. Here, we formalise this notion by defining the notion of a *coarse-graining* [19]. Coarse-graining is the operation of joining two or more outcomes of a test into a single outcome. More precisely, a test $\{\mathcal{C}_i\}_{i \in \mathsf{X}}$ is a *coarse-graining* of the test $\{\mathcal{D}_j\}_{j \in \mathsf{Y}}$ if there is a partition $\{\mathsf{Y}_i\}_{i \in \mathsf{X}}$ of Y such that, for all $i \in \mathsf{X}$

$$
\mathcal{C}_i = \sum_{j \in \mathsf{Y}_i} \mathcal{D}_j
$$

In this case, we say that the test $\{\mathcal{D}_j\}_{j \in \mathsf{Y}}$ is a *refinement* of the test $\{\mathcal{C}_i\}_{i \in \mathsf{X}}$, and that the transformations $\{\mathcal{D}_j\}_{j \in \mathsf{Y}_i}$ are a refinement of the transformation $\mathcal{C}_i$. A transformation $\mathcal{C} \in \mathsf{Transf}\,(\mathrm{A}, \mathrm{B})$ is *pure* if it has only trivial refinements, namely refinements $\{\mathcal{D}_j\}$ of the form $\mathcal{D}_j = p_j \mathcal{C}$, where $\{p_j\}$ is a probability distribution. We denote the sets of pure transformations, pure states, and pure effects as

PurTransf (A, B), PurSt (A), and PurEff (A) respectively. Similarly, PurSt$_1$ (A), and PurEff$_1$ (A) denote *normalised* pure states and effects respectively. Non-pure states are called *mixed*.

Definition 3. *Let $\rho \in$ St$_1$ (A). A normalised state σ is* contained *in ρ if we can write $\rho = p\sigma + (1 - p)\tau$, where $p \in (0, 1]$ and τ is another normalised state.*

Clearly, no states are contained in a pure state. On the other edge of the spectrum we have complete states.

Definition 4. *A state $\omega \in$ St$_1$ (A) is* complete *if every state is contained in it.*

Definition 5. *We say that two transformations $\mathcal{A}, \mathcal{A}' \in$ Transf (A, B) are* equal upon input *of the state $\rho \in$ St$_1$ (A) if $\mathcal{A}\sigma = \mathcal{A}'\sigma$ for every state σ contained in ρ. In this case we will write $\mathcal{A} =_\rho \mathcal{A}'$.*

2.4. Causality

A natural requirement of a physical theory is that it is *causal*, that is, no signals can be sent from the future to the past. In the OPT framework this is formalised as follows:

Axiom 1 (Causality [19,39]). *The probability that a transformation occurs is independent of the choice of tests performed on its output.*

Causality is equivalent to the requirement that, for every system A, there exists a unique deterministic effect u_A on A (or simply u, when no ambiguity can arise) [19]. Owing to the uniqueness of the deterministic effect, the marginals of a bipartite state can be uniquely defined as:

$$\boxed{\rho_A}\!\!-\!\!\stackrel{A}{-} \;:=\; \boxed{\rho_{AB}}\!\!\begin{array}{l} -\!\!\stackrel{A}{-} \\ -\!\!\stackrel{B}{-}\boxed{u} \end{array} ,$$

Moreover, this uniqueness forbids the ability to signal [19,53]. We will denote by Tr$_B\rho_{AB}$ the marginal on system A, in analogy with the notation used in the quantum case. We will stick to the notation Tr in formulas where the deterministic effect is applied directly to a state, e.g., Tr $\rho := (u|\rho)$.

In a causal theory it is easy to see that the norm of a state takes the form $\|\rho\| = \text{Tr}\,\rho$, and that a state can be prepared deterministically if and only if it is normalised.

3. Higher-Order Interference

The definition of higher-order interference we shall present in this section takes its motivation from the set-up of multi-slit interference experiments. In such experiments a particle passes through slits in a physical barrier and is detected at a screen. By repeating the experiment many times, one builds up a pattern on the screen. To determine if this experiment exhibits interference one compares this pattern to those produced when certain subsets of the slits are blocked. In quantum theory, for example, the two-slit experiment exhibits interference as the pattern formed with both slits open is not equal to the sum of the one-slit patterns.

Consider the state of the particle just before it passes through the slits. For every slit, there should exist states such that the particle is definitely found at that slit, if measured. Mathematically, this means that there is a face [4] of the state space, such that all states in this face give unit probability for the "yes" outcome of the two-outcome measurement "is the particle at this slit?". Recall that a face is a convex set with the property that if $px + (1 - p)y$, for $0 \le p \le 1$, is an element then x and y are also elements. These faces will be labelled F_i, one for each of the n slits $i \in \{1, \dots, n\}$. As the slits should be perfectly distinguishable, the faces associated with each slit should be perfectly distinguishable, or orthogonal. One can additionally ask coarse-grained questions of the form "Is the particle found among a certain subset of slits, rather than somewhere else?". The set of states that give outcome "yes" with probability one must contain all the faces associated with each slit in the subset. Hence the face

associated with the subset of slits $I \subseteq \{1, \ldots, n\}$ is the smallest face containing each face in this subset $F_I := \bigvee_{i \in I} F_i$, where the operation $\bigvee$ is the least upper bound of the lattice of faces where the ordering is provided by subset inclusion of one face within another. The face F_I contains all those states which can be found among the slits contained in I. The experiment is "complete" if all states in the state space (of a given system A) can be found among some subset of slits. That is, if $F_{12 \cdots n} = \mathsf{St}(A)$.

An *n*-slit experiment requires a system that has *n* orthogonal faces F_i, with $i \in \{1, \ldots, n\}$. Consider an effect *E* associated with finding a particle at a particular point on the screen. We now formally define an *n*-slit experiment.

Definition 6. *An n-slit experiment is a collection of effects e_I, where $I \subseteq \{1, \ldots, n\}$, such that*

$$(e_I|\rho) = (E|\rho), \quad \forall \rho \in F_I, \text{ and}$$
$$(e_I|\rho) = 0, \quad \forall \rho \text{ where } \rho \perp F_I.$$

The effects introduced in the above definition arise from the conjunction of blocking off the slits $\{1, \ldots, n\} \setminus I$ and applying the effect *E*. If the particle was prepared in a state such that it would be unaffected by the blocking of the slits (i.e., $\rho \in F_I$) then we should have $(e_I|\rho) = (E|\rho)$. If instead the particle is prepared in a state which is guaranteed to be blocked (i.e., $\rho' \perp F_I$) then the particle should have no probability of being detected at the screen, i.e., $(e_I|\rho') = 0$.

The relevant quantities for the existence of various orders of interference are [2,9,13,15]:

$$I_1 := (E|\rho), \tag{3}$$
$$I_2 := (E|\rho) - (e_1|\rho) - (e_2|\rho), \tag{4}$$
$$I_3 := (E|\rho) - (e_{12}|\rho) - (e_{23}|\rho) - (e_{31}|\rho) + (e_1|\rho) + (e_2|\rho) + (e_3|\rho), \tag{5}$$
$$I_n := \sum_{\varnothing \neq I \subseteq \{1, \ldots, n\}} (-1)^{n-|I|} (e_I|\rho), \tag{6}$$

for some state ρ, and defining $e_{\{1, \ldots, n\}} := E$.

Definition 7. *A theory has n-th order interference if there exists a state ρ and an effect E such that $I_n \neq 0$.*

In a slightly different formal setting, it was shown in [2] that $I_n = 0 \implies I_{n+1} = 0$, so if there is no *n*th order interference, there will be no $(n+1)$th order interference; the argument of [2] applies here.

It should be noted that there appears to be a lot of freedom in choosing a set of effects $\{e_I\}$ to test for the existence of higher-order interference. Indeed, in arbitrary generalised theories this appears to be the case [9]. However, it is natural to ask whether there exists physical transformations T_I in the theory which correspond to leaving the subset of slits I open and blocking the rest. Hence a unique e_I is assigned to each fixed *E* defined as $e_I = ET_I$. Ruling out the existence of higher-order interference then reduces to proving certain properties of the T_I. This will turn out to be the case in sharp theories with purification.

4. Sharp Theories with Purification

In this section we present the definition and important properties of sharp theories with purification. They were originally introduced in [30,49,54] for the analysis of the foundations of thermodynamics and statistical mechanics.

Sharp theories with purification are causal theories defined by three axioms. The first axiom—Purity Preservation—states that no information can leak when two pure transformations are composed:

Axiom 2 (Purity Preservation [55]). *Sequential and parallel compositions of pure transformations yield pure transformations.*

The second axiom—Pure Sharpness—guarantees that every system possesses at least one elementary property.

Axiom 3 (Pure Sharpness [54])**.** *For every system there exists at least one pure effect occurring with unit probability on some state.*

These axioms are satisfied by both classical and quantum theory. Our third axiom—Purification—signals the departure from classicality, and characterises when a physical theory admits a level of description where all deterministic processes are pure and reversible.

Given a normalised state $\rho_A \in \mathsf{St}_1(A)$, a normalised pure state $\Psi \in \mathsf{PurSt}_1(A \otimes B)$ is a *purification* of ρ_A if

$$\begin{array}{l} \Psi \begin{array}{l} \underline{\quad A \quad} \\ \underline{\quad B \quad} \; u \end{array} \end{array} = \begin{array}{l} \rho_A \underline{\quad A \quad} \end{array} ;$$

in this case B is called the *purifying system*. We say that a pure state $\Psi \in \mathsf{PurSt}(A \otimes B)$ is an *essentially unique purification of its marginal* ρ_A [39] if every other pure state $\Psi' \in \mathsf{PurSt}(A \otimes B)$ satisfying the purification condition must be of the form

$$\begin{array}{l} \Psi' \begin{array}{l} \underline{\quad A \quad} \\ \underline{\quad B \quad} \end{array} \end{array} = \begin{array}{l} \Psi \begin{array}{l} \underline{\quad\quad A \quad\quad} \\ \underline{\quad B \quad} \; \mathcal{U} \; \underline{\quad B \quad} \end{array} \end{array} ,$$

for some reversible channel $\mathcal{U}$.

Axiom 4 (Purification [19,39])**.** *Every state has a purification. Purifications are essentially unique.*

Quantum theory, both on complex and real Hilbert spaces, satisfies Purification, and also Spekkens' toy model [56]. Examples of sharp theories with purification besides quantum theory include fermionic quantum theory [57,58], a superselected version of quantum theory known as doubled quantum theory [49], and a recent extension of classical theory with the theory of codits [30].

Properties of Sharp Theories With Purifications

Sharp theories with purifications enjoy some nice properties, which were mainly derived in Refs. [30,54]. The first property is that every non-trivial system admits perfectly distinguishable states [54], and that all maximal sets of pure states have the same cardinality [30].

Proposition 1. *For every system* A *there is a positive integer* d_A, *called the* dimension *of* A, *such that all maximal sets of pure states have* d_A *elements.*

Note that we will omit the subscript A when the context is clear.

In sharp theories with purification every state can be diagonalised, i.e., written as a convex combination of perfectly distinguishable pure states (cf. Refs. [30,54]).

Theorem 5. *Every normalised state* $\rho \in \mathsf{St}_1(A)$ *of a non-trivial system can be decomposed as*

$$\rho = \sum_{i=1}^{d} p_i \alpha_i,$$

where $\{p_i\}_{i=1}^{d}$ *is a probability distribution, and* $\{\alpha_i\}_{i=1}^{d}$ *is a* pure *maximal set. Moreover, given* ρ, $\{p_i\}_{i=1}^{d}$ *is unique up to rearrangements.*

Such a decomposition is called a *diagonalisation* of ρ, the p_i's are the *eigenvalues* of ρ, and the α_i's are the *eigenstates*. Theorem 5 implies that the eigenvalues of a state are unique, and independent of its diagonalisation. Sharp theories with purification have a unique invariant state χ [19], which can be diagonalised as $\chi = \frac{1}{d} \sum_{i=1}^{d} \alpha_i$, where $\{\alpha_i\}_{i=1}^{d}$ is *any* pure maximal set [30]. Furthermore, the

diagonalisation result of Theorem 5 can be extended to every vector in $\mathsf{St}_\mathbb{R}(A)$, but here the eigenvalues will be generally real numbers [30].

One of the most important consequences for this paper of the axioms defining sharp theories with purification is a duality between normalised pure states and normalised pure effects.

Theorem 6 (States-effects duality [30,54]). *For every system* A, *there is a bijective correspondence* †: $\mathsf{PurSt}_1(A) \to \mathsf{PurEff}_1(A)$ *such that if* $\alpha \in \mathsf{PurSt}_1(A)$, $\alpha^\dagger$ *is the* unique *normalised pure effect such that* $\left(\alpha^\dagger\middle|\alpha\right) = 1$. *Furthermore this bijection can be extended by linearity to an isomorphism between the vector spaces* $\mathsf{St}_\mathbb{R}(A)$ *and* $\mathsf{Eff}_\mathbb{R}(A)$.

With a little abuse of notation we will use † also to denote the inverse map $\mathsf{PurEff}_1(A) \to \mathsf{PurSt}_1(A)$, by which, if $a \in \mathsf{PurEff}_1(A)$, $a^\dagger$ is the unique pure state such that $\left(a\middle|a^\dagger\right) = 1$. Pure maximal sets $\{\alpha_i\}_{i=1}^d$ have the property that $\sum_{i=1}^d \alpha_i^\dagger = u$ [30].

A diagonalisation result holds for vectors of $\mathsf{Eff}_\mathbb{R}(A)$ as well [30]: they can be written as $X = \sum_{i=1}^d \lambda_i \alpha_i^\dagger$, where $\{\alpha_i\}_{i=1}^d$ is a pure maximal set. Again, the λ_i's are uniquely defined given X.

Another result that will be made use of in the following sections is the following. It was shown to hold in Ref. [30], and expresses the possibility of constructing non-disturbing measurements [20,59,60].

Proposition 2. *Given a system* A, *let* $a \in \mathsf{Eff}(A)$ *be an effect such that* $(a|\rho) = 1$, *for some* $\rho \in \mathsf{St}_1(A)$. *Then there exists a pure transformation* $\mathcal{T} \in \mathsf{PurTransf}(A)$ *such that* $\mathcal{T} =_\rho \mathcal{I}$, *with* $(u|\mathcal{T}|\sigma) \leq (a|\sigma)$, *for every state* $\sigma \in \mathsf{St}_1(A)$.

Note that the pure transformation $\mathcal{T}$ is non-disturbing on ρ because it acts as the identity on ρ and on all states contained in it. In other words, whenever we have an effect occurring with unit probability on some state ρ, we can always find a transformation that does not disturb ρ (i.e., a non-disturbing, non-demolition measurement) [30].

Finally, a property that we will use often is a sort of no-restriction hypothesis for tests, derived in [20] (Corollary 4).

Proposition 3. *A collection of transformations* $\{\mathcal{A}_i\}_{i\in\mathsf{X}}$ *is a valid test if and only if* $\sum_{i\in\mathsf{X}} u\mathcal{A}_i = u$.
A collection of effects $\{a_i\}_{i\in\mathsf{X}}$ *is a valid observation-test if and only if* $\sum_{i\in\mathsf{X}} a_i = u$.

5. Sharp Theories with Purification Have No Higher-Order Interference

Here we will show that sharp theories with purification do not exhibit higher-order interference. Our proof strategy will be to show that results of [4], which rule out the existence of higher-order interference from certain assumptions, hold in sharp theories with purification. To this end, we will first prove that these theories are self-dual, and that they admit *pure* orthogonal projectors which satisfy certain properties, compatible with the setting presented in Section 3.

5.1. Self-Duality

Now we will prove that sharp theories with purification are self-dual. Recall that a theory is *self-dual* if for every system A there is an inner product $\langle \bullet, \bullet \rangle$ on $\mathsf{St}_\mathbb{R}(A)$ such that $\xi \in \mathsf{St}_+(A)$ if and only if $\langle \xi, \eta \rangle \geq 0$ for every $\eta \in \mathsf{St}_+(A)$. To show that, we need to find a self-dualising inner product on $\mathsf{St}_\mathbb{R}(A)$ for every system A. The dagger will provide us with a good candidate. First we need the following lemma.

Lemma 1. *Let* $a \in \mathsf{Eff}_1(A)$ *be a normalised effect. Then* a *is of the form* $a = \sum_{i=1}^r \alpha_i^\dagger$, *with* $r \leq d$, *and the pure states* $\{\alpha_i\}_{i=1}^r$ *are perfectly distinguishable.*

Proof. We know that every effect a can be written as $a = \sum_{i=1}^{r} \lambda_i \alpha_i^\dagger$, where $r \leq d$, the pure states $\{\alpha_i\}_{i=1}^{r}$ are perfectly distinguishable, and for every $i \in \{1, \ldots, r\}$, $\lambda_i \in (0, 1]$. Since the state space is closed, and a is normalised, then there exists a (normalised) state ρ such that $(a|\rho) = 1$. One has

$$1 = (a|\rho) = \sum_{i=1}^{r} \lambda_i \left(\alpha_i^\dagger \middle| \rho \right).$$

Now, $\left(\alpha_i^\dagger \middle| \rho \right) \geq 0$, and $\sum_{i=1}^{r} \left(\alpha_i^\dagger \middle| \rho \right) \leq 1$ because

$$\sum_{i=1}^{r} \left(\alpha_i^\dagger \middle| \rho \right) \leq \sum_{i=1}^{d} \left(\alpha_i^\dagger \middle| \rho \right) = \operatorname{Tr} \rho = 1,$$

where we have used the fact that $\sum_{i=1}^{d} \alpha_i^\dagger = u$. Then $\sum_{i=1}^{r} \lambda_i \left(\alpha_i^\dagger \middle| \rho \right) \leq \lambda_{\max}$, where $\lambda_{\max}$ is the maximum of the λ_i's. Therefore, $\lambda_{\max} \geq 1$, which implies $\lambda_{\max} = 1$. Now, the condition

$$\sum_{i=1}^{r} \lambda_i \left(\alpha_i^\dagger \middle| \rho \right) = \lambda_{\max}$$

means that $\lambda_i = \lambda_{\max} = 1$ for all $i \in \{1, \ldots, r\}$. $\quad\square$

In the above, we call r the *rank* of the normalised effect. We can use this result to prove the following.

Lemma 2. *For every system* A, *the map*

$$\langle \xi, \eta \rangle := \left(\xi^\dagger \middle| \eta \right),$$

for every $\xi, \eta \in \mathsf{St}_\mathbb{R}(A)$ *is an inner product on* $\mathsf{St}_\mathbb{R}(A)$.

Proof. The map $\langle \bullet, \bullet \rangle$ is clearly bilinear by construction, because the dagger is also linear. Let us show that it is positive-definite. Take a non-null vector $\xi \in \mathsf{St}_\mathbb{R}(A)$, and diagonalise it as $\xi = \sum_{i=1}^{d} x_i \alpha_i$. Then

$$\langle \xi, \xi \rangle = \left(\xi^\dagger \middle| \xi \right) = \sum_{i,j=1}^{d} x_i x_j \left(\alpha_i^\dagger \middle| \alpha_j \right) = \sum_{i=1}^{d} x_i^2 > 0,$$

where we have used the fact that for perfectly distinguishable pure states $\left(\alpha_i^\dagger \middle| \alpha_j \right) = \delta_{ij}$ [30].

The hard part is to prove that this bilinear map is symmetric, namely $\langle \xi, \eta \rangle = \langle \eta, \xi \rangle$, for every $\xi, \eta \in \mathsf{St}_\mathbb{R}(A)$. Let us define a new (double) dagger $\ddagger$. The double dagger of a normalised state ρ is an effect $\rho^\ddagger$ whose action on normalised states σ is defined as

$$\left(\rho^\ddagger \middle| \sigma \right) := \left(\sigma^\dagger \middle| \rho \right), \tag{7}$$

where $\dagger$ is the dagger of Theorem 6. Note that Equation (7) is enough to characterise $\rho^\ddagger$ completely, and it guarantees that $\rho^\ddagger$ is a mathematically well-defined effect, because it is linear and $\left(\sigma^\dagger \middle| \rho \right) \in [0, 1]$. Consider now ρ and σ to be a normalised *pure* state ψ. Then $\left(\psi^\ddagger \middle| \psi \right) = \left(\psi^\dagger \middle| \psi \right) = 1$, this means that $\alpha^\ddagger$ is normalised. If we manage to show that $\psi^\ddagger$ is pure, then by Theorem 6 we can conclude that $\psi^\ddagger = \psi^\dagger$. By Lemma 1, $\psi^\ddagger$ is of the form $\psi^\ddagger = \sum_{i=1}^{r} \alpha_i^\dagger$, with $r \leq d$, and the pure states $\{\alpha_i\}_{i=1}^{r}$ are perfectly distinguishable. Clearly $\psi^\ddagger$ is pure if and only if $r = 1$. To prove it, first let us evaluate $\psi^\ddagger$ on χ:

$$\left(\psi^\ddagger \middle| \chi \right) = \left(\chi^\dagger \middle| \psi \right) = \frac{1}{d} \operatorname{Tr} \psi = \frac{1}{d}, \tag{8}$$

as prescribed by Equation (7). Now, since $\psi^{\ddagger} = \sum_{i=1}^{r} \alpha_i^{\dagger}$, we have

$$\left(\psi^{\ddagger}\middle|\chi\right) = \sum_{i=1}^{r} \left(\alpha_i^{\dagger}\middle|\chi\right) = \frac{r}{d}, \tag{9}$$

because $\left(\alpha_i^{\dagger}\middle|\chi\right) = \frac{1}{d}$ for every i [30]. A comparison between Equations (8) and (9), shows that $r = 1$. This means that $\psi^{\ddagger}$ is pure, whence $\psi^{\ddagger} = \psi^{\dagger}$. Now we can show that the double dagger $\ddagger$ actually coincides with the dagger of Theorem 6. Indeed, given a state ρ, diagonalise it as $\rho = \sum_{i=1}^{d} p_i \alpha_i$. One can easily show that the double dagger of Equation (7) is linear, so we have $\rho^{\ddagger} = \sum_{i=1}^{d} p_i \alpha_i^{\ddagger}$, but we have just proved that $\alpha_i^{\ddagger} = \alpha_i^{\dagger}$ for pure states, so $\rho^{\ddagger} = \sum_{i=1}^{d} p_i \alpha_i^{\dagger} = \rho^{\dagger}$. This means that $\ddagger = \dagger$, and that Equation (7) is nothing but a redefinition of the usual dagger. This means for every normalised states we have

$$\left(\rho^{\dagger}\middle|\sigma\right) = \left(\sigma^{\dagger}\middle|\rho\right), \tag{10}$$

and this extends linearly to all vectors $\xi, \eta \in \mathsf{St}_{\mathbb{R}}(A)$. We have proved that $\langle \bullet, \bullet \rangle$ is symmetric, and this concludes the proof. $\square$

Note that the above result immediately yields the "symmetry of transition probabilities" as defined in Ref. [61,62].

Now we prove that this inner product is invariant under reversible transformations.

Proposition 4. *For every $\xi, \eta \in \mathsf{St}_{\mathbb{R}}(A)$ and every reversible channel $\mathcal{U}$ one has*

$$\langle \mathcal{U}\xi, \mathcal{U}\eta \rangle = \langle \xi, \eta \rangle.$$

Proof. To prove the statement, let us first prove that for a normalised pure state α one has $(\mathcal{U}\alpha)^{\dagger} = \alpha^{\dagger}\mathcal{U}^{-1}$, for every reversible channel $\mathcal{U}$. $\alpha^{\dagger}\mathcal{U}^{-1}$ is a pure effect and one has $\left(\alpha^{\dagger}\mathcal{U}^{-1}\middle|\mathcal{U}\alpha\right) = \left(\alpha^{\dagger}\middle|\alpha\right) = 1$. By the uniqueness of the dagger for normalised pure states, $\alpha^{\dagger}\mathcal{U}^{-1} = (\mathcal{U}\alpha)^{\dagger}$. This can be extended by linearity to all vectors ξ in $\mathsf{St}_{\mathbb{R}}(A)$, so $(\mathcal{U}\xi)^{\dagger} = \xi^{\dagger}\mathcal{U}^{-1}$. Therefore, when we compute $\langle \mathcal{U}\xi, \mathcal{U}\eta \rangle$, we have

$$\langle \mathcal{U}\xi, \mathcal{U}\eta \rangle = \left(\xi^{\dagger}\middle|\mathcal{U}^{-1}\mathcal{U}\middle|\eta\right) = \left(\xi^{\dagger}\middle|\eta\right) = \langle \xi, \eta \rangle.$$

$\square$

The fact that $\langle \bullet, \bullet \rangle$ is an inner product allows us to define an additional norm in sharp theories with purification: if $\xi \in \mathsf{St}_{\mathbb{R}}(A)$, define the *dagger norm* as

$$\|\xi\|_{\dagger} := \sqrt{\langle \xi, \xi \rangle}.$$

See Appendix A.1 for an extended discussion on the properties of this norm.

Now we are ready to state the core of this subsection.

Proposition 5. *Sharp theories with purification are self-dual.*

Proof. Given a system A, we need to prove that $\xi \in \mathsf{St}_{\mathbb{R}}(A)$ is in $\mathsf{St}_{+}(A)$ if and only if $\langle \xi, \eta \rangle \geq 0$ for all $\eta \in \mathsf{St}_{+}(A)$. Note that $\xi \in \mathsf{St}_{+}(A)$ if and only if it can be diagonalised as $\xi = \sum_{i=1}^{d} x_i \alpha_i$, where the x_i's are all non-negative.

Necessity. Suppose $\xi \in \mathsf{St}_{+}(A)$, and take any $\eta \in \mathsf{St}_{+}(A)$, diagonalised as $\eta = \sum_{i=1}^{d} y_i \beta_i$. Then we have

$$\langle \xi, \eta \rangle = \sum_{i,j=1}^{d} x_i y_j \left(\alpha_i^{\dagger}\middle|\beta_j\right) \geq 0$$

because all the terms x_i, y_j, and $\left(\alpha_i^{\dagger}\middle|\beta_j\right)$ are non-negative.

Sufficiency. Take $\xi \in \mathsf{St}_{\mathbb{R}}(A)$, and assume that $\langle \xi, \eta \rangle \geq 0$ for all $\eta \in \mathsf{St}_+(A)$. Assume ξ is diagonalised as $\xi = \sum_{i=1}^{d} x_i \alpha_i$, where the x_i's are generic real numbers. We wish to prove that all the x_i's are non-negative. Then

$$\langle \xi, \eta \rangle = \sum_{i,j=1}^{d} x_i \left(\alpha_i^\dagger \big| \eta \right) \geq 0.$$

Recalling that for perfectly distinguishable pure states one has $\left(\alpha_i^\dagger \big| \alpha_j \right) = \delta_{ij}$ [30], it is enough to take η to be one of the states $\{\alpha_i\}_{i=1}^{d}$ to conclude that $x_i \geq 0$ for every $i \in \{1, \ldots, d\}$, meaning that $\xi \in \mathsf{St}_+(A)$. $\square$

The self-dualising inner product, besides being a nice mathematical tool, has some operational meaning, because it provides a measure of the distinguishability of states, as explained in Appendix A.2. Moreover, it is the starting point for extending the dagger to all transformations. This is done in Appendix B.

5.2. Existence of Pure Orthogonal Projectors

Now we show that we have orthogonal projectors on every face of the state space. A consequence of diagonalisation is that all faces are generated by perfectly distinguishable *pure* states. Indeed, every face F is generated by a state ω in its relative interior. ω can be diagonalised as $\omega = \sum_{i=1}^{r} p_i \alpha_i$, where $r \leq d$, and $p_i > 0$ for $i \in \{1, \ldots, r\}$. By definition of face, this means that the states $\{\alpha_i\}_{i=1}^{r}$ are in F, and therefore generate F. Consequently, there is an effect a that picks out the whole face as the set of states ρ such that $(a|\rho) = 1$. In the specific case considered above, it is $a = \sum_{i=1}^{r} \alpha_i^\dagger$. Such faces are called *exposed*.

Therefore the study of faces of sharp theories with purification reduces to the study of normalised effects. Thanks to Lemma 1, it is enough to consider subsets of pure maximal sets. Pick a pure maximal set $\{\alpha_i\}_{i=1}^{d}$, and consider a subset I of $\{1, \ldots, d\}$. The subset I flags the slits that are open in the experiment. Setting $a_\mathsf{I} := \sum_{i \in \mathsf{I}} \alpha_i^\dagger$, we can define the two faces

1. $F_\mathsf{I} := \{\rho \in \mathsf{St}_1(A) : (a_\mathsf{I}|\rho) = 1\}$;
2. $F_\mathsf{I}^\perp := \{\rho \in \mathsf{St}_1(A) : (a_\mathsf{I}|\rho) = 0\}$,

in analogy with those of Definition 6. Clearly the effect $a_\mathsf{I}^\perp := \sum_{i \notin \mathsf{I}} \alpha_i^\dagger$ defines the orthogonal face $F_\mathsf{I}^\perp$, as it occurs with probability one on the states of $F_\mathsf{I}^\perp$. Note that each of the effects $\left\{ \alpha_i^\dagger \right\}_{i \notin \mathsf{I}}$ occurs with zero probability on the states of F_I.

Definition 8. *An* orthogonal projector *(in the sense of [20]) on the face F_I is a transformation $P_\mathsf{I} \in \mathsf{Transf}(A)$ such that*

-    *if $\rho \in F_\mathsf{I}$, then $P_\mathsf{I}\rho = \rho$;*
-    *if $\rho \in F_\mathsf{I}^\perp$, then $P_\mathsf{I}\rho = 0$.*

We can prove the existence of a projector at least in one case, when $\mathsf{I} = \{1, \ldots, d\}$. In this case $a_\mathsf{I} = u$, so $F_\mathsf{I} = \mathsf{St}_1(A)$, and $F_\mathsf{I}^\perp = \varnothing$. Then it is enough to take $P_\mathsf{I} \doteq \mathcal{I}$. However, sharp theories with purification admit projectors on *every* face.

Proposition 6. *Sharp theories with purification have* pure *projectors on every face F_I. Furthermore one has $uP_\mathsf{I} = a_\mathsf{I}$.*

Proof. Suppose ρ is any state in F_I, then $(a_\mathsf{I}|\rho) = 1$. By Proposition 2 we know that there is a *pure* transformation P_I such that $P_\mathsf{I}\rho = \rho$ for every $\rho \in F_\mathsf{I}$. We also have $(u|P_\mathsf{I}|\sigma) \leq (a_\mathsf{I}|\sigma)$, so if $\sigma \in F_\mathsf{I}^\perp$, we have $(u|P_\mathsf{I}|\sigma) = 0$, whence $P_\mathsf{I}\sigma = 0$.

To prove that $uP_\mathsf{I} = a_\mathsf{I}$, first note that $\psi^\dagger P_\mathsf{I} = \psi^\dagger$ for every pure state $\psi \in F_\mathsf{I}$. Indeed $\psi^\dagger P_\mathsf{I}$ is pure by Purity Preservation, and we have $(\psi^\dagger|P_\mathsf{I}|\psi) = (\psi^\dagger|\psi) = 1$ because $P_\mathsf{I}\psi = \psi$ by definition. By Theorem 6, we have $\psi^\dagger P_\mathsf{I} = \psi^\dagger$. Furthermore, $\varphi^\dagger P_\mathsf{I} = 0$ for a pure state $\varphi \in F_\mathsf{I}^\perp$. Indeed, consider

$$\left(\varphi^\dagger\big|P_\mathsf{I}\big|\chi\right) = \frac{1}{d}\sum_{i\in\mathsf{I}}\left(\varphi^\dagger\big|P_\mathsf{I}\big|\alpha_i\right) + \frac{1}{d}\sum_{i\notin\mathsf{I}}\left(\varphi^\dagger\big|P_\mathsf{I}\big|\alpha_i\right).$$

The second term vanishes because $\alpha_i \in F_\mathsf{I}^\perp$ for $i \notin \mathsf{I}$. The first term vanishes because $P_\mathsf{I}\alpha_i = \alpha_i$ for $i \in \mathsf{I}$, and φ is perfectly distinguishable from any of the α_i's for $i \in \mathsf{I}$ by means of the observation-test $\{u - a_\mathsf{I}, a_\mathsf{I}\}$, implying $(\varphi^\dagger|\alpha_i) = 0$ [30]. This means that $\varphi^\dagger P_\mathsf{I}$ occurs with zero probability on all states contained in χ, and since χ is complete [19], $\varphi^\dagger P_\mathsf{I} = 0$. Now, when we calculate uP_I, we separate the contribution arising from states in orthogonal faces:

$$uP_\mathsf{I} = \sum_{i\in\mathsf{I}}\alpha_i^\dagger P_\mathsf{I} + \sum_{i\notin\mathsf{I}}\alpha_i^\dagger P_\mathsf{I} = \sum_{i\in\mathsf{I}}\alpha_i^\dagger = a_\mathsf{I}$$

This concludes the proof. □

In other words, P_I occurs with the same probability as a_I, thus satisfying one of the desiderata of Section 3. Moreover, extending some of the results in the Proof 6 by linearity, we obtain the dual statements of Definition 8, namely

- $\rho^\dagger P_\mathsf{I} = \rho^\dagger$ if $\rho \in F_\mathsf{I}$
- $\rho^\dagger P_\mathsf{I} = 0$ if $\rho \in F_\mathsf{I}^\perp$

Another consequence of Proposition 6 is that projectors actually project on their associated face, viz. for every normalised state ρ, $P_\mathsf{I}\rho = \lambda\sigma$, where σ is in F_I, and $\lambda = (a_\mathsf{I}|\rho)$. Indeed, $\lambda = (u|P_\mathsf{I}|\rho) = (a_\mathsf{I}|\rho)$. If $\lambda \neq 0$, which means $\rho \notin F_\mathsf{I}^\perp$, then and $(a_\mathsf{I}|\sigma) = \frac{1}{\lambda}(a_\mathsf{I}|P_\mathsf{I}|\rho)$. However, we know that $a_\mathsf{I}P_\mathsf{I} = a_\mathsf{I}$, so $(a_\mathsf{I}|\sigma) = 1$, showing that $\sigma \in F_\mathsf{I}$.

Furthermore, we can show that every projector P_I has a *complement* $P_\mathsf{I}^\perp$, which is the projector associated with the effect $a_\mathsf{I}^\perp = \sum_{i\notin\mathsf{I}}\alpha_i^\dagger$, which defines the orthogonal face $F_\mathsf{I}^\perp$. Clearly $P_\mathsf{I}^\perp\rho = (a_\mathsf{I}^\perp|\rho)\sigma$, with $\sigma \in F_\mathsf{I}^\perp$. In particular, $P_\mathsf{I}^\perp\rho$ vanishes if and only if $\rho \in F_\mathsf{I}$.

These properties are the starting point for proving the idempotence of projectors.

Proposition 7. *Given a fixed pure maximal set $\{\alpha_i\}_{i=1}^d$ and $\mathsf{I} \subseteq \{1,\ldots,d\}$, one has $P_\mathsf{I}^2 \doteq P_\mathsf{I}$. Moreover, if J is another subset of $\{1,\ldots,d\}$ disjoint from I, then $P_\mathsf{I}P_\mathsf{J} \doteq 0$.*

Proof. Recall that for every state ρ, $P_\mathsf{I}\rho = \lambda\sigma$, where σ is in F_I. Now, P_I leaves σ invariant by definition, so

$$P_\mathsf{I}^2\rho = \lambda P_\mathsf{I}\sigma = \lambda\sigma,$$

so $P_\mathsf{I}^2 \doteq P_\mathsf{I}$. To prove the other property, note that if I and J are disjoint, they define orthogonal faces. Indeed, suppose $\rho \in F_\mathsf{I}$, then

$$1 = \mathrm{Tr}\,\rho = (a_\mathsf{I}|\rho) + (a_\mathsf{J}|\rho) + \sum_{i\notin\mathsf{I}\cup\mathsf{J}}\left(\alpha_i^\dagger\big|\rho\right),$$

which implies $(a_\mathsf{J}|\rho) = 0$ because $(a_\mathsf{I}|\rho) = 1$. Hence $\rho \in F_\mathsf{J}^\perp$. Now, given *any* normalised state ρ, $P_\mathsf{I}P_\mathsf{J}\rho = 0$ because $P_\mathsf{J}\rho$ is proportional to a state in $F_\mathsf{I}^\perp$. This proves that $P_\mathsf{I}P_\mathsf{J} \doteq 0$. □

This result shows that, once a pure maximal set $\{\alpha_i\}_{i=1}^d$ is fixed, whenever we have a partition $\{\mathsf{I}_j\}$ of $\{1,\ldots,d\}$, the test $\{P_{\mathsf{I}_j}\}$ is a von Neumann measurement. The only thing left to check is that

$\sum_j u P_{\mathsf{I}_j} = u$, which is a sufficient condition for a set of transformations to be a test in sharp theories with purification. This is satisfied because, recalling Proposition 6,

$$\sum_j u P_{\mathsf{I}_j} = \sum_j a_{\mathsf{I}_j} = \sum_{i=1}^d \alpha_i^\dagger = u.$$

Because of the properties proved above, von Neumann measurements are repeatable and minimally disturbing measurements in the sense of Refs. [59,63]. Indeed, $a_{\mathsf{I}_j} P_{\mathsf{I}_j} = a_{\mathsf{I}_j}$, and

$$a_{\mathsf{I}_j} \sum_k P_{\mathsf{I}_k} = a_{\mathsf{I}_j} P_{\mathsf{I}_j} + \sum_{k \neq j} a_{\mathsf{I}_j} P_{\mathsf{I}_k} = a_{\mathsf{I}_j},$$

because for $k \neq j$ the P_{I_k}'s project on faces orthogonal to F_{I_j}.

The next proposition concerns the interplay between orthogonal projectors and the dagger.

Proposition 8. *For every normalised state ρ, and for every projector P_{I} on a face F_{I}, one has $(P_{\mathsf{I}}\rho)^\dagger = \rho^\dagger P_{\mathsf{I}}$.*

Proof. First of all, note that $0 \leq \|P_{\mathsf{I}}\rho\| \leq 1$, and it vanishes if and only if $\rho \in F_{\mathsf{I}}^\perp$. If $\rho \in F_{\mathsf{I}}^\perp$, then $\rho^\dagger P_{\mathsf{I}} = 0$, so the statement is trivially true. Now suppose $\|P_{\mathsf{I}}\rho\| > 0$. We will first prove the statement for normalised pure states ψ, then it is sufficient to extend it by linearity to all states. We will make use of the uniqueness of the dagger for normalised pure states. Then the statement is equivalent to proving

$$\left(\frac{P_{\mathsf{I}}\psi}{\|P_{\mathsf{I}}\psi\|} \right)^\dagger = \frac{\psi^\dagger P_{\mathsf{I}}}{\|P_{\mathsf{I}}\psi\|},$$

Noting that the term in brackets is a *normalised* pure state (by Purity Preservation), and that the RHS is a pure effect (again by Purity Preservation), by the uniqueness of the dagger for normalised pure states (cf. Theorem 6), it is enough to prove that

$$\frac{(\psi^\dagger P_{\mathsf{I}} | P_{\mathsf{I}}\psi)}{\|P_{\mathsf{I}}\psi\|^2} = 1;$$

in other words that $(\psi^\dagger P_{\mathsf{I}} | P_{\mathsf{I}}\psi) = \|P_{\mathsf{I}}\psi\|^2$. Recall that $P_{\mathsf{I}}^2 \doteq P_{\mathsf{I}}$ (Proposition 7), so $(\psi^\dagger P_{\mathsf{I}} | P_{\mathsf{I}}\psi) = (\psi^\dagger | P_{\mathsf{I}} | \psi)$. Now, $P_{\mathsf{I}}\psi = \|P_{\mathsf{I}}\psi\| \psi'$, where ψ' is a pure state in F_{I}. We have $(\psi^\dagger P_{\mathsf{I}} | P_{\mathsf{I}}\psi) = \|P_{\mathsf{I}}\psi\| (\psi^\dagger | \psi')$. We only need to prove that $(\psi^\dagger | \psi') = \|P_{\mathsf{I}}\psi\|$. Recall that $(\psi^\dagger | \psi') = \left(\psi'^\dagger | \psi \right)$ by Lemma 2, and that $\psi'^\dagger P_{\mathsf{I}} = \psi'^\dagger$ as $\psi' \in F_{\mathsf{I}}$, thus

$$\left(\psi^\dagger | \psi' \right) = \left(\psi'^\dagger | P_{\mathsf{I}} | \psi \right) = \|P_{\mathsf{I}}\psi\| \left(\psi'^\dagger | \psi' \right) = \|P_{\mathsf{I}}\psi\|.$$

By the uniqueness of the dagger for normalised pure states we conclude that $\left(\frac{P_{\mathsf{I}}\psi}{\|P_{\mathsf{I}}\psi\|} \right)^\dagger = \frac{\psi^\dagger P_{\mathsf{I}}}{\|P_{\mathsf{I}}\psi\|}$, namely $(P_{\mathsf{I}}\psi)^\dagger = \psi^\dagger P_{\mathsf{I}}$. □

A consequence of this proposition is that orthogonal projectors play nicely with the inner product of Lemma 2, namely for every $\xi, \eta \in \mathsf{St}_\mathbb{R}(A)$ one has

$$\langle P_{\mathsf{I}}\xi, \eta \rangle = \langle \xi, P_{\mathsf{I}}\eta \rangle. \tag{11}$$

In other words, projections are symmetric with respect to the inner product.

The last property we need is a generalisation of the results of Proposition 7.

Proposition 9. *Fixing a pure maximal set $\{\alpha_i\}_{i=1}^d$, and considering $\mathsf{I}, \mathsf{J} \subseteq \{1, \dots, d\}$, we have $P_{\mathsf{I}} P_{\mathsf{J}} \doteq P_{\mathsf{I} \cap \mathsf{J}}$.*

Proof. First let us prove that

$$P_\mathsf{I} P_\mathsf{J} \rho = \|P_\mathsf{I} P_\mathsf{J} \rho\| \, \rho' \tag{12}$$

for every normalised state ρ, where $\rho' \in F_{\mathsf{I}\cap\mathsf{J}}$. Let us show that $\|P_\mathsf{I}P_\mathsf{J}\rho\| = (a_{\mathsf{I}\cap\mathsf{J}}|\rho)$. By Proposition 6, $(u|P_\mathsf{I}P_\mathsf{J}|\rho) = (a_\mathsf{I}|P_\mathsf{J}|\rho)$. Now, recalling that $a_\mathsf{I} = \sum_{i\in\mathsf{I}} \alpha_i^\dagger$,

$$(a_\mathsf{I}|P_\mathsf{J}|\rho) = \sum_{i\in\mathsf{I}\cap\mathsf{J}} \left(\alpha_i^\dagger \Big| P_\mathsf{J} \Big| \rho\right) + \sum_{i\in\mathsf{I}\setminus\mathsf{J}} \left(\alpha_i^\dagger \Big| P_\mathsf{J} \Big| \rho\right) = \sum_{i\in\mathsf{I}\cap\mathsf{J}} \left(\alpha_i^\dagger \Big| \rho\right) = (a_{\mathsf{I}\cap\mathsf{J}}|\rho) \,,$$

where we have used the fact that $\alpha_i^\dagger P_\mathsf{J} = \alpha_i^\dagger$ if $i \in \mathsf{J}$, and $\alpha_i^\dagger P_\mathsf{J} = 0$ if $i \notin \mathsf{J}$. If $\rho \in F_{\mathsf{I}\cap\mathsf{J}}^\perp$, both the LHS and the RHS of Equation (12) vanish, and the statement is trivially satisfied. Now, let us assume $\rho \notin F_{\mathsf{I}\cap\mathsf{J}}^\perp$, in this case $(a_{\mathsf{I}\cap\mathsf{J}}|\rho) > 0$. We wish to prove that $(a_{\mathsf{I}\cap\mathsf{J}}|P_\mathsf{I}P_\mathsf{J}|\rho) = (a_{\mathsf{I}\cap\mathsf{J}}|\rho)$. Recalling the expression of $a_{\mathsf{I}\cap\mathsf{J}}$, we have

$$\sum_{i\in\mathsf{I}\cap\mathsf{J}} \left(\alpha_i^\dagger \Big| P_\mathsf{I}P_\mathsf{J} \Big| \rho\right) = \sum_{i\in\mathsf{I}\cap\mathsf{J}} \left(\alpha_i^\dagger \Big| P_\mathsf{J} \Big| \rho\right) = \sum_{i\in\mathsf{I}\cap\mathsf{J}} \left(\alpha_i^\dagger \Big| \rho\right) = (a_{\mathsf{I}\cap\mathsf{J}}|\rho) \,,$$

again by the properties of P_I and P_J. This means that $P_\mathsf{I}P_\mathsf{J}$ maps every normalised state to a state of $F_{\mathsf{I}\cap\mathsf{J}}$, up to normalisation.

Now let us prove that $(P_\mathsf{I}P_\mathsf{J})^2 \doteq P_\mathsf{I}P_\mathsf{J}$. First note that $F_{\mathsf{I}\cap\mathsf{J}} \subseteq F_\mathsf{I}$. Indeed, suppose $\rho \in F_{\mathsf{I}\cap\mathsf{J}}$, then

$$(a_\mathsf{I}|\rho) = \sum_{i\in\mathsf{I}\cap\mathsf{J}} \left(\alpha_i^\dagger \Big| \rho\right) + \sum_{i\in\mathsf{I}\setminus\mathsf{J}} \left(\alpha_i^\dagger \Big| \rho\right) = (a_{\mathsf{I}\cap\mathsf{J}}|\rho) = 1,$$

where we have used the fact that $\left(\alpha_i^\dagger|\rho\right) = 0$ if $i \notin \mathsf{I}\cap\mathsf{J}$. By a similar argument, $F_{\mathsf{I}\cap\mathsf{J}} \subseteq F_\mathsf{J}$. Now, $P_\mathsf{I}P_\mathsf{J}\rho = \|P_\mathsf{I}P_\mathsf{J}\rho\| \, \rho'$, with $\rho' \in F_{\mathsf{I}\cap\mathsf{J}}$. Then $(P_\mathsf{I}P_\mathsf{J})^2 \rho = \|P_\mathsf{I}P_\mathsf{J}\rho\| \, P_\mathsf{I}P_\mathsf{J}\rho'$. However, $\rho' \in F_\mathsf{J}$, so $P_\mathsf{J}\rho' = \rho'$, and, similarly, $\rho' \in F_\mathsf{I}$, so $P_\mathsf{I}\rho' = \rho'$. Consequently,

$$(P_\mathsf{I}P_\mathsf{J})^2 \rho = \|P_\mathsf{I}P_\mathsf{J}\rho\| \, \rho' = P_\mathsf{I}P_\mathsf{J}\rho,$$

proving that $(P_\mathsf{I}P_\mathsf{J})^2 \doteq P_\mathsf{I}P_\mathsf{J}$.

Now let us prove that for every $\xi \in \mathsf{St}_\mathbb{R}(A)$, we have $(P_\mathsf{I}P_\mathsf{J}\xi)^\dagger = \xi^\dagger P_\mathsf{I}P_\mathsf{J}$. Following the lines of proof of Proposition 8, let us show that this is true when ξ is a normalised pure state ψ. This boils down to showing that

$$\left(\psi^\dagger P_\mathsf{I}P_\mathsf{J} \Big| P_\mathsf{I}P_\mathsf{J}\psi\right) = \|P_\mathsf{I}P_\mathsf{J}\psi\|^2 \,.$$

The proof goes on as for Proposition 8, noting that if $\psi' \in F_{\mathsf{I}\cap\mathsf{J}}$, then $\psi'^\dagger P_\mathsf{I}P_\mathsf{J} = \psi'^\dagger$ because $\psi'^\dagger P_\mathsf{I} = \psi'^\dagger$ as $\psi' \in F_\mathsf{I}$, and, similarly, $\psi'^\dagger P_\mathsf{J} = \psi'^\dagger$ as $\psi' \in F_\mathsf{J}$. Eventually we find that for pure states $(P_\mathsf{I}P_\mathsf{J}\psi)^\dagger = \psi^\dagger P_\mathsf{I}P_\mathsf{J}$, and by linearity this means that $(P_\mathsf{I}P_\mathsf{J}\xi)^\dagger = \xi^\dagger P_\mathsf{I}P_\mathsf{J}$.

A consequence of this property is that $\langle P_\mathsf{I}P_\mathsf{J}\xi, \eta \rangle = \langle \xi, P_\mathsf{I}P_\mathsf{J}\eta \rangle$, for all $\xi, \eta \in \mathsf{St}_\mathbb{R}(A)$. These linear maps on $\mathsf{St}_\mathbb{R}(A)$ are such that $\mathsf{St}_\mathbb{R}(A) = \mathrm{im}\, P_\mathsf{I}P_\mathsf{J} \oplus \ker P_\mathsf{I}P_\mathsf{J}$, and $\ker P_\mathsf{I}P_\mathsf{J}$ is the orthogonal subspace to $\mathrm{im}\, P_\mathsf{I}P_\mathsf{J}$, hence it is uniquely defined once $\mathrm{im}\, P_\mathsf{I}P_\mathsf{J}$ is fixed. Note that for any projector P_I we have $\mathrm{im}\, P_\mathsf{I} = \mathrm{span}\, F_\mathsf{I}$, and we have just proved that $\mathrm{im}\, P_\mathsf{I}P_\mathsf{J} = \mathrm{span}\, F_{\mathsf{I}\cap\mathsf{J}} = \mathrm{im}\, P_{\mathsf{I}\cap\mathsf{J}}$. Having the same image, and consequently the same kernel, $P_\mathsf{I}P_\mathsf{J}$ and $P_{\mathsf{I}\cap\mathsf{J}}$ agree on a basis of $\mathsf{St}_\mathbb{R}(A)$, therefore they agree also on all states of A, meaning that $P_\mathsf{I}P_\mathsf{J} \doteq P_{\mathsf{I}\cap\mathsf{J}}$. $\quad\square$

5.3. Main Result

Proposition 29 of [4] asserts that theories satisfying two postulates, Strong Symmetry and Projectivity, have higher-order interference if and only if their projectors (in our terminology here) preserve purity. A close examination of its proof, and those of all lemmas and propositions used in its proof—notably Lemma 22 and Propositions 18, 25, 26, and 28 of [4]—reveals that only premises weaker than the conjunction of Strong Symmetry and Projectivity are used: self-duality, the "spectral-like decomposition" of effects as in Lemma 1 above, the fact that faces are determined by subsets of maximal distinguishable sets of states as in Section 5.2 above, the existence of projectors onto each face in the

sense of Definition 8 above, and the fact that these are symmetric with respect to the self-dualising inner product (i.e., orthogonal projectors), and satisfy Proposition 9 above. We have established these weaker premises for sharp theories with purification, and moreover, we have established in Proposition 6 that their projectors preserve purity, so we have proved:

Theorem 7. *In any sharp theory with purification there can be no nth order interference for $n \geq 3$.*

5.4. Jordan-Algebraic Structure

Our results also imply that systems, and therefore also the "subsystems" associated with their faces, are operationally equivalent to finite-dimensional Jordan-algebraic systems. These are systems A for which $\mathsf{St}_+(A)$ is the cone of squares in a finite-dimensional Euclidean Jordan algebra (EJA) and $\mathsf{Eff}_+(A)$ is identified with the same cone, with evaluation of effects on states given by the inner product and the Jordan unit as the deterministic effect. (See [37] for more on Jordan algebraic operational systems, and [61] for a mathematical treatment.)

Theorem 8. *In a sharp theory with purification, every system A has both $\mathsf{St}_+(A)$ and $\mathsf{Eff}_+(A)$ isomorphic to the cone of squares in a Euclidean Jordan algebra (EJA) via isomorphisms S and T such that $(a|\rho) = \langle Ta, S\rho \rangle$, where $\langle \bullet, \bullet \rangle$ is the canonical inner product on the EJA, and T takes the deterministic effect to the Jordan unit.*

Proof. The proof uses results of Alfsen and Shultz [64], for which we refer to [61]. Theorem 9.33 in [61] implies that finite-dimensional systems with symmetry of transition probabilities (STP), a type of projection operator they call "compression" associated with every face, and whose compressions preserve purity, have state spaces affinely isomorphic to the state spaces of Euclidean Jordan algebras. Sharp theories with purification satisfy STP, as noted following Lemma 2 above. Our projectors are easily shown to be examples of compressions by the same argument as in Theorem 17 of [4]; this argument uses only properties satisfied by our projectors (the same ones needed in the proof of Theorem 7, except for Purity Preservation) and does not need Strong Symmetry. As shown above, our projectors also preserve purity. □

Since faces of Jordan-algebraic systems are also Jordan-algebraic (to see this, combine a result of Iochum [65] (Theorem 5.32 in [61]), whose finite dimensional case is that all faces of EJAs are the positive part of the images of compressions, with the facts (cf. pp. 22–26 of [61]) that every face of the cone of squares is the image of such a compression P ([61], Lemma 1.39), and also a Jordan subalgebra whose unit is the image of the order unit under P ([61], Proposition 1.43).), so are the faces of state spaces in sharp theories with purification. However, it is not the case that in sharp theories with purification, each face of a system is necessarily isomorphic to a stand-alone system of the theory (an *object* of the category, in the categorical formulation), but, it is always possible to extend the theory such that they are. Every category has a *Cauchy completion*: this is a minimal extension of the category such that every idempotent morphism $\pi : A \to A$ can be written as a retraction-section pair, i.e., as the composition $\pi = \sigma \circ \rho$, with $\rho : A \to B$ and $\sigma : B \to A$, such that the reverse composition $\rho \circ \sigma$ is the identity morphism on B. When the idempotents are projectors P like the ones we consider here, B will be a system isomorphic to the face $\mathrm{im}_+(P)$. Of course, since there may be idempotents beyond the projectors onto faces (for example, decoherence of a set of orthogonal subspaces, or damping to a fixed state, in quantum theory), Cauchy completion of an operational theory T may add many objects in addition to ones isomorphic to faces of systems of T; indeed, for many operational theories (e.g., ones possessing idempotent decoherence maps) this will add some classical systems. This is indeed the case for quantum theory where the Cauchy completion leads to the category of finite-dimensional C*-algebras and completely positive maps [66]. The Cauchy completion can be thought of as adding in all operationally accessible systems that can be simulated on the physical system via a consistent restriction on the allowed states, effects and transformations. The Cauchy completion of a sharp theory with purification will likely satisfy the Ideal Compression postulate by virtue of containing the faces that are images of orthogonal projectors; but there are also non-Cauchy complete theories that satisfy

it, e.g., the category CPM of finite-dimensional quantum systems and CP maps, in which all systems, and also all images of orthogonal projectors as defined above, are fully coherent quantum systems, but there are no classical systems.

In [37], some categories, including dagger-compact-closed categories, of Jordan algebraic systems were constructed; these categories are equivalent to operational theories as we use the term here. Although sharp theories with purification also have Jordan algebraic state and effect spaces, it is interesting to note that some of the explicit examples in [30,49] involve composites different from those that would be obtained in the categories considered in [37] for systems with the same state spaces. On the other hand, the category combining real and quaternionic systems in [37] does *not* satisfy Purity Preservation by parallel composition and hence falls outside the class of sharp theories with purification, although its filters do preserve purity. Of course, the failure of Purity Preservation by parallel composition seems likely to allow phenomena like the nonextensiveness of entropy when products of states are taken, which could warrant focusing on sharp theories with purification in thermodynamically motivated work such as [30].

That Jordan-algebraic systems lack higher-order interference was shown by Barnum and Ududec ([12]; announced in [67]) and by Niestegge [68]; combining this with Theorem 8 gives another way to see that our results on sharp theories with purification imply the absence of higher-order interference. Moreover, as not all EJAs satisfy our postulates, it is clear that our postulates are sufficient but not necessary conditions for ruling out higher-order interfence.

6. Discussion and Conclusions

We proved that in sharp theories with purification multi-slit experiments must have a pure projector structure and, moreover, such theories exhibit at most second-order interference. Hence these theories are, at least conceptually, very "close" to quantum theory. Moreover, recent work has shown that sharp theories with purification are close to quantum theory in terms of other physical and information processing features. Indeed, such theories possess quantum-like contextuality behaviour [59,63], quantum-like computation [7,8], and quantum-like thermodynamic Properties [30,49,54]. Recall from Section 4 that quantum theory is not the only example of a generalised probabilistic theory satisfying these principles. Hence Causality, Purity Preservation, Pure Sharpness, and Purification do not recover the entire quantum formalism.

However, if one were to introduce the Ideal Compression and Local Discriminability principles of the reconstruction of quantum theory due to Chiribella, D'Ariano, and Perinotti [20], one would indeed regain the entire quantum formalism. Indeed, both additional principles are necessary: Local Discriminability to preclude real quantum theory and Ideal Compression to preclude the contrived—yet admissible—example of the theory in which all systems are composites of qubits. Sharp theories with purification thus serve as a fertile test-bed for physics that is conceptually quite close to that predicted by the quantum world, but which may diverge from it in certain small, yet interesting, ways.

Finding Higher Order Interference

To date there has been no experiment that has found higher-order interference, at least, none that cannot be explained by taking into account the fact that the "sets of histories are not mutually exclusive" [2,35]. However, this might be due to the specific experimental set-up employed, rather than a fundamental preclusion of higher-order interference in nature. We show here that many of the properties needed to rule out observing higher-order interference are in fact quite natural assumptions which appear to be suggested by the experimental set-up employed. This suggests that the experimental set-up itself may implicitly rule out observing higher-order interference from the outset.

The main result of the current work is that sharp theories with purification can never exhibit higher-order interference in any experiment. However, in a wider class of theories, we still will not observe higher-order interference in a particular experiment if the following three conditions are met;

hence, to have any chance of observing higher-order interference, experiments must be designed in order to try to violate these conditions.

1. The transformations corresponding to blocking slits satisfy: $T_I T_J = T_{I \cap J}$. By this we mean that they share several properties with the projectors P_I of Section 5: if we define the effects $a_I = u T_I$ and the faces F_I and $F_I^{\perp}$ as in Section 5.2, i.e., as the 1-set and 0-set of a_I, then the T_I are assumed to be *orthogonal projectors* in the sense of Definition 8, and to be both idempotent and "orthogonal" ($T_I T_J = 0$) if I and J are disjoint (as in Proposition 7).
2. The T_I's map pure states to pure states
3. The T_I's are self-adjoint.

The first of these is generally expected as only those slits belonging to both I and J will not be blocked by either T_I or T_J, and so should hold in this experimental set-up for any theory that can describe it.

The second assumption, which is also natural given the multi-slit set-up, is that, in an idealised scenario, the slits should not introduce fundamental noise. That is, if an input state ρ is pure, i.e., has no classical noise associated with it, then $T_I \rho$ should also be pure. Hence it appears natural to assume that T_I maps pure states to pure states. Violating this principle by just adding noise to the experiment does not seem likely to demonstrate higher-order interference. A more plausible way to violate this however would be if the particle passing through the slits were to become entangled with some degree of freedom associated with them, if we do not have access to this degree of freedom then this would send a pure input to a mixed state.

The final assumption is far less general than the others, as it places a constraint on the theory. That is, to even discuss whether a transformation is self-adjoint (cf. also Appendix B), one requires that the theory itself be self-dual. To fully understand what this assumption entails, one needs an operational or physical interpretation of the self-dualising inner product (see [69] for an example of such an interpretation). However, intuitively this notion reflects the inherent symmetry of the experimental set-up. Here one could consider propagation from the source to the effect or from the effect to the source as being "dual" to one another and, moreover, that the physical blocking of slits has an equivalent effect in either situation. That is, the assumption of self-adjointness corresponds to the statement that the projector has an equivalent action on the effects associated with a particular slit as it does on the states which can pass through them.

If an experiment satisfies these assumptions then for any self-dual theory it was shown in [4] (Proposition 29) that we will not see higher-order interference in this experiment. Hence any set of physical principles which ensure these assumptions hold will rule out higher-order interference. Because the mathematical assumptions involved in formalising a multi-slit experiment are so natural when interpreted operationally, perhaps one should search for higher-order interference in set-ups that don't seem to preclude it from the outset. This could involve "asymmetric" multi-slit set-ups that are not obviously time-symmetric in an arbitrary generalised probabilistic theory. One could also consider experiments that search for higher-order phases [8], a reformulation of higher-order interference that makes no reference to projectors and hence does not preclude certain generalised theories from the outset. The assumption that nature is self-dual could also be rejected; this poses the question as to whether it is possible to find a direct experimental test of this principle.

Appendix A. Norms and Fidelity

Appendix A.1. Operational Norm and Dagger Norm

In Ref. [19] the operational norm for every vector $\xi \in \mathsf{St}_{\mathbb{R}}(A)$ was introduced:

$$\|\xi\| := \sup_{a \in \mathsf{Eff}(A)} (a|\xi) - \inf_{a \in \mathsf{Eff}(A)} (a|\xi)$$

As pointed out in [19], in quantum theory the operational norm coincides with the trace norm. The analogy is apparent also in sharp theories with purification.

Proposition A1. *Let $\xi \in \mathsf{St}_{\mathbb{R}}(A)$ be diagonalised as $\xi = \sum_{i=1}^{d} x_i \alpha_i$. Then $\|\xi\| = \sum_{i=1}^{d} |x_i|$.*

Proof. Let us separate the terms with non-negative eigenvalues from the terms with negative eigenvalues, so that we can write $\xi = \xi_+ - \xi_-$, where $\xi_+ := \sum_{x_i \geq 0} x_i \alpha_i$, and $\xi_- = \sum_{x_i < 0} (-x_i) \alpha_i$. Clearly, $\xi_+, \xi_- \in \mathsf{St}_+(A)$. In order to achieve the supremum of $(a|\xi)$ we must have $(a|\xi_-) = 0$. Moreover,

$$(a|\xi_+) = \sum_{x_i \geq 0} x_i (a|\alpha_i) \leq \sum_{x_i \geq 0} x_i$$

since $(a|\alpha_i) \leq 1$ for every i. The supremum of $(a|\xi_+)$ is achieved by $a = \sum_{x_i \geq 0} \alpha_i^\dagger$. Hence $\sup_a (a|\xi) = \sum_{x_i \geq 0} x_i$. By a similar argument, one shows that $\inf_a (a|\xi) = \sum_{x_i < 0} x_i$. Therefore

$$\|\xi\| = \sum_{x_i \geq 0} x_i + \sum_{x_i < 0} (-x_i) = \sum_{i=1}^{d} |x_i|.$$

□

For $p \geq 1$, the p-norm of a vector $\mathbf{x} \in \mathbb{R}^d$ is defined as $\|\mathbf{x}\|_p := \left(\sum_{i=1}^{d} |x_i|^p \right)^{\frac{1}{p}}$, thus we have $\|\xi\| = \|\mathbf{x}\|_1$, where $\mathbf{x}$ is the spectrum of ξ.

In sharp theories with purification we have an additional norm, the dagger norm, defined in Section 5.1. The dagger norm of a vector $\xi \in \mathsf{St}_{\mathbb{R}}(A)$ is $\|\xi\|_\dagger = \sqrt{\sum_{i=1}^{d} x_i^2}$, where the x_i's are the eigenvalues of ξ. It is obvious from the very definition that $\|\xi\|_\dagger = \|\mathbf{x}\|_2$. Thanks to these results following from diagonalisation, we can derive the standard bounds between the two norms, by making use of the well-known bounds $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{d} \|\mathbf{x}\|_2$, which imply

$$\|\xi\|_\dagger \leq \|\xi\| \leq \sqrt{d} \|\xi\|_\dagger. \tag{A1}$$

Note that, unlike Ref. [70], here the bounds are derived without assuming Bit Symmetry [4,71].

If we take ξ to be a normalised state ρ, its eigenvalues form a probability distribution, and we have $\|\rho\|_\dagger \leq 1$, with equality if and only if ρ is pure. Note that $\|\rho\|_\dagger$ is a Schur-convex function [72] of the eigenvalues of ρ, so it is a purity monotone [30]. As such, it attains its minimum on the invariant state, which is $\|\chi\|_\dagger = \frac{1}{\sqrt{d}}$, so for every normalised state one has

$$\frac{1}{\sqrt{d}} \leq \|\rho\|_\dagger \leq 1,$$

consistently with the bounds (A1). The square of the dagger norm, still a Schur-convex function, was called *purity* in Refs. [70,73]. Consequently $1 - \|\rho\|_+^2$ is a measure of mixedness, sometimes called the *impurity* $I(\rho)$ of ρ. The impurity can be extended to subnormalised states by defining it as $I(\rho) := (\text{Tr}\,\rho)^2 - \|\rho\|_+^2$ [4].

The two norms behave differently under channels applied to states. In Ref. [19] it was shown that in causal theories the operational norm of a state ρ is preserved by channels: $\|\mathcal{C}\rho\| = \|\rho\|$ for every channel $\mathcal{C}$, because channels are such that $u\mathcal{C} = u$.

Instead the dagger norm shows a different behaviour. To describe it, it is useful to divide channels into two classes: unital and non-unital channels [49].

Definition A1. *A channel* $\mathcal{D} \in \text{Transf}(A, B)$ *is* unital *if* $\mathcal{D}\chi_A = \chi_B$.

Unital channels do not increase the dagger norm of states.

Proposition A2. *If* $\mathcal{D}$ *is a unital channel, then* $\|\mathcal{D}\rho\|_+ \leq \|\rho\|_+$, *for every normalised state* ρ.

Proof. Unital channels can be chosen as free operations for the resource theory of purity [49]. In Ref. [49] it was shown that the spectrum of $\mathcal{D}\rho$ is majorised by the spectrum of ρ (see Ref. [72] for a definition of majorisation and Schur-convex functions). Since the dagger norm is a Schur-convex function, we have $\|\mathcal{D}\rho\|_+ \leq \|\rho\|_+$. □

Clearly if $\mathcal{D}$ is reversible, the dagger norm is preserved, by Proposition 4.

For non-unital channels there is at least one state—the invariant state χ—for which the dagger norm increases. Indeed, if $\mathcal{C}$ is non-unital, χ is majorised by $\mathcal{C}\chi$, whence $\|\chi\|_+ \leq \|\mathcal{C}\chi\|_+$. Is it true, then, that non-unital channels increase the dagger norm of all states? The answer is clearly negative. Consider the non-unital channel mapping all states to a fixed *mixed* state $\rho_0 \neq \chi$. For some states, e.g., the invariant state, the dagger norm will increase, for others, e.g., pure states, the dagger norm will decrease because it is a purity monotone. In short, for non-unital channels there is no uniform behaviour of the dagger norm.

Appendix A.2. Dagger Fidelity

The inner product defined in Section 5.1 allows us to define a fidelity-like quantity, called the *dagger fidelity*.

Definition A2. *Given two normalised states* ρ *and* σ, *the* dagger fidelity *is defined as*

$$F_+(\rho, \sigma) = \frac{\langle \rho, \sigma \rangle}{\|\rho\|_+ \|\sigma\|_+}.$$

The dagger fidelity measures the overlap between two states. It shares some properties with the fidelity in quantum theory (cf. for instance Ref. [74]), despite *not* coinciding with it. The first, obvious one, is that $F_+(\rho, \sigma) = F_+(\sigma, \rho)$.

To prove the other properties we need the following lemma, generalising one of the results of Ref. [30].

Lemma A1. *Let* $\{\rho_i\}_{i=1}^n$ *be perfectly distinguishable states. Then* $(\rho_i^\dagger | \rho_j) = \|\rho_i\|_+^2 \delta_{ij}$.

Proof. Clearly what we need to prove is that $(\rho_i^\dagger | \rho_j) = 0$ if $i \neq j$. Let $\{a_i\}_{i=1}^n$ be the perfectly distinguishing test, and let ρ_i be diagonalised as $\rho_i = \sum_{k=1}^{r_i} p_{k,i} \alpha_{k,i}$, where $p_{k,i} > 0$ for all $k = 1, \ldots, r$. We have $(a_i | \rho_i) = 1$, hence by Proposition 2 there exists a non-disturbing pure transformation $\mathcal{T}_i$ such that $\mathcal{T}_i =_{\rho_i} \mathcal{I}$. Specifically, we have that $\mathcal{T}_i \alpha_{k,i} = \alpha_{k,i}$. Moreover if $i \neq j$, we have $(u | \mathcal{T}_i | \rho_j) \leq (a_i | \rho_j) = 0$, whence $(u | \mathcal{T}_i | \rho_j) = 0$. This means that $\mathcal{T}_i \rho_j = 0$ for all $j \neq i$.

Now, consider

$$\left(\alpha_{k,i}^\dagger \middle| \mathcal{T}_i \middle| \alpha_{k,i}\right) = \left(\alpha_{k,i}^\dagger \middle| \alpha_{k,i}\right) = 1,$$

where we have used the fact that $\mathcal{T}_i \alpha_{k,i} = \alpha_{k,i}$. Since $\alpha_{k,i}^\dagger \mathcal{T}_i$ is a pure effect, it must be $\alpha_{k,i}^\dagger \mathcal{T}_i = \alpha_{k,i}^\dagger$ by Theorem 6. By linearity we have $\rho_i^\dagger \mathcal{T}_i = \rho_i^\dagger$. Now, using this fact, for all $j \neq i$

$$\left(\rho_i^\dagger \middle| \rho_j\right) = \left(\rho_i^\dagger \middle| \mathcal{T}_i \middle| \rho_j\right) = 0,$$

because $\mathcal{T}_i \rho_j = 0$. $\square$

Recalling that $\left(\rho^\dagger \middle| \sigma\right) = \langle \rho, \sigma \rangle$, this lemma means that perfectly distinguishable states form an orthogonal set. Specifically, if the states are pure, the set is orthonormal.

The following proposition extends and generalises the properties of the self-dualising inner product of Ref. [71].

Proposition A3. *The dagger fidelity has the following properties, for all normalised states ρ and σ.*

1. $0 \leq F_\dagger\left(\rho, \sigma\right) \leq 1$;
2. $F_\dagger\left(\rho, \sigma\right) = 0$ *if and only if ρ and σ are perfectly distinguishable;*
3. $F_\dagger\left(\rho, \sigma\right) = 1$ *if and only if $\rho = \sigma$;*
4. $F_\dagger\left(\mathcal{U}\rho, \mathcal{U}\sigma\right) = F_\dagger\left(\rho, \sigma\right)$, *for every reversible channel $\mathcal{U}$.*

Proof. Let us prove the various properties.

1. Recall that $\langle \rho, \sigma \rangle = \left(\rho^\dagger \middle| \sigma\right) \geq 0$, whence $F_\dagger\left(\rho, \sigma\right) \geq 0$. Moreover, by Schwarz inequality, $\langle \rho, \sigma \rangle \leq \|\rho\|_\dagger \|\sigma\|_\dagger$, so $F_\dagger\left(\rho, \sigma\right) \leq 1$.
2. Suppose ρ and σ are perfectly distinguishable, then by Lemma A1 $\langle \rho, \sigma \rangle = 0$, implying $F_\dagger\left(\rho, \sigma\right) = 0$. Now suppose $F_\dagger\left(\rho, \sigma\right) = 0$; then $\langle \rho, \sigma \rangle = 0$. Let $\rho = \sum_{i=1}^r p_i \alpha_i$ be a diagonalisation of ρ, with $p_i > 0$, for all $i = 1, \ldots, r$, and $r \leq d$. We have $\sum_{i=1}^r p_i \left(\alpha_i^\dagger \middle| \sigma\right) = 0$, which means that $\left(\alpha_i^\dagger \middle| \sigma\right) = 0$ for $i = 1, \ldots, r$. This means that we can build an observation-test that distinguishes ρ and σ perfectly by taking $\{a, u - a\}$, where $a = \sum_{i=1}^r \alpha_i^\dagger$.
3. Clearly, if $\rho = \sigma$, $\langle \rho, \sigma \rangle = \|\rho\|_\dagger^2$, whence $F_\dagger\left(\rho, \sigma\right) = 1$. Conversely, suppose $F_\dagger\left(\rho, \sigma\right) = 1$. This means that $\langle \rho, \sigma \rangle = \|\rho\|_\dagger \|\sigma\|_\dagger$. By Schwarz inequality, this is true if and only if $\rho = \lambda \sigma$, for some $\lambda \in \mathbb{R}$. Since both states are normalised, $\lambda = 1$, yielding $\rho = \sigma$.
4. This property follows by Proposition 4, because the inner product and the dagger norm are invariant under reversible channels.

$\square$

Note that Property 3 captures the sharpness of the dagger for all normalised states [69].

A property involving tensor product of states is the following.

Proposition A4. *For all normalised states ρ_1, ρ_2, σ_1, σ_2 one has*

$$F_\dagger\left(\rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2\right) = F_\dagger\left(\rho_1, \sigma_1\right) F_\dagger\left(\rho_2, \sigma_2\right)$$

The proof needs the following easy lemma.

Lemma A2. *Let $\rho, \sigma \in \mathsf{St}_1\left(\mathrm{A}\right)$, then $\left(\rho \otimes \sigma\right)^\dagger = \rho^\dagger \otimes \sigma^\dagger$.*

Proof. Let us prove the result for ρ and σ pure, the general result will follow by linearity. By Purity Preservation, $\rho \otimes \sigma$ and $\rho^\dagger \otimes \sigma^\dagger$ are pure, and one has $\left(\rho^\dagger \otimes \sigma^\dagger \middle| \rho \otimes \sigma\right) = 1$. By Theorem 6, $\left(\rho \otimes \sigma\right)^\dagger = \rho^\dagger \otimes \sigma^\dagger$. $\square$

Now comes the actual proof.

Proof of Proposition A4. We have

$$F_{\dagger}\left(\rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2\right) = \frac{\langle \rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2 \rangle}{\|\rho_1 \otimes \rho_2\|_{\dagger} \|\sigma_1 \otimes \sigma_2\|_{\dagger}}.$$

Now, by Lemma A2,

$$\langle \rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2 \rangle = \left(\rho_1^{\dagger} \otimes \rho_2^{\dagger} \middle| \sigma_1 \otimes \sigma_2\right) = \left(\rho_1^{\dagger} \middle| \sigma_1\right)\left(\rho_2^{\dagger} \middle| \sigma_2\right) = \langle \rho_1, \sigma_1 \rangle \langle \rho_2, \sigma_2 \rangle.$$

Furthermore,

$$\|\rho_1 \otimes \rho_2\|_{\dagger} = \sqrt{\langle \rho_1 \otimes \rho_2, \rho_1 \otimes \rho_2 \rangle} = \sqrt{\langle \rho_1, \rho_1 \rangle \langle \rho_2, \rho_2 \rangle} = \|\rho_1\|_{\dagger} \|\rho_2\|_{\dagger}.$$

Putting everything together,

$$F_{\dagger}\left(\rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2\right) = \frac{\langle \rho_1, \sigma_1 \rangle}{\|\rho_1\|_{\dagger} \|\sigma_1\|_{\dagger}} \cdot \frac{\langle \rho_2, \sigma_2 \rangle}{\|\rho_2\|_{\dagger} \|\sigma_2\|_{\dagger}} = F_{\dagger}\left(\rho_1, \sigma_1\right) F_{\dagger}\left(\rho_2, \sigma_2\right).$$

□

Appendix B. Dagger of All Transformations

Inspired by the results of Lemma 2, in sharp theories with purification, we can extend the dagger to all transformations, a feature often present in process theories [44,45,69,75].

Definition A3. *Given the transformation $\mathcal{A} \in$ Transf (A, B), its* dagger *(or* adjoint*) is a linear transformation $\mathcal{A}^{\dagger}$ from B to A defined as*

(A2)

for every system S, and every state $\rho \in \mathsf{St}_1 (B \otimes S)$.

This definition specifies the dagger of a transformation completely, thanks to Equation (2). Note that Lemma 2 allows us to formulate Equation (10) in term of effects and their dagger:

$$\left(a \middle| b^{\dagger}\right) = \left(b \middle| a^{\dagger}\right)$$

for all effects a, and b. In this way, Definition A3 can be recast in equivalent terms by taking b as the term in round brackets in the RHS of Equation (A2). This yields

(A3)

for every system S, every state $\rho \in \mathsf{St}_1 (B \otimes S)$, and every effect $E \in \mathsf{Eff} (A \otimes S)$.

The dagger of a transformation may not be a physical transformation, i.e., it may send physical states to non-physical ones. Indeed, the action of $\mathcal{A}^{\dagger} \otimes \mathcal{I}$ on a generic state (the LHS of Equation (A2)) is defined as the dagger of an effect. However, not all daggers of effects are physical states. For instance, take the deterministic effect $u = \sum_{i=1}^{d} \alpha_i^{\dagger}$, where $\{\alpha_i\}_{i=1}^{d}$ is a pure maximal set. Its dagger is $u^{\dagger} = \sum_{i=1}^{d} \alpha_i = d\chi$, which is a supernormalised (and hence non-physical) state.

For channels, we can give a necessary condition for the existence of a physical dagger of the channel.

Proposition A5. *Let* $\mathcal{C} \in$ Transf (A, B) *be a channel. If* $\mathcal{C}^\dagger$ *is a physical transformation, then* $\mathcal{C}$ *is unital, and* $\mathcal{C}^\dagger$ *itself is a unital channel.*

Proof. If $\mathcal{C}^\dagger$ is a physical transformation, then, for every normalised state $\rho \in$ St$_1$ (B), we have $\|\mathcal{C}^\dagger \rho\| \leq 1$, or in other words, $(u|\mathcal{C}^\dagger|\rho) \leq 1$. By Equation (A3), $(u|\mathcal{C}^\dagger|\rho) = (\rho^\dagger|\mathcal{C}|u^\dagger)$, so the condition $\|\mathcal{C}^\dagger \rho\| \leq 1$ is equivalent to

$$\left(\rho^\dagger \middle| \mathcal{C} \middle| \chi\right) \leq \frac{1}{d}, \tag{A4}$$

with equality if and only if $\mathcal{C}^\dagger$ is a channel. Suppose by contradiction that $\mathcal{C}$ is not unital, then $\mathcal{C}\chi = \rho_0 \neq \chi$. Diagonalise ρ_0 as $\rho_0 = \sum_{i=1}^d p_i \alpha_i$, where $p_1 \geq p_2 \geq \ldots \geq p_d \geq 0$, and $p_1 > \frac{1}{d}$. Then taking ρ to be α_1 in $\left(\rho^\dagger|\mathcal{C}|\chi\right)$ yields p_1, but $p_1 > \frac{1}{d}$, contradicting Equation (A4).

Being $\mathcal{C}$ unital, we have that

$$\left(\rho^\dagger \middle| \mathcal{C} \middle| \chi\right) = \left(\rho^\dagger \middle| \chi\right) = \frac{1}{d}\text{Tr}\,\rho = \frac{1}{d},$$

showing that $\mathcal{C}^\dagger$ is itself a channel. Let us prove it is unital. The action of $\mathcal{C}^\dagger$ on χ is defined in Equation (A2), so

$$\mathcal{C}^\dagger \chi = \left(\chi^\dagger \mathcal{C}\right)^\dagger = \frac{1}{d}\left(u\mathcal{C}\right)^\dagger = \frac{1}{d}u^\dagger = \chi,$$

where we have used the fact that $\mathcal{C}$ is a channel, so $u\mathcal{C} = u$. This proves that $\mathcal{C}^\dagger$ is unital. $\square$

We can prove that the dagger of a transformation has some nice properties.

Proposition A6. *For every transformation* $\mathcal{A} \in$ Transf (A, B), *one has* $\left(\mathcal{A}^\dagger\right)^\dagger = \mathcal{A}$.

Proof. By Equation (A3) given any system S, any state $\rho \in$ St$_1$ $(A \otimes S)$, and any effect $E \in$ Eff $(B \otimes S)$, we have

$$\begin{array}{c}\underbrace{\rho}_{\substack{A \quad \boxed{\left(\mathcal{A}^\dagger\right)^\dagger} \quad B \\ S}} \underbrace{E} = \underbrace{E^\dagger}_{\substack{B \quad \boxed{\mathcal{A}^\dagger} \quad A \\ S}} \underbrace{\rho^\dagger}. \end{array} \tag{A5}$$

A linear extension of Equation (A3) to cover the case when $E^\dagger$ is not a physical state, applied to the RHS of Equation (A5) yields

$$\underbrace{E^\dagger}_{\substack{B \quad \boxed{\mathcal{A}^\dagger} \quad A \\ S}} \underbrace{\rho^\dagger} = \underbrace{\rho}_{\substack{A \quad \boxed{\mathcal{A}} \quad B \\ S}} \underbrace{E}.$$

Comparing this with Equation (A5), we get the thesis. $\square$

We can give a characterisation of the dagger of reversible channels, which are unital channels.

Proposition A7. *If* $\mathcal{U} \in$ Transf (A, B) *is a reversible channel,* $\mathcal{U}^\dagger = \mathcal{U}^{-1}$.

Proof. We have

$$\underbrace{\rho}_{\substack{B \quad \boxed{\mathcal{U}^\dagger} \quad A \\ S}} \underbrace{E} = \underbrace{E^\dagger}_{\substack{A \quad \boxed{\mathcal{U}} \quad B \\ S}} \underbrace{\rho^\dagger},$$

for any S, ρ, E. Recalling Lemma 2, the RHS is $\langle \rho, (\mathcal{U} \otimes \mathcal{I}) E^\dagger \rangle$. By Proposition 4 $\langle \rho, (\mathcal{U} \otimes \mathcal{I}) E^\dagger \rangle = \langle (\mathcal{U}^{-1} \otimes \mathcal{I}) \rho, E^\dagger \rangle$, and by symmetry of the inner product we have that

$$\left\langle \left(\mathcal{U}^{-1} \otimes \mathcal{I}\right) \rho, E^\dagger \right\rangle = \left\langle E^\dagger, \left(\mathcal{U}^{-1} \otimes \mathcal{I}\right) \rho \right\rangle = \left(\begin{array}{c} \rho \quad \boxed{\mathcal{U}^{-1}} \quad E \end{array} \right),$$

whence the thesis follows. □

In particular we have that the dagger of the SWAP channel between two systems is the SWAP with the input and output systems reversed.

The orthogonal projectors of Section 5.2, on the other hand, are self-adjoint on single system.

Proposition A8. *Given the orthogonal projector P_l on a face F_l, we have $P_\mathsf{l}^\dagger \doteq P_\mathsf{l}$.*

Proof. For every ρ and E, we have $\left(E|P_\mathsf{l}^\dagger|\rho\right) = \left(\rho^\dagger|P_\mathsf{l}|E^\dagger\right)$. The RHS is $\left\langle \rho, P_\mathsf{l} E^\dagger \right\rangle$. By the properties of projectors,

$$\left\langle \rho, P_\mathsf{l} E^\dagger \right\rangle = \left\langle P_\mathsf{l}\rho, E^\dagger \right\rangle = \left\langle E^\dagger, P_\mathsf{l}\rho \right\rangle = \left(E|P_\mathsf{l}|\rho\right).$$

This shows that $P_\mathsf{l}^\dagger \doteq P_\mathsf{l}$. □

Finally we prove some properties of the dagger with respect to compositions. We need an easy lemma first.

Lemma A3. *For every $\mathcal{A} \in \mathsf{Transf}\,(A, B)$, every system S, and every vector $\xi \in \mathsf{St}_\mathbb{R}\,(A \otimes S)$ we have*

$$\left(\begin{array}{c} \xi \quad \boxed{\mathcal{A}} \end{array} \right)^\dagger = \begin{array}{c} \boxed{\mathcal{A}^\dagger} \quad \xi^\dagger \end{array}.$$

Proof. Recall that $\mathcal{A} = \left(\mathcal{A}^\dagger\right)^\dagger$; by Definition A3 we have $\left(\mathcal{A}^\dagger\right)^\dagger \xi = \left(\xi^\dagger \mathcal{A}^\dagger\right)^\dagger$

$$\begin{array}{c} \xi \quad \boxed{\mathcal{A}} \end{array} = \begin{array}{c} \xi \quad \boxed{\left(\mathcal{A}^\dagger\right)^\dagger} \end{array} = \left(\begin{array}{c} \boxed{\mathcal{A}^\dagger} \quad \xi^\dagger \end{array} \right)^\dagger.$$

Taking the dagger of this equation yields the desired result. □

Now we can state the main results. The first concerns sequential composition.

Proposition A9. *For all transformations $\mathcal{A} \in \mathsf{Transf}\,(A, B)$, $\mathcal{B} \in \mathsf{Transf}\,(B, C)$, one has $(\mathcal{BA})^\dagger = \mathcal{A}^\dagger \mathcal{B}^\dagger$.*
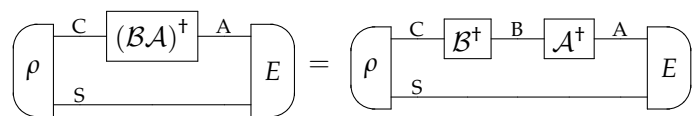
Proof. Take any system S, any state $\rho \in \mathsf{St}_1\,(C \otimes S)$, and any effect $E \in \mathsf{Eff}\,(A \otimes S)$. By Equation (A3) we have

$$\left(\begin{array}{c} \rho \quad \boxed{(\mathcal{BA})^\dagger} \quad E \end{array} \right) = \left(\begin{array}{c} E^\dagger \quad \boxed{\mathcal{BA}} \quad \rho^\dagger \end{array} \right) = \left(\begin{array}{c} E^\dagger \quad \boxed{\mathcal{A}} \quad \boxed{\mathcal{B}} \quad \rho^\dagger \end{array} \right).$$

Define ξ as $\xi := (\mathcal{A} \otimes \mathcal{I}) E^\dagger$, so

$$\left(\begin{array}{c} \rho \quad \boxed{(\mathcal{BA})^\dagger} \quad E \end{array} \right) = \left(\begin{array}{c} \xi \quad \boxed{\mathcal{B}} \quad \rho^\dagger \end{array} \right) = \left(\begin{array}{c} \rho \quad \boxed{\mathcal{B}^\dagger} \quad \xi^\dagger \end{array} \right).$$

By Lemma A3 $\xi^\dagger = \left[(\mathcal{A} \otimes \mathcal{I}) \, E^\dagger \right]^\dagger = E \left(\mathcal{A}^\dagger \otimes \mathcal{I} \right)$, then
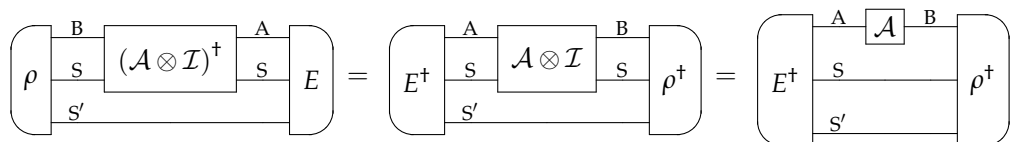
$$
\left(\rho \begin{array}{c} {}^{C}\!\!-\!\!\boxed{(\mathcal{B}\mathcal{A})^\dagger}\!\!-\!\!{}^{A} \\ {}_{S} \end{array} \!\! E \right) = \left(\rho \begin{array}{c} {}^{C}\!\!-\!\!\boxed{\mathcal{B}^\dagger}\!\!-\!\!{}^{B}\!\!-\!\!\boxed{\mathcal{A}^\dagger}\!\!-\!\!{}^{A} \\ {}_{S} \end{array} \!\! E \right),
$$

therefore $(\mathcal{B}\mathcal{A})^\dagger = \mathcal{A}^\dagger \mathcal{B}^\dagger$. $\square$

Finally the dagger respects parallel composition. Again we need a lemma.

Lemma A4. *For every* $\mathcal{A} \in \mathsf{Transf}\,(A, B)$, *every systems* S *and* S′, *we have* $(\mathcal{I}_S \otimes \mathcal{A} \otimes \mathcal{I}_{S'})^\dagger = \mathcal{I}_S \otimes \mathcal{A}^\dagger \otimes \mathcal{I}_{S'}$.

Proof. As a first step, let us prove that, for every system S, we have $(\mathcal{A} \otimes \mathcal{I}_S)^\dagger = \mathcal{A}^\dagger \otimes \mathcal{I}_S$. Take any system S′, any state $\rho \in \mathsf{St}_1 \, (B \otimes S \otimes S')$, and any effect $E \in \mathsf{Eff}\,(A \otimes S \otimes S')$, Equation (A3) yields
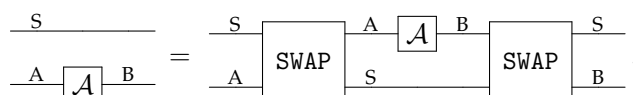
$$
\left(\rho \begin{array}{c} {}^{B} \\ {}^{S}\!\!-\!\!\boxed{(\mathcal{A}\otimes\mathcal{I})^\dagger}\!\!-\!\!{}^{A} \\ {}_{S'} \end{array} \!\! E \right) = \left(E^\dagger \begin{array}{c} {}^{A} \\ {}^{S}\!\!-\!\!\boxed{\mathcal{A}\otimes\mathcal{I}}\!\!-\!\!{}^{B} \\ {}_{S'} \end{array} \!\! \rho^\dagger \right) = \left(E^\dagger \begin{array}{c} {}^{A}\!\!-\!\!\boxed{\mathcal{A}}\!\!-\!\!{}^{B} \\ {}^{S} \\ {}_{S'} \end{array} \!\! \rho^\dagger \right).
$$

Specialising Equation (A3) to the case of a composite system, we have

$$
\left(E^\dagger \begin{array}{c} {}^{A}\!\!-\!\!\boxed{\mathcal{A}}\!\!-\!\!{}^{B} \\ {}^{S} \\ {}_{S'} \end{array} \!\! \rho^\dagger \right) = \left(\rho \begin{array}{c} {}^{B}\!\!-\!\!\boxed{\mathcal{A}^\dagger}\!\!-\!\!{}^{A} \\ {}^{S} \\ {}_{S'} \end{array} \!\! E \right),
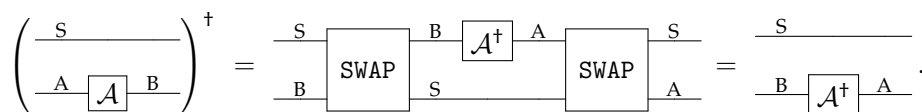$$
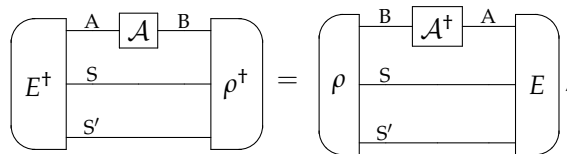
whence we conclude that $(\mathcal{A} \otimes \mathcal{I}_S)^\dagger = \mathcal{A}^\dagger \otimes \mathcal{I}_S$.

Now let us prove that, for every system S, $(\mathcal{I}_S \otimes \mathcal{A})^\dagger = \mathcal{I}_S \otimes \mathcal{A}^\dagger$. Note that

$$
\begin{array}{c} {}^{S}\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\! \\ {}^{A}\!\!-\!\!\boxed{\mathcal{A}}\!\!-\!\!{}^{B} \end{array} = \begin{array}{c} {}^{S}\!\!-\!\!\boxed{\mathrm{SWAP}}\!\!\begin{array}{c}{}^{A}\!\!-\!\!\boxed{\mathcal{A}}\!\!-\!\!{}^{B}\\{}_{S}\end{array}\!\!\boxed{\mathrm{SWAP}}\!\!-\!\!{}^{S} \\ {}^{A} \qquad\qquad\qquad {}_{B} \end{array}.
$$

By Proposition A9, and recalling what we have just proved, we have

$$
\left(\begin{array}{c} {}^{S}\!\!-\!\!-\!\!-\!\!- \\ {}^{A}\!\!-\!\!\boxed{\mathcal{A}}\!\!-\!\!{}^{B} \end{array} \right)^{\!\dagger} = \begin{array}{c} {}^{S}\!\!-\!\!\boxed{\mathrm{SWAP}}\!\!\begin{array}{c}{}^{B}\!\!-\!\!\boxed{\mathcal{A}^\dagger}\!\!-\!\!{}^{A}\\{}_{S}\end{array}\!\!\boxed{\mathrm{SWAP}}\!\!-\!\!{}^{S} \\ {}^{B} \qquad\qquad\qquad {}_{A} \end{array} = \begin{array}{c} {}^{S}\!\!-\!\!-\!\!-\!\!-\!\!- \\ {}^{B}\!\!-\!\!\boxed{\mathcal{A}^\dagger}\!\!-\!\!{}^{A} \end{array}.
$$

To get the thesis, note that $(\mathcal{I}_S \otimes \mathcal{A} \otimes \mathcal{I}_{S'})^\dagger = [(\mathcal{I}_S \otimes \mathcal{A}) \otimes \mathcal{I}_{S'}]^\dagger$. We have just proved that

$$
[(\mathcal{I}_S \otimes \mathcal{A}) \otimes \mathcal{I}_{S'}]^\dagger = (\mathcal{I}_S \otimes \mathcal{A})^\dagger \otimes \mathcal{I}_{S'},
$$

and that $(\mathcal{I}_S \otimes \mathcal{A})^\dagger = \mathcal{I}_S \otimes \mathcal{A}^\dagger$, therefore we conclude that $(\mathcal{I}_S \otimes \mathcal{A} \otimes \mathcal{I}_{S'})^\dagger = \mathcal{I}_S \otimes \mathcal{A}^\dagger \otimes \mathcal{I}_{S'}$. $\square$

Proposition A10. *Let* $\mathcal{A} \in \mathsf{Transf}\,(A, B)$, *and* $\mathcal{B} \in \mathsf{Transf}\,(C, D)$. *We have* $(\mathcal{A} \otimes \mathcal{B})^\dagger = \mathcal{A}^\dagger \otimes \mathcal{B}^\dagger$.

Proof. Take any system S, any state $\rho \in \mathsf{St}_1\,(B \otimes D \otimes S)$, and any effect $E \in \mathsf{Eff}\,(A \otimes C \otimes S)$, we have

$$\rho \;\boxed{(\mathcal{A} \otimes \mathcal{B})^\dagger}\; E \;=\; E^\dagger \;\boxed{\mathcal{A} \otimes \mathcal{B}}\; \rho^\dagger \;=\; E^\dagger \;\boxed{\substack{\mathcal{A}\\ \mathcal{B}}}\; \rho^\dagger .$$

Now define $\xi := (\mathcal{I}_\mathrm{A} \otimes \mathcal{B} \otimes \mathcal{I}_\mathrm{S})\, E^\dagger$, hence

$$\rho \;\boxed{(\mathcal{A} \otimes \mathcal{B})^\dagger}\; E \;=\; \xi \;\boxed{\mathcal{A}}\; \rho^\dagger \;=\; \rho \;\boxed{\mathcal{A}^\dagger}\; \xi^\dagger .$$

By Lemmas A3 and A4, we have that $\xi^\dagger = E\left(\mathcal{I}_\mathrm{A} \otimes \mathcal{B}^\dagger \otimes \mathcal{I}_\mathrm{S}\right)$, so

$$\rho \;\boxed{(\mathcal{A} \otimes \mathcal{B})^\dagger}\; E \;=\; \rho \;\boxed{\substack{\mathcal{A}^\dagger\\ \mathcal{B}^\dagger}}\; E ,$$

whence the thesis. $\quad\square$

This means that the dagger respects the composition of diagrams, and corresponds to the action of flipping a diagram with respect to a vertical axis.

References

1. Feynman, R.P.; Leighton, R.; Sands, M. *The Feynman Lectures on Physics. The Definitive and Extended Edition*; Addison Wesley: Boston, MA, USA, 2005.
2. Sorkin, R.D. Quantum mechanics as quantum measure theory. *Mod. Phys. Lett. A* **1994**, *9*, 3119–3127.
3. Sorkin, R.D. *Quantum Classical Correspondence: The 4th Drexel Symposium on Quantum Nonintegrability*; Chapter Quantum Measure Theory and Its Interpretation; International Press: Boston, MA, USA, 1997; pp. 229–251.
4. Barnum, H.; Müller, M.P.; Ududec, C. Higher-order interference and single-system postulates characterizing quantum theory. *New J. Phys.* **2014**, *16*, 123029.
5. Bolotin, A. On the ongoing experiments looking for higher-order interference: What are they really testing? *arXiv* **2016**, arXiv:1611.06461.
6. Dakić, B.; Paterek, T.; Brukner, Č. Density cubes and higher-order interference theories. *New J. Phys.* **2014**, *16*, 023028.
7. Lee, C.M.; Selby, J.H. Deriving grover's lower bound from simple physical principles. *New J. Phys.* **2016**, *18*, 093047.
8. Lee, C.M.; Selby, J.H. Generalised phase kick-back: The structure of computational algorithms from physical principles. *New J. Phys.* **2016**, *18*, 033023.
9. Lee, C.M.; Selby, J.H. Higher-order interference in extensions of quantum theory. *Found. Phys.* **2017**, *47*, 89–112.
10. Niestegge, G. Three-slit experiments and quantum nonlocality. *Found. Phys.* **2013**, *43*, 805–812.
11. Ududec, C. Perspectives on the Formalism of Quantum Theory. Ph.D. Thesis, University of Waterloo, Waterloo, ON, Canada, 2012.
12. Ududec, C.; Barnum, H.; Emerson, J. Probabilistic Interference in Operational Models. 2009, in preparation.
13. Ududec, C.; Barnum, H.; Emerson, J. Three slit experiments and the structure of quantum theory. *Found. Phys.* **2011**, *41*, 396–405.
14. Lee, C.M.; Selby, J.H. A no-go theorem for theories that decohere to quantum mechanics. *arXiv* **2017**, arXiv:1701.07449.

15. Barnum, H.; Barrett, J.; Leifer, M.; Wilce, A. Generalized no-broadcasting theorem. *Phys. Rev. Lett.* **2007**, *99*, 240501.

16. Barnum, H.; Wilce, A. Information processing in convex operational theories. *Electron. Notes Theor. Comput. Sci.* **2011**, *270*, 3–15.

17. Barrett, J. Information processing in generalized probabilistic theories. *Phys. Rev. A* **2007**, *75*, 032304.

18. Barrett, J.; de Beaudrap, N.; Hoban, M.J.; Lee, C.M. The computational landscape of general physical theories. *arXiv* **2017**, arXiv:1702.08483.

19. Chiribella, G.; D'Ariano, G.M.; Perinotti, P. Probabilistic theories with purification. *Phys. Rev. A* **2010**, *81*, 062348.

20. Chiribella, G.; D'Ariano, G.M.; Perinotti, P. Informational derivation of quantum theory. *Phys. Rev. A* **2011**, *84*, 012311.

21. Chiribella, G.; Spekkens, R.W. (Eds.) *Quantum Theory: Informational Foundations and Foils*; Fundamental Theories of Physics; Springer: Dordrecht, The Netherlands, 2016; Volume 181.

22. Dakić, B.; Brukner, Č. *Quantum Theory and Beyond: Is Entanglement Special*; Cambridge University Press: Cambridge, UK, 2011; pp. 365–392.

23. Hardy, L. Quantum Theory From Five Reasonable Axioms. *arXiv* **2001**, arXiv:quant-ph/0101012.

24. Hardy, L. *Foliable Operational Structures for General Probabilistic Theories*; Cambridge University Press: Cambridge, UK, 2011; pp. 409–442.

25. Lee, C.M.; Barrett, J. Computation in generalised probabilistic theories. *New J. Phys.* **2015**, *17*, 083001.

26. Lee, C.M.; Hoban, M.J. Bounds on the power of proofs and advice in general physical theories. *Proc. R. Soc. A* **2016**, *472*, 20160076.

27. Lee, C.M.; Hoban, M.J. The information content of systems in general physical theories. In Proceedings of the 7th International Workshop on Physics and Computation, Manchester, UK, 14 July 2016; Volume 214, pp. 22–28.

28. Masanes, L.; Müller, M.P. A derivation of quantum theory from physical requirements. *New J. Phys.* **2011**, *13*, 063001.

29. Hardy, L. Reformulating and reconstructing quantum theory. *arXiv* **2011**, arXiv:1104.2066.

30. Chiribella, G.; Scandolo, C.M. Entanglement as an axiomatic foundation for statistical mechanics. *arXiv* **2016**, arXiv:1608.04459.

31. Krumm, M.; Barnum, H.; Barrett, J.; Müller, M.P. Thermodynamics and the structure of quantum theory. *New J. Phys.* **2017**, *19*, 043025.

32. Jin, F.; Liu, Y.; Geng, J.; Huang, P.; Ma, W.; Shi, M.; Duan, C.; Shi, F.; Rong, X.; Du, J. Experimental test of born's rule by inspecting third-order quantum interference on a single spin in solids. *Phys. Rev. A* **2017**, *95*, 012107.

33. Kauten, T.; Keil, R.; Kaufmann, T.; Pressl, B.; Brukner, Č.; Weihs, G. Obtaining tight bounds on higher-order interferences with a 5-path interferometer. *New J. Phys.* **2017**, *19*, 033017.

34. Park, D.K.; Moussa, O.; Laflamme, R. Three path interference using nuclear magnetic resonance: A test of the consistency of born's rule. *New J. Phys.* **2012**, *14*, 113025.

35. Sinha, A.; Vijay, A.H.; Sinha, U. On the superposition principle in interference experiments. *Sci. Rep.* **2015**, *5*, 10304.

36. Sinha, U.; Couteau, C.; Jennewein, T.; Laflamme, R.; Weihs, G. Ruling out multi-order interference in quantum mechanics. *Science* **2010**, *329*, 418–421.

37. Barnum, H.; Graydon, M.; Wilce, A. Composites and categories of Euclidean Jordan algebras. *arXiv* **2016**, arXiv:1606.09331.

38. Chiribella, G. Dilation of states and processes in operational-probabilistic theories. In Proceedings of the 11th workshop on Quantum Physics and Logic, Kyoto, Japan, 4–6 June 2014; Volume 172, pp. 1–14.

39. Chiribella, G.; D'Ariano, G.M.; Perinotti, P. *Quantum Theory: Informational Foundations and Foils*; Chapter Quantum from Principles; Springer: Dordrecht, The Netherlands, 2016; pp. 171–221.

40. Hardy, L. *Quantum Theory: Informational Foundations and Foils*; Chapter Reconstructing Quantum Theory; Springer: Dordrecht, The Netherlands, 2016; pp. 223–248.

41. Abramsky, S.; Coecke, B. A categorical semantics of quantum protocols. In Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science, Turku, Finland, 13–17 July 2004; pp. 415–425.

42. Coecke, B. Kindergarten quantum mechanics: Lecture notes. *AIP Conf. Proc.* **2006**, *810*, 81–98.

43. Coecke, B. Quantum picturalism. *Contemp. Phys.* **2010**, *51*, 59.

44. Coecke, B.; Duncan, R.; Kissinger, A.; Wang, Q. *Quantum Theory: Informational Foundations and Foils*; Chapter Generalised Compositional Theories and Diagrammatic Reasoning; Springer: Dordrecht, The Netherlands, 2016; pp. 309–366.

45. Coecke, B.; Kissinger, A. *Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning*; Cambridge University Press: Cambridge, UK, 2017.

46. Selinger, P. A survey of graphical languages for monoidal categories. In *New Structures for Physics*; Coecke, B., Ed.; Springer: Berlin, Germany, 2011; pp. 289–356.

47. Wootters, W.K. Local accessibility of quantum states. In *Complexity, Entropy and the Physics of Information*; Zurek, W.H., Ed.; Westview Press: Boulder, CO, USA, 1990; pp. 39–46.

48. Chiribella, G.; Scandolo, C.M. Entanglement and thermodynamics in general probabilistic theories. *New J. Phys.* **2015**, *17*, 103027.

49. Chiribella, G.; Scandolo, C.M. Purity in microcanonical thermodynamics: A tale of three resource theories. *arXiv* **2016**, arXiv:1608.0446.

50. Gour, G.; Müller, M.P.; Narasimhachar, V.; Spekkens, R.W.; Yunger Halpern, N. The resource theory of informational nonequilibrium in thermodynamics. *Phys. Rep.* **2015**, *583*, 1–58.

51. Horodecki, M.; Horodecki, P.; Oppenheim, J. Reversible transformations from pure to mixed states and the unique measure of information. *Phys. Rev. A* **2003**, *67*, 062104.

52. Selby, J.H.; Coecke, B. Leaks: Quantum, classical, intermediate, and more. *Entropy* **2017**, *19*, 174.

53. Coecke, B. Terminality implies non-signalling. In Proceedings of the 11th workshop on Quantum Physics and Logic, Kyoto, Japan, 4–6 June 2014; Volume 172, pp. 27–35.

54. Chiribella, G.; Scandolo, C.M. Operational axioms for diagonalizing states. In Proceedings of the 12th International Workshop on Quantum Physics and Logic, Oxford, UK, 15–17 July 2015; Volume 195, pp. 96–115.

55. Chiribella, G.; Scandolo, C.M. Conservation of information and the foundations of quantum mechanics. *EPJ Web Conf.* **2015**, *95*, 03003.

56. Disilvestro, L.; Markham, D. Quantum protocols within Spekkens' toy model. *Phys. Rev. A* **2017**, *95*, 052324.

57. D'Ariano, G.M.; Manessi, F.; Perinotti, P.; Tosini, A. Fermionic computation is non-local tomographic and violates monogamy of entanglement. *Europhys. Lett.* **2014**, *107*, 20009.

58. D'Ariano, G.M.; Manessi, F.; Perinotti, P.; Tosini, A. The Feynman problem and fermionic entanglement: Fermionic theory versus qubit theory. *Int. J. Mod. Phys. A* **2014**, *29*, 1430025.

59. Chiribella, G.; Yuan, X. Bridging the gap between general probabilistic theories and the device-independent framework for nonlocality and contextuality. *Inf. Comput.* **2016**, *250*, 15–49.

60. Pfister, C.; Wehner, S. An information-theoretic principle implies that any discrete physical theory is classical. *Nat. Commun.* **2013**, *4*, 1851.

61. Alfsen, E.M.; Shultz, F.W. *Geometry of State Spaces of Operator Algebras*; Mathematics Theory & Applications; Birkhäuser: Basel, Switzerland, 2003.

62. Barnum, H.; Barrett, J.; Krumm, M.; Müller, M.P. Entropy, majorization and thermodynamics in general probabilistic theories. In Proceedings of the 12th International Workshop on Quantum Physics and Logic, Oxford, UK, 15–17 July 2015; Volume 195, pp. 43–58.

63. Chiribella, G.; Yuan, X. Measurement sharpness cuts nonlocality and contextuality in every physical theory. *arXiv* **2014**, arXiv:1404.3348.

64. Alfsen, E.M.; Shultz, F.W. State spaces of Jordan algebras. *Acta Math.* **1978**, *140.1*, 155–190.

65. Iochum, B. *Cônes Autopolaires et Algèbres de Jordan*; Lecture Notes in Mathematics; Springer: Berlin/Heidelberg, Germany, 1984; Volume 1049, doi: 10.1007/BFb0071358. (In French)

66. Coecke, B.; Selby, J.; Tull, S. Two roads to classicality. *arXiv* **2017**, arXiv:1701.07400.

67. Barnum, H. Spectrality as a Tool for Quantum Reconstruction: Higher-Order Interference, Jordan State Space Characterizations, Aug. 2009. Talk Given at the Conference "Reconstructing Quantum Theory", August 9–11, Perimeter Institute for Theoretical Physics. Available online: http://pirsa.org/09080016/ (accessed on 26 May 2017).

68. Niestegge, G. Conditional probability, three-slit experiments, and the jordan algebra structure of quantum mechanics. *Adv. Math. Phys.* **2012**, *2012*, 156573.

69. Selby, J.H.; Coecke, B. Process-theoretic characterisation of the hermitian adjoint. *arXiv* **2016**, arXiv:1606.05086.

70. Müller, M.P.; Oppenheim, J.; Dahlsten, O.C.O. The black hole information problem beyond quantum theory. *J. High Energy Phys.* **2012**, *2012*, 9.
71. Müller, M.P.; Ududec, C. Structure of reversible computation determines the self-duality of quantum theory. *Phys. Rev. Lett.* **2012**, *108*, 130401.
72. Marshall, A.W.; Olkin, I.; Arnold, B.C. *Inequalities: Theory of Majorization and Its Applications*; Springer Series in Statistics; Springer: New York, NY, USA, 2011.
73. Müller, M.P.; Dahlsten, O.C.O.; Vedral, V. Unifying typical entanglement and coin tossing: On randomization in probabilistic theories. *Commun. Math. Phys.* **2012**, *316*, 441–487.
74. Wilde, M.M. *Quantum Information Theory*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2017.
75. Selinger, P. Dagger compact closed categories and completely positive maps. *Electron. Notes Theor. Comput. Sci.* **2007**, *170*, 139–163.