Available online at www.sciencedirect.com
ScienceDirect
Electronic Notes in Theoretical Computer Science 270 (1) (2011) 3–15
Electronic Notes in Theoretical Computer Science
www.elsevier.com/locate/entcs

Information Processing in Convex Operational Theories

Howard Barnum[1]

*CCS-3: Information Sciences; Quantum Institute; and Center for Nonlinear Studies
Los Alamos National Laboratory
Los Alamos, NM, USA*

Alexander Wilce[1]

*Department of Mathematical Sciences
Susquehanna University
Selinsgrove, PA, USA*

Abstract

In order to understand the source and extent of the greater-than-classical information processing power of quantum systems, one wants to characterize both classical and quantum mechanics as points in a broader space of possible theories. One approach to doing this, pioneered by Abramsky and Coecke, is to abstract the essential categorical features of classical and quantum mechanics that support various information-theoretic constraints and possibilities, e.g., the impossibility of cloning in the latter, and the possibility of teleportation in both. Another approach, pursued by the authors and various collaborators, is to begin with a very conservative, and in a sense very concrete, generalization of classical probability theory – which is still sufficient to encompass quantum theory – and to ask which "quantum" informational phenomena can be reproduced in this much looser setting. In this paper, we review the progress to date in this second programme, and offer some suggestions as to how to link it with the categorical semantics for quantum processes offered by Abramsky and Coecke.

Keywords: Ordered linear spaces, convex sets, operational theories, categories, enriched categories, quantum theory, quantum mechanics, information processing, bit commitment, teleportation

1 Introduction

The advent of quantum information theory has been accompanied by a resurgence of interest in the convex (or ordered linear spaces) framework for operational theories, as researchers seek to understand the nature of information processing in

[1] Thanks to Samson Abramsky, John Armstrong, Rick Blute and Bob Coecke for useful discussions about categories, linear logic, and information processing

[2] Email: barnum@lanl.gov

[3] Email: wilce@susqu.edu

doi:10.1016/j.entcs.2011.01.002

increasingly abstract terms, both in order to illuminate the sources of the difference between the information processing power of quantum versus classical theory, and because quantum information has occasioned renewed interest in foundational aspects of quantum theory, often with the new twist that axioms or principles concerning information processing are considered. A representative (but by no means exhaustive) sample of work in this vein might include the work of Hardy [14,15], D'Ariano [12], and Barrett [9].

At the same time, a fascinating and illuminating categorical approach to the formulation of quantum physics has crystallized around the notions of compact closed and dagger compact closed categories that exhibit key features of quantum theory, but allow many other models as well. The main work along these lines has been done by Abramsky and Coecke [2], by Selinger [19,20], and by Baez [3].

In [2], Abramsky and Coecke established that many of the most striking phenomena associated with quantum information processing – notably, various forms of teleportation – arise much more generally in any compact closed category, including, for instance, the category of sets and relations. An important observation here is that *the unit and co-unit defining a dual object in such a category can be interpreted as a teleportation protocol.* On the other hand, working in the much more concrete but structurally much looser convex framework (in which essentially arbitrary compact convex sets serve as abstract state spaces), our coauthors (Jonathan Barrett and Matthew Leifer) and we have shown ([3]-[5]) that many of the same phenomena – in particular, many aspects of entanglement, as well as no-cloning and no-broadcasting theorems – are quite generic features of probabilistic models. In this framework, the existence of a teleportation protocol is a nontrivial constraint, moving one somewhat closer to quantum theory; but even so, one can construct many models of teleportation – and even of deterministic teleportation – that are neither classical nor quantum. An important observation here is that a teleportation protocol is *just a special case of conditioning.*

This paper reviews work by ourselves and various collaborators, especially Jon Barrett, Matt Leifer, Oscar Dahlsten, Leifer, and Ben Toner, on information processing in the ordered linear spaces framework, and then proceeds to discuss how this work may be related to broad project of describing information-processing using categories of processes. The work reviewed shows that certain information-processing properties which had sometimes been taken to be "peculiarly quantum," are actually common to all nonclassical theories in the framework. These include the existence of information about states which cannot be obtained without disturbing them, and generalizations of the quantum no-cloning and no-broadcasting theorems.

The impossibility of bit commitment has been suggested (for example by Brassard [11] and by Fuchs [13]) as a potential fundamental information-processing principle, shared by classical and quantum mechanics, that might, in combination with other principles, characterize quantum mechanics. The other principles proposed by Brassard and by Fuchs are the possibility of secure secret key distribution, which is intimately connected with no-cloning, no-broadcasting, and information-disturbance tradeoffs and which, as we shall see, rules out classical theory, and the impossibility of instantaneous signaling between systems (which is built into the notion of composite system used in our version of the ordered linear spaces framework). We will also present some results on bit commitment in our framework, to the effect that *all* nonclassical theories that lack entanglement permit exponentially secure bit commitment, and some results on how the presence in a theory of certain kinds of entangled states can defeat the bit commitment protocol we used for the unentangled case. Closely related states can permit teleportation, another information-processing task whose possibility helps distinguish between classes of nonclassical theories in our framework. We summarize some of our recent work with Barrett and Leifer on multipartite composite systems and teleportation in the ordered linear spaces framework. In particular, we report necessary and sufficient conditions for a composite of three systems to support a conclusive teleportation protocol, and interesting sufficient conditions for deterministic teleportation.

We then make some first steps towards a category-theoretic formulation of our results. The abstract state spaces that we consider naturally form a category; however, this is far from being compact closed. For one thing, the dual of an abstract state space is usually not, in any natural way, another state space, but a different sort of beast altogether. Nor are our categories generally monoidal: more typically, they support a profusion of possible mechanisms for coupling systems, bounded by a maximal (and maximally entangled) tensor product $\otimes_{max}$, and a minimal (unentangled) product $\otimes_{min}$. On the other hand, there are various constructions by which one can embed our category of state spaces in a larger category of *processes* having a better behaved – in particular, monoidal and self-dual – structure. Moving in the opposite direction, one can focus on restricted categories that are, in a sense (made precise below) “closed under teleportation”: as it happens, the entangled state and effect corresponding to a correction-free teleportation protocol are precisely the unit and co-unit of a duality.

Rather than building categories of processes from (categories of) abstract state spaces, one might start from the opposite direction, by treating categories of processes axiomatically. A natural starting point for such an effort would be to consider categories *enriched* over ordered linear spaces, or over abstract state spaces. This paper ends with a sketch of this idea. We envision such categorical formulations as a first step toward comparing the necessary and/or sufficient conditions for various information processing protocols or informational properties of theories, obtained in the convex framework, with properties such as compact closure, dagger compact closure, non-cartesianity and so forth that have been used in the literature on categorical descriptions of information processing. This section of the paper has benefited from discussions with Abramsky, Armstrong, Coecke, and others and may be viewed as describing work early-in-progress in collaboration with at least some of them.

2 Abstract State Spaces

By an *abstract state space*, we mean a pair (A, u_A) where A is a finite-dimensional ordered real vector space, with positive cone A_+, and where $u_A : A \to \mathbb{R}$ is a distinguished linear functional, called the *order unit*, that is *strictly* positive on $A_+ \setminus \{0\}$. A state is *normalized* iff $u_A(\alpha) = 1$. We write Ω_A for the convex set of normalized states in A_+. By way of illustration, if A is the space $\mathbb{R}^X$ of real-valued functions on a set X, ordered pointwise on X, with $u_A(f) = \sum_{x \in X} f(x)$, then $\Omega_A = \Delta(X)$, the simplex of probability weights on X. If A is the space $\mathcal{L}(\mathcal{H})$ of hermitian operators on a (finite-dimensional) complex Hilbert space $\mathcal{H}$, with the usual operator ordering, and if $u_A(a) = \mathrm{Tr}(a)$, then Ω_A is the set of density operators on $\mathcal{H}$. On any abstract state space A, there is a canonical norm (the *base norm*) such that for $\alpha \in A_+$, $\|\alpha\| = u_A(\alpha)$. For $\mathbb{R}^X$, this is just the norm on X; for $\mathcal{L}(\mathcal{H})$, it is the trace norm.

Events (e.g., measurement outcomes) associated with an abstract state space A are represented by *effects*, i.e., positive linear functionals $a \in A^*$, with $0 \leq a \leq u_A$ in the dual ordering. Note that 0 and u_A are, by definition, the least and greatest effects. If α is a *normalized* state in A – that is, if $u_A(\alpha) = 1$ – then we interpret $a(\alpha)$ as the *probability* that the event represented by the effect a will occur if measured. Accordingly, a discrete *observable* on A is a list $(a_1, ..., a_n)$ of effects with $a_1 + a_2 + \cdots + a_n = u_A$. We represent a physical process with initial state space A and final state space B by a positive mapping $\tau : A \to B$ such that, for all $\alpha \in A_+$, $u_B(\tau(\alpha)) \leq u_A(\alpha)$ – equivalently, τ is norm-contractive. We can regard $\|\tau(\alpha)\| = u_B(\tau(\alpha))$ as the probability that the process represented by τ takes place in initial state α; this *event* is represented by the effect $u_B \circ \tau$ on A.

It is important to note that, in the framework just outlined, the state space A and its dual space A^* have (in general) quite different structures: A is a *cone-base space*, i.e., an ordered space with a preferred base, Ω_A, for A_+, while A^* is an *order-unit* space, i.e., an ordered space with a preferred *element* in its positive cone. Indeed, the spaces A and A^* are generally not even isomorphic as ordered spaces. Where there exists a linear order-isomorphism (that is, a positive linear mapping with positive inverse) between A and A^*, we shall say that A is *weakly self-dual.* Where this isomorphism induces an inner product on A such that $A_+ = \{b \in A | \langle b, a \rangle \geq 0 \ \forall a \in A_+\}$, we say that A is *self-dual.* Finite dimensional quantum and classical state spaces are self-dual in this sense. A celebrated theorem of Vinberg and Koecher [8,9] tells us that if A is an irreducible, finite-dimensional self-dual state space, and if the group of affine automorphisms of A_+ acts transitively on the interior of A_+, then the space Ω_A of normalized states is affinely isomorphic to the set of density operators on an n-dimensional Hilbert space, or to a ball, or to the set of 3×3 trace-one positive matrices over the octonions.

3 Composite Systems

For our purposes, it will be convenient to identify the tensor product, $A \otimes B$, of two state spaces with the space $B(A^*, B^*)$ of bilinear forms on $A^* \times B^*$, interpreting the pure tensor $\alpha\otimes\beta$ of states $\alpha \in A, \beta \in B$ as the form given by $(\alpha\otimes\beta)(f, g) = f(\alpha)g(\beta)$ where $f \in A^*, g \in B^*$. We call a form $\omega \in A \otimes B$ *positive* iff $\omega(a, b) \geq 0$ for all $(a, b) \in A^*_+ \times B^*_+$. If ω is positive and $\omega(u_A, u_B) = 1$, then $\omega(a, b)$ can be interpreted as a joint probability for effects $a \in A^*$ and $b \in B^*$. Conversely, one can show (see [4] and [10]) that any assignment of joint probabilities consistent with a no-signalling requirement must be bilinear. Thus, the most general model of a composite of A and B consistent with such a requirement, is the space $A \otimes B$, ordered by the cone of all positive forms, and with order unit given by $u_A \otimes u_B : \omega \mapsto \omega(u_A, u_B)$. This gives us an abstract state space, which we term the *maximal tensor product* of A and B, and denote $A \otimes_{max} B$. At the other extreme, we might wish to allow only product states $\alpha \otimes \beta$, and mixtures of these, to count as bipartite (normalized) states. This gives us the *minimal tensor product*, $A \otimes_{min} B$. These coincide if A and B are classical – that is, if Ω_A and Ω_B are simplices [11]; in general, however, the maximal tensor product allows many more states than the minimal. A state in $\Omega_{A\otimes_{max}B}$ not belonging to $\Omega_{A\otimes_{min}B}$ is *entangled.*

More generally, we define a *composite* of A and B to be *any* state space AB consisting of bilinear forms on $A^* \times B^*$, ordered by a cone AB_+ of positive forms containing every product state $\alpha \otimes \beta$, where $\alpha \in \Omega_A$ and $\beta \in \Omega_B$ – equivalently, AB is a composite iff $A \otimes_{min} B \leq AB \leq A \otimes_{max} B$ (where, for abstract state spaces A and B, $A \leq B$ means that A is a subspace of B, that $A_+ \subseteq B_+$, and that u_A is the restriction of u_B to A.) More generally still, a composite of n state spaces $A_1, ..., A_n$ is a state space A of n-linear forms on $A_1^* \times \cdots \times A_n^*$, ordered by any cone of positive forms containing all product states.

4 Information-disturbance tradeoffs

With Barrett and Leifer, we have shown (as described in [9]) that in nonclassical theories, the only information that can be obtained about the state without disturbing it is inherently classical information—information about which of a set of irreducible direct summands of the state cone the state lies in. Call a positive map $T : A \to A$ *nondisturbing* on state ω if $T(\omega) = c_\omega \omega$ for some positive constant c_ω that in principle could depend on the state. Say such a map is *nondisturbing* if it is nondisturbing on all pure states.[4] Nevertheless, a norm-nonincreasing map nondisturbing in this sense is precisely the type of map that can appear associated with some measurement outcome in an operation that, averaged over measurement outcomes, leaves the state unchanged.

A cone C in a vector space V is a *direct sum* of cones D and E if D and E span disjoint (except for 0) subspaces of V, and every element of C is a positive combination of vectors in D and E. A cone is irreducible if if it is not a nontrivial direct sum of cones. Every finite-dimensional cone is uniquely expressible as a direct sum $C = \oplus_i C_i$ of irreducible cones C_i. Information about which of the summands

[4] Of course if we condition on information obtained, this definition permits mixed states to be disturbed by a nondisturbing map—that can be viewed as something like an inevitable "epistemic" disturbance associated with obtaining information.

a state is in should be thought of as “inherently classical” information about the state.

Theorem 4.1 *The nondisturbing maps on a cone that is a sum $C = \oplus_i C_i$ of irreducible C_i, are precisely the maps $M = \sum_i c_i \mathrm{id}_i$, where id_i is the identity operator on the summand V_i and the zero operator elsewhere, and c_i are arbitrary nonnegative constants.*

So for a nondisturbing map, c_ω can depend only on the irreducible component a state is in. That is, the fact that a nondisturbing map has occurred can give us no information about the state within an irreducible component: in other words, as claimed, only inherently classical information is contained in the fact that a nondisturbing map has occurred.

The existence of information that cannot be obtained without disturbance is often taken to be the principle underlying the possibility of quantum key distribution, so the fact that it is generic in nonclassical theories in the framework leads us (with Barrett and Leifer) to conjecture that secure key distribution, given an authenticated public channel, is possible in all nonclassical models.

4.1 No-cloning and no-broadcasting theorems

The security of quantum key distribution is also often ascribed to the quantum no-cloning or no-broadcasting theorem—certainly no-cloning is at least *necessary* for security. A map $T : A \to A \otimes A$ *clones* a state ω if $T(\omega) = \omega \otimes \omega$. A *set* S of normalized states can be (deterministically) cloned if there is a single dynamically allowed map T that clones every $\omega \in S$.

No-cloning can be closely related to the information-disturbance principle, by an argument introduced in the quantum context but that generalizes to our setting, since if two non-identical states in the same irreducible component of a cone could be cloned, we could—by, for instance, doing an informationally complete measurement on the clone—obtain information about which state we have without disturbing it, contradicting our information-disturbance theorem.

In quantum mechanics, only orthogonal sets of states—sets S such that for all pairs $\rho, \sigma \in S$, $\rho\sigma = 0$—can be cloned [6]. As a special case of this, in a classical probability theory with a finite sample space, properly mixed states (distributions) cannot be cloned. Because of this, and because it is natural to consider *commuting* rather than mutually orthogonal sets of density matrices to be “classical subsets” of the quantum states of a system, [6] introduced the notion of *broadcasting* in order to better pick out classical subsets of the state spaces of quantum systems. A map $T : A \to A \otimes A$ *broadcasts* a state ω if both marginals of $T(\omega)$ are equal to ω; thus this notion allows correlation, or even entanglement, in the broadcast state. This is to be contrasted with the mixed-state extension of the notion of cloning for which we used the term “cloning” above, which produces a product state. Of course, the notion of broadcasting also extends, in a different way, the notion of cloning pure states, since it reduces to cloning on pure states. A set S of states is *broadcastable* if there is a norm-preserving dynamical map T that broadcasts all the states in S

(i.e. the *same* map broadcasts all the states).

The *no-broadcasting theorem* [6] asserts that it is precisely the mutually commuting sets of quantum states that can be broadcast using completely positive maps. Recently, with Barrett and Leifer we have shown [4].

Theorem 4.2 *In an arbitrary convex operational theory in our framework, a set $S \subseteq \Omega$ of states is broadcastable if, and only if, it is contained in a simplex $\Delta \subseteq \Omega$ whose vertices are distinguishable by a single measurement. For each positive map $B : V \to V \otimes V$, the set of states it broadcasts is* precisely *such a simplex.*

This combines Theorems 2 and 3 of [4]. It can be interpreted as saying that broadcastable sets of states are classical sets of states—but the sort of classicality involved is different from the *inherent* classicality of the information that can be obtained without disturbance.

The proof of the theorem uses a generalized no-cloning theorem, also proved in [4], to the effect that a set of states is clonable if, and only if, the states in it are all distinguishable from each other simultaneously via a one-shot measurement. Given this, proving Theorem 4.2 reduced to proving that a broadcastable set of states is contained in (and the states broadcast by B are *precisely*) the convex hull (necessarily a simplex) of a clonable set of states. The proof of the no-cloning result is essentially to show that if one can clone a set of states, one can distinguish them by repeatedly cloning to create many independent copies, performing an informationally complete measurement on each copy, and using the statistics of the measurement results to identify the state. Conversely if one can distinguish the states, one can clone them using a map that, conditional on distinguishing state ω, prepares $\omega \otimes \omega$. More precisely: for any ω there is a norm-nonincreasing positive map $Prep_\omega$ that prepares ω, i.e. outputs ω no matter what normalized state goes in. The cloning map is $\sum_i Prep_{\omega_i \otimes \omega_i} \circ T_i$, where $\{T_i\}$ are a set of maps such that the effects $u \circ T_i$ are a measurement distinguishing the $\rho_i \in S$; such a measurement must exist by our assumption the ρ_i were one-shot distinguishable, and we assumed as part of our general framework that every effect has at least one associated map T_i. One immediately sees that this map that clones the ω_i will also broadcast any state in the convex hull of the ω_i, giving us the easy direction of the generalized no-broadcasting theorem.

5 Nonuniqueness of extremal decomposition, and bit commitment in unentangled theories

Quantum theory has mixed states whose representation as a convex combination of pure states is not unique. So do *all* nonclassical theories: uniqueness of the decomposition of mixed states into pure states is an easy characterization–sometimes used as a definition—of simplices (see, for example, the proof in [8]). While we are not aware of any quantum information processing task whose possibility is directly traced to the non-unique decomposability of mixed states into pure, this was certainly proposed as a possible basis for quantum bit commitment schemes, though (as shown in [10] for their proposed scheme, and in [18,17] for more elaborate schemes) these schemes do not work because of entanglement.

In [7] it is shown that the existence of bit commitment protocols is universal in nonclassical theories, provided that the tensor products used do not permit entanglement. Consider a theory generated by a finite set Σ of "elementary" systems modeled by finite-dimensional abstract state spaces, containing at least one non-classical system, and closed under the minimal, or separable, tensor product, which we write with the ordinary tensor product symbol $\otimes$.

The protocol. Let a system have a non-simplicial, convex, compact state space Ω of dimension d, embedded as the base of a cone of unnormalized states in a vector space V of dimension $d + 1$. The protocol uses a state μ that has two distinct decompositions into finite disjoint sets $\{\mu_i^0\}, \{\mu_j^1\}$ of exposed states, that is,

$$\omega = \sum_{i=1}^{N_0} p_i^0 \mu_i^0 = \sum_{j=1}^{N_1} p_j^1 \mu_j^1, \tag{1}$$

A state μ_i^b is *exposed* if there is a measurement outcome a_i^b that has probability 1 when, and only when, the state is μ_i^b; the protocol exists for all nonclassical systems because, as we show, any non-simplicial convex set of affine dimension d always has a state ω with two decompositions (as above), into *disjoint* set of states whose total number $N_0 + N_1$ is $d+1$ (the disjointness and the bound on cardinality are used in the proof of exponential security).

In the honest protocol, Alice first decides on a bit $b \in \{0, 1\}$ to commit to. She then draws n samples from p^b, obtaining a string $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. To commit, she sends the state $\mu_{\mathbf{x}}^b = \mu_{x_1}^b \otimes \mu_{x_2}^b \otimes \ldots \otimes \mu_{x_n}^b$ to Bob. To reveal the bit, she sends b and $\mathbf{x}$ to Bob. Bob measures each subsystem of the state he has. On the k-th subsystem, he performs a measurement, (which will depend on b) containing the distinguishing effect for $\mu_{x_k}^b$ and rejects if the result is not the distinguishing effect. If he obtains the appropriate distinguishing effect for every system, he accepts. The protocol is perfectly sound (if Alice is honest, Bob never accuses her of cheating and always obtains the correct bit), perfectly hiding (if Alice is honest, Bob cannot gain any information about the bit until Alice reveals it), and has an exponentially low probability of Alice's successfully cheating.

6 Conditioning and teleportation protocols

If AB is a composite of state spaces A and B, we can define, for any normalized state $\omega \in AB_+$ and any effect $a \in A$, both a *marginal* state $\omega_A(-) = \omega(-, u_B)$ and a *conditional* state $\omega_{B|a}(b) = \omega(a, b)/\omega_A(a)$ (with the usual proviso that if $\omega_A(a) = 0$, the conditional state is also 0). We shall also refer to the partially evaluated state $\omega_B(a) := \omega(a, -)$ as an *un-normalized* conditional state.

More generally, if A is a composite of state spaces $A_1, ..., A_n$, with order-units $u_1, ..., u_n$, then for all subsets $J \subseteq \{1, ..., n\}$, and all $a := (a_i) \in \Pi_{i \notin J} A_i^*$, we can define an un-normalized conditional state – that is, a partially evaluated state – ω_J^a, a $|J|$-linear form on $\Pi_{j \in J} A_j^*$. We define the *J-th subsystem* to the the ordered space spanned by the cone generated by these conditional states, with order unit $u_J := \otimes_{j \in J} u_j$. We call A a *regular* composite iff it is closed under taking products of such multi-partite conditional states. All state spaces constructed from a single, associative, bilinear product are regular, but one can also build regular composites using "mixed" constructions. For instance, it is not difficult to show that $A \otimes_{min} (B \otimes_{max} C)$ is a regular composite of A, B and C. An example of a non-regular composite is $(A \otimes_{min} B) \otimes_{max} (C \otimes_{min} D)$ where A, B, C and D are four copies of a weakly self-dual, but non-classical, state space.

A state $\omega \in AB$ gives rise to a positive operator $\widehat{\omega} : A^* \to B$, given by $\widehat{\omega}(a)(b) = \omega(a, b)$. We can regard $\widehat{\omega}(a)$ as an "un-normalized" conditional state. As a partial converse, any positive operator $\psi : A^* \to B$ with $\psi(u_A) \in \Omega_B$ – that is, with $\psi^*(u_B) := u_B \circ \psi = u_A$ – corresponds to a state in the maximal tensor product $A \otimes_{max} B$. Dually, any effect $f \in (AB)^*$ yields an operator $\widehat{f} : A \to B^*$, given by $\widehat{f}(\alpha)(\beta) = f(\alpha \otimes \beta)$; and any positive operator $\phi : A \to B^*$ with $\phi(\alpha) \leq u_B$ for all $\alpha \in \Omega_A$ – that is, with $\|\phi\| \leq 1$ – corresponds to an effect in $(A \otimes_{min} B)^*$. We have the following result (easily verified by checking that it holds for elementary tensors):

Lemma 6.1 *Let ABC be a regular composite. If f is an effect in $(AB)^*$ and ω is a state in BC, then, for any $\alpha \in A$,*

$$(\alpha \otimes \omega)^B_f = \|\widehat{\omega} \circ \widehat{f}(\alpha)\| \widehat{\omega}(\widehat{f}(\alpha)). \tag{2}$$

If ABC is in a state $\alpha \otimes \omega$, with α unknown, then conditional on securing measurement outcome f on $A \otimes B$, the state of C is, up to normalization, a *known function of* α. We call this *remote evaluation.* This is very like a teleportation protocol. Indeed, suppose that C is a copy of A, and that $\eta : A \to C$ is a specified isomorphism allowing us to match up states in the former with those in the latter:

Definition 6.2 With notation as above, (f, ω) is a (one-outcome, post-selected) *teleportation protocol* iff there exists a positive, norm-contractive *correction map* $\tau : C \to C$ such that, for all $\alpha \in A$, $\tau(\alpha \otimes \omega)^C_f = \eta(\alpha)$.[5]

By Lemma 6.1, the un-normalized conditional state of $\alpha \otimes \omega$ is exactly $\widehat{\omega}(\widehat{f}(\alpha))$. If we let $\mu := \widehat{\omega} \circ \widehat{f}$, the normalized conditional state can be written as $\mu(\alpha)/u(\mu(\alpha))$. Thus, (f, ω) is a teleportation protocol iff there exists a norm-contractive mapping τ with $(\tau \circ \mu)(\alpha) = \|\mu(\alpha)\| \eta$ for all $\alpha \in \Omega_A$.

Theorem 6.3 *With notation as above, (f, ω) is a teleportation protocol iff $\mu := \widehat{\omega} \circ \widehat{f}$ is proportional to an isomorphism $(A, u_A) \simeq (C, u_C)$; in this case, the correction $\tau : (C, u_C) \simeq (C, u_C)$ is also an isomorphism.*

Henceforth, we simply identify C with A, suppressing η. Note that if (f, ω) is a teleportation protocol on a regular composite ABA of A, B and (a copy of) A, then, as f lives in $(AB)^* \leq (A \otimes_{min} B)^*$ and ω lives in $BA \leq B \otimes_{max} C$, one can also regard (f, ω) as a teleportation protocol on $A \otimes_{min} (B \otimes_{max} A)$.

[5] One could also allow protocols in which the correction has a nonzero probability to fail. For details, see [5].

Theorem 6.4 *$A \otimes_{min} (B \otimes_{max} A)$ supports a conclusive teleportation protocol iff A_1 is order-isomorphic to the range of a compression (a positive idempotent mapping) $P : A_2^* \to A_1$.*

Corollary 6.5 *If A can be teleported through a copy of itself, then A is weakly self-dual.*

In order to *deterministically* teleport an unknown state $\alpha \in A$ through B, we need not just one entangled effect f, but an entire observable's worth. Here, we specialize to the case in which $ABC = A \otimes_{min} (B \otimes_{max} C)$:

Definition 6.6 A *deterministic teleportation protocol* for A through B consists of an observable $E = (f_1, ..., f_n)$ on $A \otimes B$ and a state ω in $B \otimes A$, such that for all $i = 1, ..., n$, the operator $\widehat{f_i} \circ \widehat{\omega}$ is physically invertible.

The following result provides a sufficient condition (satisfied, e.g., by any state space A with Ω_A a regular polyhedron) for such a protocol to exist.

Theorem 6.7 *Let $A = B$. Suppose that G is a finite group acting transitively on the pure states of A, and let ω be a state such that $\widehat{\omega}$ is a G-equivariant isomorphism. For all $g \in G$, let $f_g \in (A \otimes_{max} A)^*$ correspond to the operator*

$$\widehat{f_g} = \frac{1}{|G|} \widehat{\omega}^{-1} \circ g.$$

Then $E = \{f_g | g \in G\}$ is an observable, and (E, ω) is a deterministic teleportation protocol.

7 Categories of Abstract State Spaces

If an abstract state space and its dual "effect space" provide an abstract probabilistic model, one should like to say that a probabilistic *theory* is a class of such models, closed under appropriate operations. To make this systematic, one should consider categories of state spaces. Let **Asp** denote the category whose objects are finite-dimensional abstract state spaces (A, u_A), and whose morphisms are norm-contractive positive linear mappings. This category has a preferred object $I = \mathbb{R}$, ordered as usual, with $u_I = 1$ and $\Omega_I = \{1\}$. For any A in **Asp**, there is a preferred morphism, namely u_A, from A to I. (Indeed, the mappings $\tau \mapsto u_A \circ \tau$ define a natural transformation $\mathbf{Asp}(-, A) \to \mathbf{Asp}(-, I)$.) We can model effect spaces, i.e., dual state spaces, by Hom sets $\mathbf{Asp}(A, I)$. However, as remarked above, there is no natural *internal* duality for **Asp**; nor is **Asp** naturally a monoidal category, owing to the the existence of *two* canonical tensor products $\otimes_{max}$ and $\otimes_{min}$. These interact in a way that will be familiar to linear-logicians, namely, for any state spaces A, B and C, there is a canonical embedding

$$A \otimes_{min} (B \otimes_{max} C) \leq (A \otimes_{min} B) \otimes_{max} C.$$

Thus, we can regard $(\mathbf{Asp}, \otimes_{min}, \otimes_{max})$ as a *linearly distributive* category [6] (albeit without negation). As to duality, there are various constructions whereby a useful self-duality can be supplied. Applied to **Asp**, these result in what may be regarded as categories of *process spaces.* To take the simplest case, consider the category $\mathbf{C}^2 = \mathbf{C} \times \mathbf{C}$, i.e., the category whose objects are ordered pairs (A, B) of state spaces in **Asp**, and in which

$$\mathbf{C}^2((A,B),(C,D)) = \mathbf{C}(C,A) \times \mathbf{C}(B,D),$$

with composition defined in the obvious way.[6] The idea is that the pair (A, B) represents a space of possible processes from A to B, and that a pair $f : C \to A$, $g : B \to D$ takes a process $\tau : A \to B$ to the process $g \circ \tau \circ f$. Indeed, the functor $(A, B) \mapsto \mathbf{C}(A, B)$ endows $\mathbf{C}^2$ with exactly this interpretation. Thus, (A, I) encodes A^*, while (I, A) encodes A. Thus, $\mathbf{C}^2$ allows us to consider state spaces and effect spaces on an equal footing. Moreover, $\mathbf{C}^2$ has a natural self-duality, given by $(A, B)^* = (B, A)$ and, for $(f, g) \in \mathbf{C}^2((A, B),(C, D))$,

$$(f,g)^* = (g,f) \in \mathbf{C}^2((C,D)^*,(A,B)^*) = \mathbf{C}^2((D,C),(B,A)).$$

Where $\mathbf{C}$ is closed under both maximal and minimal tensor products (in particular, for $\mathbf{C} = \mathbf{Asp}$), the category $\mathbf{C}^2$ has a natural symmetric monoidal structure given by[7]

$$(A,B) \otimes (C,D) = (A \otimes_{min} B, C \otimes_{max} D).$$

Note that we then have the expected identity

$$((A,B)^* \otimes (C,D)^*)^* = (A \otimes_{max} B, C \otimes_{min} D).$$

Rather than enlarging the category **Asp**, one can also look within it for subcategories with a desirable structure. Let $\mathbf{C}$ be a subcategory of **Asp**. If A and B are state spaces in $\mathbf{C}$, let us agree that, as in Definition 6.2, a *teleportation protocol* for *A through B* consists of (i) a regular composite $ABA \in \mathbf{C}$; and (ii) a pair $f \in \mathbf{C}(AB, I)$, $\omega \in BA = \mathbf{C}(I, BA)$ such that, for some *correction* $\tau \in \mathbf{C}(A, A)$,

$$\tau((f \otimes -)(- \otimes \omega)) = \mathrm{id}_A.$$

We shall say that this protocol is *correction free* if τ can be taken to be the identity morphism id_A. This may look familiar. Recall that a *dual* for an object A in a symmetric monoidal category $(\mathbf{C}, \otimes)$ is an object B, together with morphisms $\eta_A : I \to B \otimes A$ (the *unit*) and $\epsilon_A : A \otimes B \to I$ (the *co-unit*) such that (here identifying $I \otimes B$ and $B \otimes I$ with B, and suppressing the canonical association morphism $(A \otimes B) \otimes C \simeq A \otimes (B \otimes C)$)

$$(\mathrm{id}_B \otimes \epsilon_A) \circ (\eta_A \otimes \mathrm{id}_B) = \mathrm{id}_B.$$

If $\mathbf{C}$ is a monoidal category of state spaces (that is, a sub-category of **Asp** equipped with a symmetric, associative product $A, B \mapsto AB$), then $\eta \in \mathbf{C}(I, A \otimes B)$ is simply a positive, sub-normalized state in $A \otimes B$, while $\epsilon_A \in \mathbf{C}(A \otimes B, I)$ is simply an effect on $A \otimes B$, and we see that this amounts to the definition of a correction-free teleportation protocol in $\mathbf{C}$.

The foregoing discussion suggest one way to make contact between the structurally loose, but (so to say) ontologically rigid world of abstract state spaces, and

[6] This is essentially the category $\mathbf{C}^d$ described in [16]

[7] Another possibility for a category of processes would be to apply the Int [16] (or "GoI" [1]) construction to **Asp**. We shall not pursue this here.

the more highly structured but ontologically fluid categorical semantics of Abramsky and Coecke: begin with a particular category of abstract state spaces, and, from this, construct, either by enlarging it or by paring it down, a theory having, at a minimum, a sensible duality and monoidal structure.

It is also worth considering a different, more "top-down" approach: to proceed axiomatically, by laying down at the outset a minimum of constraints on what could count as a category of processes, and exploring the consequences of further requirements, e.g, that such a theory support teleportation, or that it not allow cloning, or bit-commitment. Of course, this is very close to the approach of Abramsky and Coecke and their collaborators; but we want to suggest that it may be fruitful to add an extra structural ingredient at the outset – namely, convexity. This is naturally captured by the notion of a category of processes *enriched over ordered linear spaces.* In the operational framework, the set of states, the set of measurement outcomes, the set of dynamics on a system or of dynamics turning one type of system into another, are all compact convex sets determined by imposing natural normalization conditions on convex cones of "un-normalized" states, outcomes, or dynamics, and the convexity is motivated by saying that we should be able to prepare two states, or perform two measurements, or implement two dynamics, conditional on the outcome of some random event with some definite ascribed probabilities p and $1-p$, such as the flip of a coin with known bias.

In other words, states, outcomes, and dynamics should belong to convex cones. In a process-centred approach, we will represent states of a system A by morphisms from a distinguished object, the "unit" (not to be confused with the "order unit" associated with a system), measurement outcomes by morphisms from A *to* the unit object, and dynamics on system A by morphisms from A to itself; dynamics changing a system of type A to one of type B may be represented by morphisms from A to B. The convexity requirement operationally motivated above will be implemented in this framework by requiring all hom-sets in a category describing an operational theory to be pointed, generating, closed convex cones; more formally, by requiring a category describing an operational theory to be enriched over a certain category of ordered linear spaces. This suggests the following

Definition 7.1 A *convex operational category of processes* is a category $\mathcal{C}$ enriched over the category of finite-dimensional ordered real vector spaces, with a *unit object* I, such that each object A is equipped with a distinguished morphism $u_A \in \mathcal{C}(A, I)$.

In future work, this definition will be elucidated in more detail. We intend it, and related categorical formulations of convex operational theories that we and collaborators are embarked on, to enable the comparison of the categorical formulation of information-processing centred around dagger compact closed categories, with the convex operational formalism. In the convex formalism, we have been concerned with obtaining necessary and/or sufficient conditions for the possibility of particular kinds of information processing, such as the ones we have reviewed in this paper. In some cases, it appears these conditions may be weaker than those employed in existing categorical constructions. We hope that the project of categorifying the convex approach (and convexifying the categorical approach!) may shed more light on *categorically formulated* necessary and sufficient conditions for various information-processing protocols (about which much, especially sufficient conditions, is known already), in part by enabling us to abstract from some of the more concrete content of the convex formalism while retaining some of its structural looseness.

References

[1] S. Abramsky. Retracing some paths in process algebra. In *Proceedings of CONCUR '96: Concurrency Theory, 7th International Conference*, U. Montanari and V. Sassone, eds., Springer-Verlag, pages 1–17, 1996.

[2] S. Abramsky and B. Coecke. A categorical semantics of quantum protocols. *Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science (LICS '04)*, pages 415–425, 2004.

[3] J. Baez. Quantum quandaries: a category-theoretic perspective. `quant-ph/0404040`, 2004.

[4] H. Barnum, J. Barrett, M. Leifer, and A. Wilce. Cloning and broadcasting in generic probabilistic models. arXiv.org e-print `quant-ph/0611295`, 2006.

[5] H. Barnum, J. Barrett, M. Leifer and A. Wilce, Teleportation in general probabilistic theories. In Proceedings of the Clifford Lectures, Tulane University, March 12-15, 2008. To appear in *Proceedings of Symposia in Applied Mathematics* (AMS); also arXiv:0805.3553, 2008.

[6] H. Barnum, C. M. Caves, C. A. Fuchs, and B. Schumacher. Noncommuting mixed states cannot be broadcast. *Phys. Rev. Lett.*, 76:2818–2821, 1996.

[7] H. Barnum, O. Dahlsten, M. Leifer, and B. Toner. Nonclassicality without entanglement enables bit commitment. accepted to IEEE Information Theory Workshop, Porto, May 2008, 2007.

[8] J. Barrett. Information processing in generalized probabilistic theories. arXiv.org e-print `quant-ph/0508211`. To appear in Physical Review A, 2005.

[9] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts. Nonlocal correlations as an information-theoretic resource. *Phys. Rev. A*, 71:022101, 2005.

[10] C. H. Bennett and G. Brassard. Quantum cryptography: Public-key distribution and coin-tossing. In Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India, pages 175–179. IEEE Press, 1985.

[11] G. Brassard. Is information the key? *Nature Physics*, 1:2–4, 2005.

[12] G. M. d'Ariano. How to derive the Hilbert-space formulation of quantum mechanics from purely operational axioms. arXiv.org e-print `quant-ph/0603011`, 2006.

[13] C. A. Fuchs. Quantum mechanics as quantum information, mostly. *J. Mod. Opt.*, 50:987, 2003.

[14] L. Hardy. Quantum theory from five reasonable axioms. arXiv.org e-print `quant-ph/0101012`, 2001.

[15] L. Hardy. Why quantum theory? arXiv.org e-print `quant-ph/0111068`. Contribution to NATO Advanced Research Workshop "Modality, Probability, and Bell's Theorem", Krakow, Poland 19–23.8.01, 2001.

[16] J. Hyland and A. Schalk. Glueing and orthogonality for models of linear logic. *Theoretical Computer Science*, 294:405–407, 2003.

[17] H.-K. Lo and H. F. Chau. Is quantum bit-commitment really possible? *Phys. Rev. Lett.*, 78:3410–3413, 1997.

[18] D. Mayers. Unconditionally secure quantum bit commitment is impossible. *Physical Review Letters*, 78:3413, 1997.

[19] P. Selinger. Towards a semantics for higher-order quantum computation. In Proceedings of the 2nd International Workshop on Quantum Programming Languages, Turku Finland, pages 127–143. Turku Center for Computer Science, 2004. Publication No. 33.

[20] P. Selinger. Dagger compact closed categories. *Electronic Notes in Theoretical Computer Science*, 170:139–163, 2007. Proceedings of the 3rd International Workshop on Quantum Programming Languages (QPL 2005), Chicago.