Using BV to Describe Causal Quantum Evolution

Prakash Panangaden*
Computing Laboratory
Oxford University
Oxford, U.K.
prakash@cs.mcgill.ca

July 5, 2004

1 Basic Notions and Definitions

In this note I describe how to capture the *kinematics* of quantum causal evolution using a logic called BV developed by the Calculus of Structures group at Dresden. The setting is discrete quantum mechanics. We imagine a finite "web" of spacetime points. The points are events in spacetime. They are viewed as verhtees in a directed acyclic graph (DAG); the edges of the DAG represent causal links mediated by the propagation of matter. The fact that the graph is acyclic captures a basic causality requirement: there are no closed causal trajectories.

The DAG represents a discrete approximation to the spacetime on which a quantum system evolves. The graph is technically a *dangling* graph; there is a set of *half edges* - in addition to the ordinary edges - divided into two disjoint subsets: the *incoming edges* and the *outgoing edges*. An incoming edge has no initial point but has a terminal point, and dually for outgoing edges.

Definition 1.1 *A* ***spacetime graph*** G *consists of:*

1. *a finite set* P *of* points*,*
2. *a finite set* E *of* edges*,*
3. *disjoint finite sets* I *and* O *of incoming and outgoing edges,*
4. *maps* $\mathsf{source}, \mathsf{target}$ *from* E *to* P*, and*
5. *maps* $\mathsf{past} : I \longrightarrow P$ *and* $\mathsf{future} : O \longrightarrow P$*.*

We imagine particles propagating along the edges and engaging in interactions at the vertices. The incoming and outgoing edges represent particles arriving from and going off to distant asymptotic regimes. In spirit these graphs are like Feynman diagrams. We say that an edge e *immediately precedes* e' if $\mathsf{target}(e) = \mathsf{source}(e')$. The reflexive, transitive closure of the "immediately precedes" relation is pronounced "precedes" and is written $e_1 \leq e_2$. Since the graph is acyclic this is clearly

*On leave from: School of Computer Science, McGill University, Montréal, Québec, Canada a partial order. We can define a similar relation between vertices and even between vertices and edges. Incoming edges and outgoing edges can also be ordered - relative to each other and ordinary edges - by the same relation in the evident way. No edges strictly precede incoming edges nor do outgoing edges strictly precede any other edges. The precedence relation gives each spacetime graph a *causal structure*.

A set of edges such that no two edges in the set are related by causal precedence is called a *slice* or *spacelike slice*; these are *anti-chains* in the language of posets. A slice may be maximal or not. Any subset of the incoming edges is a slice as is any subset of the outgoing edges. Given slices S_1 and S_2 we can define a partial order as follows:

$$S_1 \sqsubseteq S_2 \iff \forall p \in S_1 \exists q \in S_2 \; p \leq q \text{ and } \forall q \in S_2 \exists p \in S_1 \; p \leq q.$$

This is the well known Egli-Milner ordering from concurrency theory in computer science.

The dynamics of quantum systems is described as follows. With each edge we associate a Hilbert space $\mathcal{H}$ and a density matrix ρ associated with the subsystem on that edge. Such a density matrix is a positive operator on $\mathcal{H}$ with trace *less than or equal to* 1[1]. At each vertex we imagine that we have an *interaction*, which may be any one of the following:

1. two, or more, subsystems coming together and interacting;
2. a subsystem breaking into pieces;
3. a subsystem being subject to a unitary transformation;
4. a subsystem being subject to a measurement;
5. a subsystem being partly discarded.

When subsystems come together we form the tensor product of their state spaces. If they have no interaction we form the tensor product of their density matrices, otherwise we have a unitary operator acting on the combined density matrices. When a system breaks apart we can have a single density matrix for all the pieces; if, however, we wish to separate the density matrices of the individual components we compute partial traces. This, of course, has the effect of removing information about nonlocal correlations. A unitary transformation U acts on a density matrix ρ by $\rho \mapsto U\rho U^\dagger$ and the effect of a measurement M is to apply a projector P: $\rho \mapsto P\rho P$. The most general physical transformation of a system is described by a *superoperator*, which is a trace-nonincreasing *completely positive map* acting on density matrices. A well known theorem, the Kraus representation theorem, states that every such map $\mathcal{E}$ can be written as

$$\mathcal{E}(\rho) = \sum_i^n E_i \rho E_i^\dagger$$

where the E_i are any linear operators such that

$$\sum_i^n E_i^\dagger E_i \leq I.$$

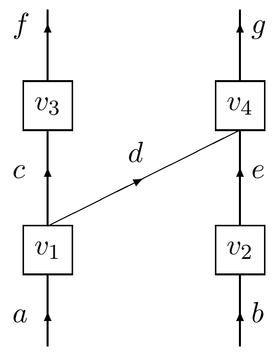

Figure 1: A simple causal graph.

Now density matrices can be associated with any slice. If we keep all the data associated with **maximal** slices then we cannot guarantee that information does not propagate between acausal paths.

Consider the causal graph shown in figure 1. If we always take maximal slices then we cannot be sure that the state of the density matrix at edge b does not influence the density matrix at edge f. One solution to guaranteeing causal propagation is to only propagate along the individual edges. In this scheme we would only allow the operators (completely positive maps) at the vertices to act on density matrices associated with single edges. This would indeed guarantee causal propagation but would kill all nonlocal correlations.

Consider the system shown in Figure 2. Here the density matrices ρ_b and ρ_c do not express non-local correlations that might exist as a result of their common origin as subsystems of ρ_a.

The solution to the problem of ensuring causal evolution while preserving important non-local correlations is to work with slices called *locative* slices defined below.

Fix any subset of incoming edges. These always form a slice. Suppose S is a slice and v is a vertex such that all the incoming edges of v are in S. We write $\mathsf{In}(v)$ for $\{e|\mathsf{source}(e) = v\} \cup \{i|\mathsf{past}(i) = v\}$ and similarly for $\mathsf{Out}(v)$. Then, clearly,

$$(S \setminus \mathsf{In}(v)) \cup \mathsf{Out}(v)$$

is always a slice. It is the slice obtained by *propagating* S through v.

Definition 1.2 *A **locative** slice is defined by induction.*

- *Any subset of the incoming edges of the graph forms a locative slice.*
- *If S is a locative slice and v is a vertex with $\mathsf{In}(v) \subset S$ then the slice obtained by propagating S through v is locative.*

The point is that if S is locative then the density matrix on S can be computed without ever computing partial traces: no information is lost.

The idea behind the prescription for evolving can now be simply stated. Each edge - more generally, each slice - has a density matrix. Pick a family of slices linearly ordered so that each

[1]We allow our density matrices to be not normalized.

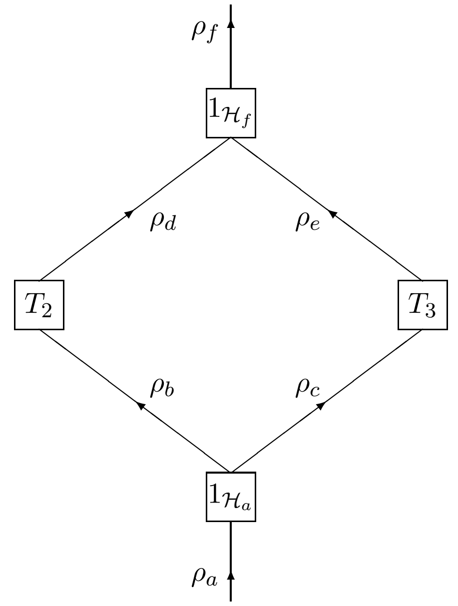

Figure 2: A system with non-local correlations.

slices differs from the immediately preceding one by propagating through one vertex. Work with a family of locative slices and propagate the density matrix from one slice to the next using the superoperator at the vertex through which the propagation is being carried out.

We have said before that every vertex corresponds to a superoperator. For trivial reasons this superoperator will depend on the family of slices. Consider again the causal graph shown in figure 2. One possible family of slices is a, bc, dc, de, f; another one is a, bc, be, de, f. The type of the superoperator at vertex v_1 is:

$$T_1 : \mathcal{DM}(\mathcal{H}_a) \longrightarrow \mathcal{DM}(\mathcal{H}_b \otimes \mathcal{H}_c).$$

In the first slicing, the type of the superoperator at v_2 is

$$T_2 : \mathcal{DM}(\mathcal{H}_b \otimes \mathcal{H}_c) \longrightarrow \mathcal{DM}(\mathcal{H}_d \otimes \mathcal{H}_c)$$

while with the second slicing we get the type:

$$T_2 : \mathcal{DM}(\mathcal{H}_b \otimes \mathcal{H}_e) \longrightarrow \mathcal{DM}(\mathcal{H}_d \otimes \mathcal{H}_e).$$

Each version of T_2 is padded out with the appropriate identity operators: the “real action” of T_2 transforms the b-piece of the density matrix into the d-piece. The important point is that superoperators at spacelike separated vertices commute.

One can now state the prescription more precisely. In order to compute the density matrix for an edge e, one first computes the density matrix on the **minimal** - in the Egli-Milner ordering - locative slice containing e, then one takes the appropriate partial traces.

Suppose L is a locative slice and u and v are two minimal - in the causal order - vertices above L. Clearly u and v are acausal with respect to each other so their superoperators commute. Thus we can go $L \longrightarrow L_u \longrightarrow L_{uv}$ or $L \longrightarrow L_v \longrightarrow L_{uv}$. Clearly L_u, L_v and L_{uv} are all locative and the density matrix on L_{uv} will be the same calculated either way. We can piece together such "diamonds" inductively and prove slicing independence by an easy inductive argument. The fact that we are calculating the density matrix at an edge by working with the *minimal* locative slice guarantees causal propagation: essentially because the only vertices that can affect the outcome are to the causal past of the edge. It is not hard to formalize these observations.

What this framework does not do very well is to deal with spatially distributed pure states of a single subsystem; for example the states that arise when one uses a Mach-Zedner type interferometer as a beam splitter. That will be the subject of later work.

2 A Logic for Causal Propagation: Problems

In the previous section I described the "physics" of discrete quantum causal propagation. Here I want to describe a logic capturing the essence of the concept of "locative" slice. The idea is to have a propositional logic where the atoms represent edges, vertices correspond to axioms and locative slices correspond to derivable sequents.

The key unit of deduction that we [BIP03] took originally is the sequent; its typical form is:

$$A_1, A_2, \ldots, A_n \vdash B_1, B_2, \ldots, B_m.$$

Here the atoms are names of edges appearing in some causal graph. We note that for purposes of this paper sequents should always be considered "up to permutation", i.e. one may rearrange the order of premises and conclusions as one sees fit. Our system will have only one inference rule, called the *Cut rule*, which states:

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma', A \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'}$$

Axioms are of the form $A_1, A_2, \ldots, A_n \vdash B_1, B_2, \ldots, B_m$, where $A_1, A_2, \ldots, A_n$ are the incoming edges of some vertex in our causal graph, and $B_1, B_2, \ldots, B_m$ will be the outgoing edges. There will be one such axiom for each vertex. For example, consider Figure 3. Then we will have the following axioms:

$$a \overset{1}{\vdash} c \quad b \overset{2}{\vdash} d, e, f \quad c, d \overset{3}{\vdash} g, h \quad e \overset{4}{\vdash} i \quad f, g \overset{5}{\vdash} j \quad h, i \overset{6}{\vdash} k$$

where we have labelled each entailment symbol with the name of the corresponding vertex. The following is an example of a deduction in this system of the sequent $a, b \vdash f, g, h, i$.

$$\dfrac{\dfrac{b \vdash d, e, f \quad \dfrac{a \vdash c \quad c, d \vdash g, h}{a, d \vdash g, h}}{a, b \vdash e, f, g, h} \quad e \vdash i}{a, b \vdash f, g, h, i}$$

As a first attempt at capturing quantum evolution on a causal graph G axiomatically we Then one would (tentatively) define a set Δ of edges to be *valid* if there is a deduction in the logic generated by G of $\Gamma \vdash \Delta$ where Γ is a set of initial edges.

However, with this notion of validity, we would fail to capture all locative slices, and thus our tentative notion of validity will have to be modified. For example, consider the dag underlying the system of Figure 2 shown in Figure 4.

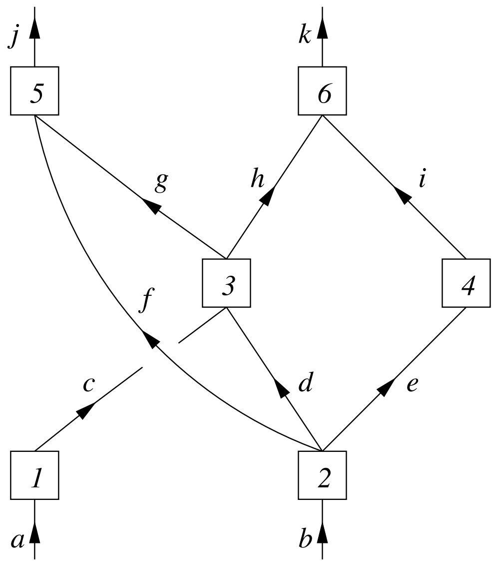

Figure 3: A typical causal graph

Corresponding to this dag, we get the following basic morphisms (axioms):

$$a \vdash b, c \quad b \vdash d \quad c \vdash e \quad d, e \vdash f.$$

Evidently, the set $\{f\}$ is a locative slice, and yet the sequent $a \vdash f$ is not derivable. The sequent $a \vdash d, e$ is derivable, and one would like to cut it against $d, e \vdash f$, but one is only allowed to cut a single formula. Such "multicuts" are usually expressly forbidden, as they lead to undesirable logical properties [Blu93]. However, one viewpoint is that it is time to think about logics with such multicuts and face up to their consequences rather than just avoiding them. That is the subject of other discussions: we will not pursue it further here.

Physically, the reason for this problem is that the sequent $d, e \vdash f$ does not encode the information that the two states at d and e are correlated. It is precisely the fact that they are correlated that implies that one would need to use a multicut. To avoid this problem, one must introduce some notation, specifically a syntax for specifying such correlations. We will use the *logical connectives* of the multiplicative fragment of *linear logic* [Gir87, Gir95] to this end. The multiplicative disjunction of linear logic, denoted $⅋$ and called the *par* connective, will express such nonlocal correlations. In our example, we will write the sequent corresponding to vertex 4 as $d ⅋ e \vdash f$ to express the fact that the subsystems associated with these two edges are possibly entangled through interactions in their common past.

Note that whenever two (or more) subsystems emerge from an interaction, they are correlated. In linear logic, this is reflected by the following rule called the (right) *Par rule*:

$$\frac{\Gamma \vdash \Delta, A, B}{\Gamma \vdash \Delta, A ⅋ B}$$

Thus we can always introduce the symbol for correlation in the right hand side of the sequent.

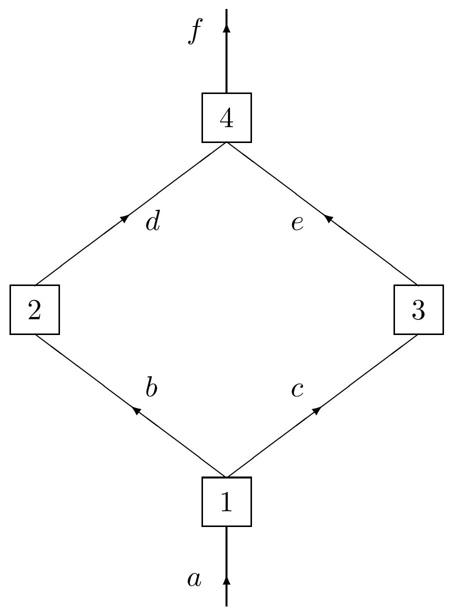

Figure 4:

Notice that we can cut along a compound formula without violating any logical rules. So in the present setting, we would have the following deduction:

$$\dfrac{\dfrac{\dfrac{\dfrac{a \vdash b, c \quad b \vdash d}{a \vdash c, d} \qquad c \vdash e}{a \vdash d, e}}{a \vdash d \invamp e} \qquad d \invamp e \vdash f}{a \vdash f}$$

All the cuts in this deduction are legitimate; instead of a multicut we are cutting along a compound formula in the last step.

The above logical rule determines how one introduces a par connective on the righthand side of a sequent. For the lefthand side, one introduces pars in the axioms by the following general prescription. Given a vertex in a multigraph, we suppose that it has incoming edges $a_1, a_2, \ldots, a_n$ and outgoing edges $b_1, b_2, \ldots, b_m$. In the previous formulation, this vertex would have been labelled with the axiom $\Gamma = a_1, a_2, \ldots, a_n \vdash b_1, b_2, \ldots, b_m$. We will now introduce several pars ($\invamp$) on the lefthand side to indicate entanglements of the sort described above. Begin by defining a relation $\sim$ by saying $a_i \sim a_j$ if there is an initial edge c and directed paths from c to a_i and from c to a_j. This is not an equivalence relation, but one takes the equivalence relation generated by the relation $\sim$. Call this new relation $\cong$. This equivalence relation, like all equivalence relations, partitions the set Γ into a set of equivalence classes. One then "pars" together the elements of each equivalence class, and this determines the structure of the lefthand side of our axiom. For example, consider vertices 5 and 6 in Figure 3. Vertex 5 would be labelled by $f \invamp g \vdash j$ and vertex 6 would be labelled by $h \invamp i \vdash k$. On the other hand, vertex 3 would be labelled by $c, d \vdash g, h$.

Just as the par connective indicates the existence of past correlations, we use the more familiar tensor symbol $\otimes$, which is also a connective of linear logic, to indicate the lack of nonlocal correlation. This connective also has a logical rule:

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma' \vdash \Delta', B}{\Gamma, \Gamma' \vdash \Delta, \Delta', A \otimes B}$$

But we note that unlike in ordinary logic, this rule can only be applied in situations that are physically meaningful. We will say that two deductions π and π' are *spacelike separated* if all the the vertices of π and π' are pairwise spacelike separated. In the above formula, we require that the deductions of $\Gamma \vdash \Delta, A$ and $\Gamma' \vdash \Delta', B$ are spacelike separated. This restriction of application of inference rules is similar to the restrictions of *ludics* [Gir01].

Summarizing, to every causal graph G we associate its "logic", namely the edges are considered as formulas and vertices are axioms. We have the usual linear logical connective rules, including the cut rule which in our setting is interpreted physically as propagation. The par connective denotes possible correlation, and the tensor lack of correlation. Note that every deduction in our system will conclude with a sequent of the form $\Gamma \vdash \Delta$, where Γ is a set of initial edges.

Now one would like to modify the definition of validity to say that a set of edges Δ is *valid* if in our extended logic, one can derive a sequent $\Gamma \vdash \hat{\Delta}$ such that the list of edges appearing in $\hat{\Delta}$ was precisely Δ, and Γ is a set of initial edges. However this is still not sufficient as an axiomatic approach to capturing all locative slices. We note the example in Figure 5.

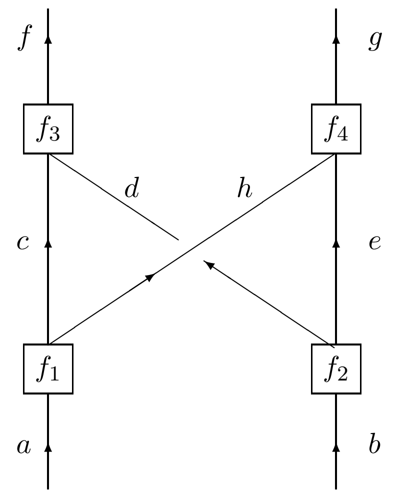

Figure 5: Induced entanglement

Evidently the slice $\{f, g\}$ is locative, but we claim that it cannot be derived even in our extended logic. To this causal graph, we would associate the following axioms:

$$a \vdash c, h \quad b \vdash d, e \quad c, d \vdash f \quad h, e \vdash g$$

Note that there are no correlations between c and d or between h and e. Thus no $\mathbin{\rotatebox[origin=c]{180}{\&}}$-combinations can be introduced. Now if one attempts to derive $a, b \vdash f, g$, we proceed as follows:

$$\dfrac{\dfrac{a \vdash c, h \quad b \vdash d, e}{a, b \vdash c \otimes d, h, e} \quad \dfrac{c, d \vdash f}{c \otimes d \vdash f}}{a, b \vdash h, e, f}$$

At this point, we are unable to proceed. Had we attempted the symmetric approach tensoring h and e together, we would have encountered the same problem.

The problem is that our logical system is still missing one crucial aspect, and that is that correlations develop dynamically as the system evolves, or equivalently as the deduction proceeds. One approach that we have considered is to let the axioms evolve dynamically. I feel that this is a very unsatisfactory hack and puts the dyanmics in from the outside; though this "works" I cannot see what we learn from it. For a while I was convinced that there was no choice but to do it this way. However, Alessio Guglielmi showed me the logic of the next section which seems to have the features that we need.

3 The Logic BV and Quantum Evolution

We take the system CBV described as follows. There are *atoms*, which we take to be purely positive though in general one has positive and negative atoms; and there are structures which are to be thought of as compounds. The syntax of structures are as follows:

$$S ::== A|[S, \ldots, S]|\langle S,\ \ldots S\rangle|(S, \ldots, S)|0.$$

The special atom 0 is a unit for the connectives $[,]$, $\langle,\ \rangle$ and $(,)$. The connectives themselves are associative and commutative.

We use the following rules: for *cut* we have $\frac{0}{[a,\bar{a}]}$ and for *axiom* we have $\frac{(a,\bar{a})}{0}$. We do use negated atoms in our physics application so far; instead we will introduce axioms for vertices as in the previous section. This logic is not based on sequent calculus notions; instead it uses *deep inference*, which allows one to rewrite inside a term. The connectives are all commutative and associative. The following *switch* rule is very useful:

$$\frac{(R, [T, U])}{[(R, T), U]}.$$

The remaining two rules are

$$\frac{\langle [R, U],\ [T, V]\rangle}{[\langle R,\ T\rangle, \langle U,\ V\rangle]} \qquad \text{and} \qquad \frac{(\langle R,\ U\rangle, \langle T,\ V\rangle)}{\langle (R, T),\ (U, V)\rangle}.$$

They are enigmatically named $q\downarrow$ and $q\uparrow$ respectively.

We use the notation $\alpha \multimap \beta$ to mean that we can give a derivation of β starting from α on top. However, we usuually build the proofs from the bottom going up. As a warm-up for novices (like me) I will show $\langle R,\ T\rangle \multimap [R, T]$.

$$\dfrac{\dfrac{\dfrac{\langle R,\ T\rangle}{\langle [R, 0],\ [0, T]\rangle}}{[\langle R,\ 0\rangle, \langle 0,\ T\rangle]}}{[R, T]}$$

We will use this as a lemma in the next proof.

Now turning to the causal graph of figure 5 we can introduce axioms as follows: $\frac{\langle c,\ d\rangle}{a}$, $\frac{\langle e,\ f\rangle}{b}$, $\frac{g}{\langle c,\ e\rangle}$ and $\frac{h}{\langle d,\ f\rangle}$. Now we show that $\langle g,\ h\rangle\ -\!\circ\ [a,b]$; in order to see this in the spirit of the last section we have to read the derivations backwards. Everthing in this calculus seems to have a beautiful symmetry!

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\langle g,\ h\rangle}{\langle g,\ \langle d,\ f\rangle\rangle}}{\langle g,\ [d,f]\rangle}}{\langle\langle c,\ e\rangle,\ [d,f]\rangle}}{\langle [c,e],\ [d,f]\rangle}}{[\langle c,\ d\rangle,\langle e,\ f\rangle]}}{[\langle c,\ d\rangle,b]}}{[a,b]}$$

Reading the proof from the bottom up the justifications are: graph axiom, graph axiom, $q\downarrow$, lemma, graph axiom, lemma and graph axiom.

One can easily see that - at least for this graph - there are no derivations of non locative slices and we have got the derivation that we wanted without the time dependence. The key seems to be the appearance of a third connective, which mediates between the other two.

Recently Abramsky and Coecke have shown a very nice and quite tight connection between compact closed categgories with biproducts and quantum mechanics. The logic of compact closed categories has long been looked down upon, especially by those enamoured of linear logic, but now it is clear that we need to take compact closed proof nets seriously. A recent paper by Abramsky and Duncan does just that; it gives a typed language for expressing quantum algorithms. Using compact closed category corresponds, roughly speaking, to allowing multicut. This would solve the problems associated with expressing correlation but it would not deal with the time dependence of entanglement. The problem with the induced entanglement is something that only this logic with its three connectives seems to solve in a nice way. What is the underlying category theory?

Acknowledgements

Instead of thanking all my famous friends and name dropping shamelessly I will thank the people who actually helped me to understand this, hopefully they will also be famous one day: first and foremost Alessio Guglielmi who saw this logic in the middle of my babble about entanglement and, secondly Ross Duncan who worked out a nice treatment of entanglement swapping in this logic.

References

[BIP03] R. F. Blute, I. T. Ivanov, and P. Panangaden. Discrete quantum causal dynamics. *International Journal of Theoretical Physics*, 2003. in press.

[Blu93] R. Blute. Linear logic, coherence and dinaturality. *Theoretical Computer Science*, 115:3–41, 1993.

[Gir87] J.-Y. Girard. Linear logic. *Theoretical Computer Science*, 50:1–102, 1987.

[Gir95] J.-Y. Girard. Linear logic: its syntax and semantics. In J.-Y. Girard, Y. Lafont, and L. Regnier, editors, *Advances in Linear logic*, number 222 in London Mathematics Society Lecture Note Series, pages 1–42. Cambridge University Press, 1995.

[Gir01] J.-Y. Girard. Locus solum. *Mathematical Structures in Computer Science*, 2001.