# Is Schrödinger's Cat Real? Exploring Quantum Bayesianism with Monoidal Categories

Edward Ayers

April 17, 2016

**Abstract**

This is an essay outlining Categorical Quantum Mechanics (CQM) for a reader familiar with quantum mechanics. No experience with category theory is assumed. CQM is a new formulation of quantum information that uses the diagrammatic language of monoidal categories to reframe concepts and problems that arise in the subject. In this paper dagger compact closed categories will be defined and its relation to Hilbert Space will be quantified. The Selinger CPM construction will be defined and will be used to model classical and quantum information. This gives rise to a mathematical distinction between ontology and epistemology. It will then be shown that the language of density matrices and unitary transformations arises naturally from this approach. It will be discussed how CQM can be used to describe and reason about a physical setup involving both quantum and classical elements and it will be linked to the usual postulates of quantum mechanics such as wavefunction collapse. Then the example of quantum teleportation in terms of CQM will be explored. Finally, some original work will be presented on how CQM and quantum Bayesianism may lead to a surprising result regarding the Schrödinger's Cat thought experiment.

## 1 Introduction

This is an essay about quantum mechanics and about how a lot of its features that we take to be real could actually be features of how we think. This essay will have succeeded if it pushes you towards believing the ideas of quantum Bayesianism (also known as QBism) [1, 2]. Bayesian probability is the idea that probability theory is purely a theory about one's state of knowledge. For instance, a Bayesian would argue that the 50/50 chance associated with a coin toss is not a physical property of the coin but instead a reflection of its chaotic mechanics and our lack of knowledge about the initial conditions of the toss. More information about Bayesian probability can be found at [3, 4]. Bayesianism is certainly not unanimously accepted among the scientific community. This is mainly because Bayesianism gives a controversial and not entirely complete answer to thermodynamics and also because the mainstream formulation of quantum mechanics (states, measurements, collapsing etc.) incorporates probabilities at a fundamental level. In this essay, I will show how categorical quantum mechanics can help provide a formulation that defines finite dimensional quantum mechanics in a way that restores the Bayesian sense of probability. While this is not a proof of Bayesianism, it should help to quell some doubts that it is incompatible with quantum mechanics. We will also see that this formulation produces some controversies of its own.

In any scientific theory, there is a division between what the theory considers to be real and what is a result of the way we gather information about the world (called epistemology). In most theories, probability lies on the epistemological side of the divide - a Bayesian would argue that it always does. For instance; in planetary astronomy, the planets are considered to be points with absolute positions. When we measure the positions using a telescope, we pick up some error from the inaccuracy of the equipment. This error is not a property of the planets themselves. It is a property of the equipment and the way we gathered information about the planets, it is epistemological. There are also fields of science where it is extremely difficult to measure anything without interfering with the system (such as cellular biology), but we still maintain the belief that there is a physical thing 'out there' that exists and runs independently of whether we choose to find out about it.

Our (Copenhagen) theory of quantum mechanics disobeys this division. In quantum mechanics, we are forced to embrace the measurement as an intrinsic aspect of the physics. It is difficult to formulate quantum mechanics in a way that explains measurement and wavefunction collapse without invoking something we would regard as epistemology in any other physical theory. This is true even if we view a measurement as an observer-less interaction with the environment because we are still forced to invoke thermodynamical arguments. As mentioned before, thermodynamics also suffers from a great deal of disagreement about epistemology.

In this essay I will outline the new field of categorical quantum mechanics (CQM) [5, 6] with a focus on how it may help point towards a clean division between the physical and the epistemological aspects of quantum theory.

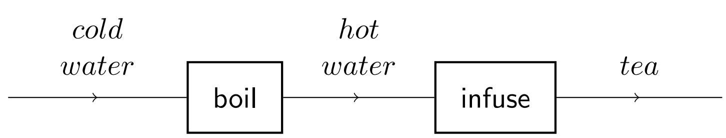
Along the way I will also try to derive the postulates of quantum mechanics in a way that appeals very strongly to intuition to help dispel some of the mystery that often shrouds discussions of the quantum world. I will also explain the correspondence between the new conceptual machinery and the orthodox formulation of QM to assure readers that the new results match up with our current body of knowledge about QM.

## 2 Dagger compact categories

Category theory is usually introduced as a system for building bridges between seemingly disparate areas of mathematics. Here, we instead take the view that a category is a generalisation of the idea of a process.

**Definition 2.1** (Category)**.** A **category** $\mathcal{C}$ is a family of objects $\mathcal{C}_0$ and a family of morphisms $\mathcal{C}_1$. Each $f \in \mathcal{C}_1$ has a source object $s(f) \in \mathcal{C}_0$ and a target object $t(f) \in \mathcal{C}_0$. Suppose $s(f) = A$ and $t(f) = B$, then we can write this concisely as $f : A \to B$. A category also has a composition operation which assigns to each $f : A \to B$ and $g : B \to C$ a morphism $g \circ f : A \to C$. Composition is associative; $(h \circ g) \circ f = h \circ (g \circ f)$. Furthermore, each object $A$ has a special morphism $1_A : A \to A$ with $1_B \circ f = f$ and $f \circ 1_A = f$ for all $f : A \to B$.

To understand the motivation for this definition, consider the flow chart below.




(1)

We will call these **process diagrams**. Considered in terms of category theory; we can think of each wire in the flow chart as an object and each box as a morphism. For instance, boil is a morphism and *tea* is an object. So the composition law of category theory is simply the statement that we can attach the output of one process to the input of another process to make a new process. The identity law expresses the fact that we always have a 'do nothing' process. A lot of objects in mathematics are also examples of categories. For example; the category of sets and functions between them and the category of groups and group homomorphisms between them.

We also want a way of mapping one category into another. However, we only want to consider mappings that preserve the categorical structure. This is the role of the functor.

**Definition 2.2** (Functor)**.** Given two categories $\mathcal{C}$ and $\mathcal{D}$, we define a **functor** $F : \mathcal{C} \to \mathcal{D}$ as a function on objects $F_0$ and a function on morphisms $F_1$ such that given $f : A \to B$ we have $F_1(f) : F_0(A) \to F_0(B)$ and also $F_1(g \circ f) = F_1(g) \circ F_1(f)$. From now on we drop the subscripts on $F$, since it will be obvious from context which map we mean.

Now we add some extra features to a category:

**Definition 2.3** ((Strict) monoidal Category)**.** A **strict monoidal category** is a category $\mathcal{C}$ equipped with the following items;

- an operation on objects $\otimes : \mathcal{C}_0 \times \mathcal{C}_0 \to \mathcal{C}_0$.
- a parallel composition operation $\otimes$ on morphisms which takes $f : A \to B$ and $g : C \to D$ to $f \otimes g : (A \otimes C) \to (B \otimes D)$.
- a special unit object $I \in C_0$.

We demand that these obey the following rules for all $A, B, C \in \mathcal{C}_0$;

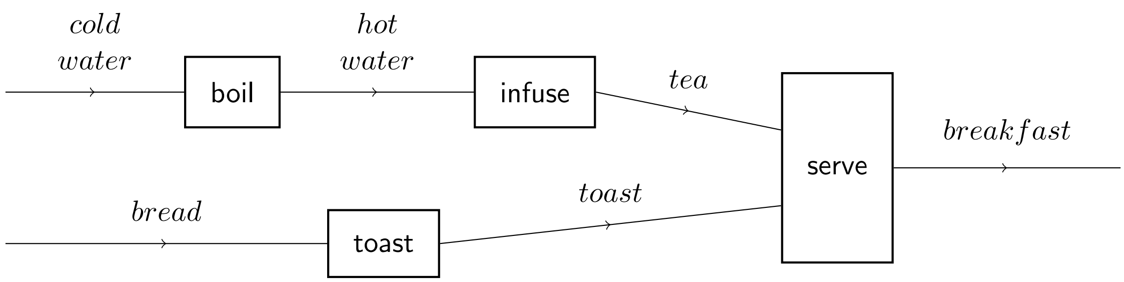
1. $I \otimes A = A = A \otimes I$
2. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
3. Given $f : A \to B$, $g : B \to C$, $h : A' \to B'$, $k : B' \to C'$, we have the interchange law: $(g \circ f) \otimes (k \circ h) = (g \otimes k) \circ (f \otimes h)$.
4. $1_{A \otimes B} = 1_A \otimes 1_B$

We could also define a **non-strict monoidal category** as a strict monoidal category where $\otimes$ obeys the identity and associativity laws only up to isomorphism. That is, we have a invertible morphisms $\alpha : (A \otimes B) \otimes C \to A \otimes (B \otimes C)$, $\rho : A \to A \otimes I$, $\lambda : A \to I \to A$. These isomorphisms also need to satisfy some coherence relations. We do not need to concern ourselves too carefully with the details here because of the coherence theorem in [7, VII.2]. In essence, the coherence theorem shows that in all monoidal categories we may as well reason as if the category is strict.

**Definition 2.4** (Symmetry)**.** We can also add a **symmetry** to the monoidal structure. This is a morphism $c_{A,B} : (A \otimes B) \to (B \otimes A)$ for each $A, B \in \mathcal{C}_0$ satisfying the following 5 axioms for all $A, B, C \in \mathcal{C}_0$ and $f : A \to B$:

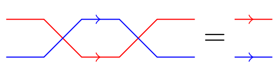
1. $c_{B,A} \circ c_{A,B} = 1_{A \otimes B}$
2. $c_{B,C} \circ (f \otimes 1_C) = (1_C \otimes f) \circ c_{A,C}$
3. $c_{C,B} \circ (1_C \otimes f) = (f \otimes 1_C) \circ c_{C,A}$
4. $(1_B \otimes c_{A,C}) \circ (c_{A,B} \otimes 1_C) = c_{A,(B \otimes C)}$
5. $(c_{A,C} \otimes 1_B) \circ (1_A \otimes c_{B,C}) = c_{(A \otimes B),C}$

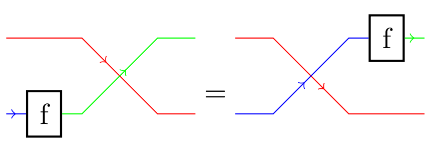
A monoidal category can be thought of as a category that allows for running two processes in parallel.

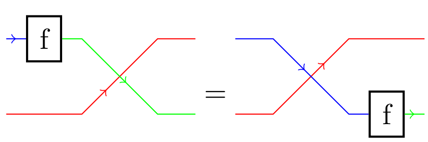


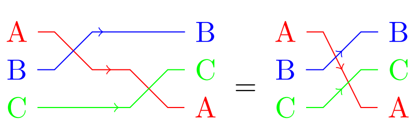

(2)

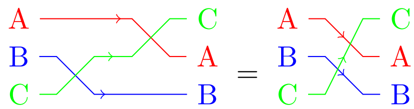
If we have a symmetry $c_{A,B}$ then we can represent this in a process diagram as a pair of crossed wires  . Under this representation, the 5 axioms in definition 2.4 look like the following diagrams.

1. 

2. 


3. 


4. 

5. 


Surprisingly, diagrams such as these are more than visual aids. We can argue quite rigorously that the symmetric monoidal category is the mathematical formalisation of process diagrams like the ones shown above. Joyal and Street [8] showed that the diagrammatic language of boxes and wires is logically coherent with the mathematics of monoidal categories. That is to say, every proof that two expressions are equal in a (symmetric) monoidal category corresponds to a demonstration that the given expressions have isotopic process diagrams. Here, two diagrams are isotopic when the various boxes can be translated and moved over each other without any of the attached wires doubling back on themselves. There are many more types of categories which are coherent with graphical languages. For a summary of the various graphical languages, read [9].

**Definition 2.5** (State, effect, number)**.** We call a morphism $I \to A$ a **state** and a morphism $A \to I$ an **effect**. A morphism $I \to I$ is called a **number**. I'll sometimes represent states and effects using kets and bras respectively. Eg. $|\psi\rangle : I \to A$. Then $\langle\phi| \circ |\psi\rangle = \langle\phi|\psi\rangle : I \to I$ and so on.

**Proposition 2.1.** *Numbers commute: for all $\lambda, \mu : I \to I$ we have $\lambda \circ \mu = \mu \circ \lambda$.*

*Proof.*

$$\begin{aligned} \lambda \circ \mu &= (\lambda \otimes 1_I) \circ (1_I \otimes \mu) = \lambda \otimes \mu \\ &= (1_I \otimes \mu) \circ (\lambda \otimes 1_I) = \mu \circ \lambda \end{aligned} \tag{3}$$

□

We now add a new feature to a category which doesn't have a clear analogue to our example of making breakfast.

**Definition 2.6** (Dagger compact category)**.** A **dagger compact category** (or DCC for short) is a symmetric monoidal category $\mathcal{C}$ equipped with a **dagger** operation on morphisms $(f : A \to B) \mapsto (f^\dagger : B \to A)$ and for each object $A$ we have

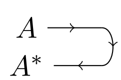
- a **dual** object $A^* \in \mathcal{C}_0$.
- a morphism $\epsilon_A : A \otimes A^* \to I$.
- a morphism $\eta_A : I \to A^* \otimes A$.

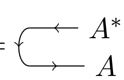
obeying these rules for all $A, B \in \mathcal{C}_0$ and all $f : A \to B$, $g : B \to C$;

1. $(\epsilon_A \otimes 1_A) \circ (1_A \otimes \eta_A) = 1_A$
2. $(1_{A^*} \otimes \epsilon_{A^*}) \circ (\eta_{A^*} \otimes 1_{A^*}) = 1_{A^*}$
3. $c_{A,A^*} \circ \eta_{A^*} = \eta_A$
4. $(f^\dagger)^\dagger = f$
5. $g \circ f^\dagger = f^\dagger \circ g^\dagger$
6. $g \otimes f^\dagger = g^\dagger \otimes f^\dagger$
7. ${c_{A,B}}^\dagger = c_{B,A}$
8. ${\epsilon_A}^\dagger = \eta_{A^*}$

**Definition 2.7** (isometry, unitary)**.** An **isometry** of a DCC is a morphism $f : A \to B$ such that $f^\dagger \circ f = 1_A$. A **unitary** is a morphism $f : A \to A$ where $f^\dagger \circ f = 1_A = f \circ f^\dagger$.

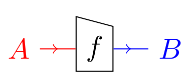
As usual, DCC's also come with a diagrammatic language. We draw $A$ as a wire with an arrow on it, $A^*$ as a wire with an arrow going in the opposite direction and $\epsilon_A$, $\eta_A$ are drawn as U-turns in the wires.

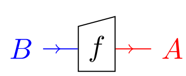
$\epsilon_A =$ 

, $\eta_A =$ 

(4)

The dagger operation is depicted as flipping the image vertically. To take advantage of this, we draw our morphism boxes with a sloping edge so that we can distinguish between $f$ and $f^\dagger$ diagrammatically.

$f =$ 

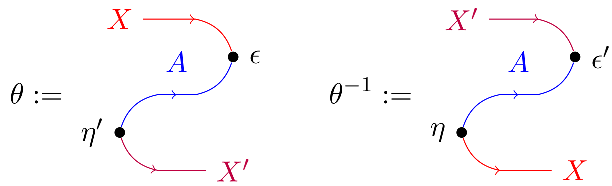
, $f^\dagger =$ 

(5)

For an example of diagrammatic reasoning, consider the following proof;

**Proposition 2.2.** *All duals are unique up to isomorphism. Here, a dual of $A$ is defined as an object equipped with an $\epsilon$ and an $\eta$ which obey the first two rules in definition 2.6.*

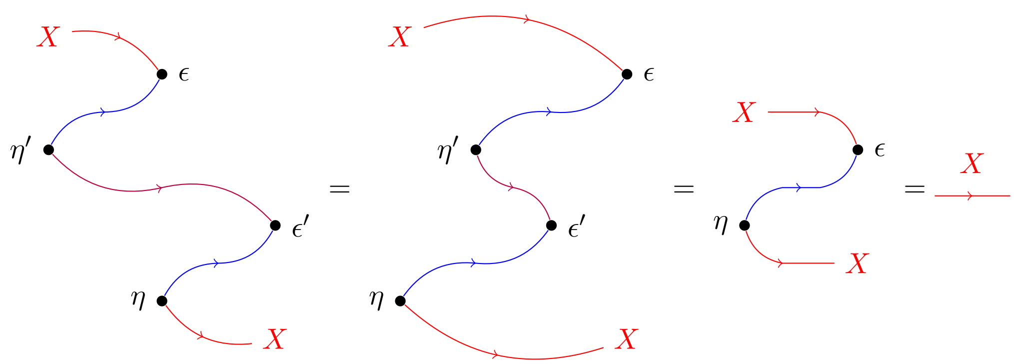
*Proof.* Suppose we had that $X$ and $X'$ are both duals of $A$. Then we have an invertible map $\theta : X \to X'$ defined as;




(6)

Which gives;

$\theta^{-1} \circ \theta =$



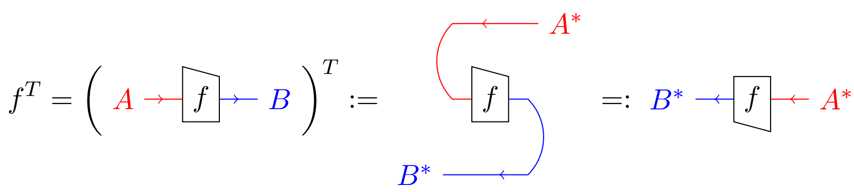

(7)

The first step in this proof is the interchange law for monoidal categories in definition 2.3 and the final two steps follow from the first two rules in definition 2.6. Similar for $\theta \circ \theta^{-1} = 1_{X'}$. So we have constructed an isomorphism between $X$ and $X'$. $\square$

We call these **string diagrams** because of how proofs involving manipulation of them have the appearance of translating and rotating boxes and moving the wires as if they were flexible pieces of string. When this happens we say the diagrams are **isotopic**.

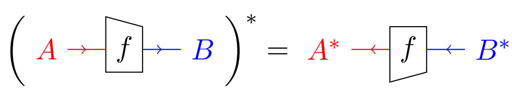
There is a coherence theorem for DCC diagrams. But Coecke and Kissinger claim in their book [10, p. 157] that this theorem 'is widely regarded as folklore'. My understanding of this is that it is assumed to be true because there are lots of very similar theorems about other diagrammatic languages which have been proven (eg. Joyal and Street [8]) and because all string diagram proofs encountered so far do obey the coherence theorem. Fortunately, we do not have to rely on the coherence theorem to perform mathematics because whenever presented with a diagrammatic proof, one can verify by consulting the laws of DCC's that the proof works.

**Definition 2.8** (Transpose and conjugate)**.** Given a morphism $f : A \to B$ in a DCC, define its **transpose** $f^T : B^* \to A^*$ as:



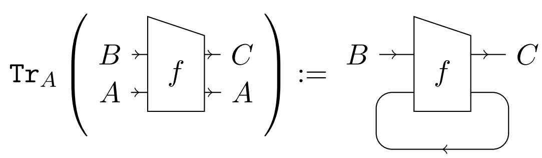

(8)

And define its **conjugate** $f^* := (f^T)^\dagger$.




(9)

**Definition 2.9** (Trace)**.** Given a morphism $f : A \otimes B \to A \otimes C$, the **partial trace** $\mathtt{Tr}_A f$ is defined as;

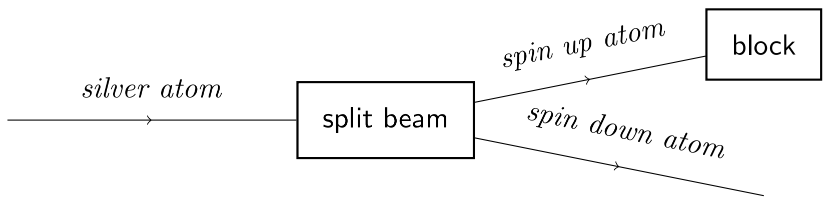



(10)

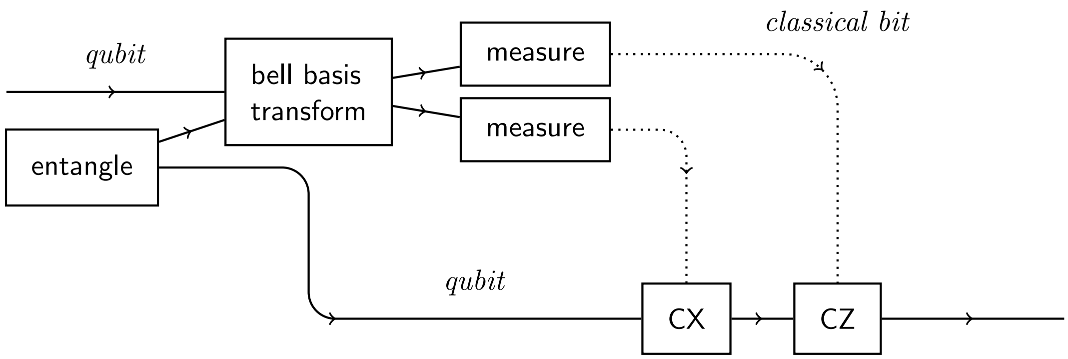
## 3 Science with categories

Let us sketch a rough dogma for how science works. We observe and experience the world and find things that we wish to understand. To understand them we form a theory, which is a model of this thing and a way of mapping between this model and our experience of the world. Once we have a theory, we may look at a science experiment, convert it into the mathematical model and then ask questions about what the model predicts.

This dogma can be implemented in category theory in a fairly clean way. First, note that scientific experiments can often be represented as process diagrams. Consider these two examples from quantum mechanics; a simple Stern-Gerlach setup (11) and quantum teleportation (12).




(11)




(12)

There are obviously lots of caveats to this and we need to be careful that what we draw really is a process diagram and is an accurate representation of the experiment we wish to model. But pressing ahead, since we know that process diagrams are monoidal categories, we can make a monoidal category of science experiments; that is, the category of all possible ways of connecting up the various pieces of lab equipment. Again, defining this rigorously is an open ended problem, but that doesn't mean we can't get something which is useful. Consider researchers into fluid dynamics: while the predictions they make are very accurate, the link from the model to real experiments is usually obvious without resort to formalisation.

Now we need to map these science experiments to our category representing the mathematical model that we postulate as describing the physics. But this is exactly the purpose for which we defined a functor.

Remarkably, categorical quantum mechanics lets us do precisely this. Moreover, it can also take into account the fact that we may not be able to model the experiment with complete accuracy. That is, we can also build model diagrams that include our epistemology.

# 4 Hilbert space

This essay focusses on the epistemology of quantum mechanics rather than the theory for the real underlying physics. We can do this because our theory of the epistemology in the next section requires only one premiss: the physics of quantum mechanics is modelled in a DCC.

Let's unpack this a little. What exactly does 'the physics of quantum mechanics' mean? It may help to consider the example of Newtonian astronomy again. What we mean by 'the physics of astronomy' is a mathematical model that lets us reason about the positions of planets as if we knew everything. That is, without having to worry about measurements, errors and so on. In Newtonian astronomy, that physical theory is 3D Euclidean space and a set of objects which come with masses, time dependent velocities and other properties. We also have differential equations for calculating the time dependence of these properties. Before Newton, we didn't know the differential equations but we knew that the appropriate system to reason about the planets was 3D Euclidean space. The premiss above is similar. We don't need to have our rules about how reality operates precisely pinned down before we can start to make reasonable assumptions about the stage on which the drama will unfold.

Deciding which DCCto use is still an active area of research. For instance in [11, 12], a DCC theory of the physics of quantum mechanics is derived with respect to some information theoretical axioms. Here, for the sake of familiarity, we will use Hilbert space as our DCC.

**Definition 4.1.** We define **FdHilb** as the category whose objects are finite dimensional Hilbert spaces and the morphisms are linear maps between them.

**Proposition 4.1.** *The category of finite dimensional Hilbert spaces* **FdHilb** *is a DCC.*

*Proof.* Take $I$ to be the one dimensional Hilbert space; the complex numbers $\mathbb{C}$ and take $\otimes$ to be the usual tensor product. A map $f : A \to B$ in **FdHilb** is an $n \times m$ matrix where $n = \dim A$ and $m = \dim B$. We set the $\dagger$ operation to be the hermitian adjoint of this matrix. Next, define $A*$ as the set of all linear functionals $A \to I$ and $\epsilon_A : A \otimes A* \to \mathbb{C}$ to be the linear map given by $(|\psi\rangle, \langle\phi|) \mapsto \langle\phi|\psi\rangle$. We then define $\eta_A = {\epsilon_{A*}}^\dagger$ and all of the defining axioms of the DCC follow easily. $\square$

Furthermore, we have a completeness theorem for string diagrams and **FdHilb**:

**Theorem 4.2.** **FdHilb** *is* complete *for string diagrams. That is, we can prove that two string diagrams are isotopic if and only if any assignment of terms in the diagram to morphisms in* **FdHilb** *results in the diagrams evaluating to the same thing.*

*Proof.* See Selinger [13]. □

Our category **FdHilb** has been very successful at describing the quantum mechanical world. This tells us that DCC's are the right way to go in trying to frame quantum mechanics in category theory. While this method of rebuilding all of our language in terms of diagrams may seem redundant since we could just use matrices like everyone else, it offers some conceptual advantages. CQM gives us formal justification to use diagrammatic reasoning, a lot of the proofs and definitions in quantum information and quantum computing have very elegant diagrammatic proofs. Indeed, a very similar notation was developed by Roger Penrose [14] as an alternative to indices in tensor algebra. These diagrams are more than aesthetics, the human brain is extremely well adapted to visual reasoning. As various psychological studies such as [15, 16] have explored, the ability to (appropriately) visualise a problem greatly enhances a person's ability to understand and solve it.

# 5 Adding epistemology

Now we want to add epistemology to our model; we want to account for our incomplete knowledge about the physical situation. Here, rather than using **FdHilb** we will use an arbitrary DCC $\mathcal{C}$ as our underlying category of pure quantum processes. We will construct a new category out of $\mathcal{C}$ that models the physics and takes into account our incomplete knowledge of the situation too. This is called the CPM construction first defined in [17][1].

Consider the following functor;

**Definition 5.1** (Doubling)**.** Suppose $\mathcal{C}$ is a DCC. Define $\mathsf{double}(\mathcal{C})$ as the functor $\mathsf{double}(A) := A \otimes A^*$ and $\mathsf{double}(f : A \to B) := f \otimes f^*$. Define $\mathsf{double}(\mathcal{C})$ to be the category of all objects and morphisms in the range of the $\mathsf{double}$ functor.

**Definition 5.2** (Selinger CPM construction)**.** Now define $\mathbf{CPM}(\mathcal{C})$ as the category with all objects in $\mathsf{double}(\mathcal{C})_0$ and morphisms as the set

$$\mathsf{double}(\mathcal{C})_1 \cup \{(1_B \otimes \epsilon_C \otimes 1_{B^*}) \circ \mathsf{double}(f) \mid f : A \to B \otimes C \in \mathcal{C}_1\} \tag{13}$$

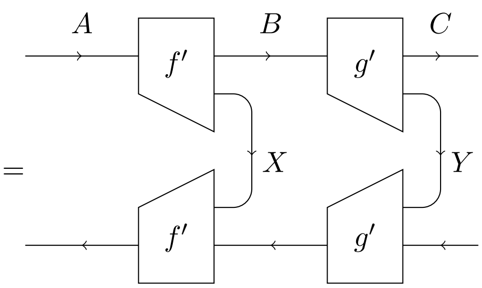
[1] This construction has also found use in linguistics to model ambiguity in the meanings of sentences [18]

That is, the set of all morphisms in $\mathsf{double}(\mathcal{C})$ and all possible ways of applying $\epsilon_A$ to $\mathsf{double}(A)$ objects for all $A \in \mathcal{C}$. Any morphism which can be expressed as $\mathsf{double}(f)$ for some $f$ in $\mathcal{C}$ is called **pure**.

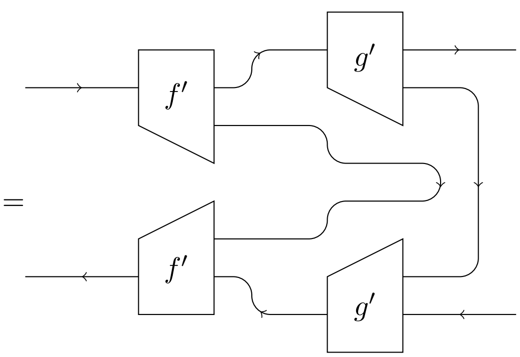
**Theorem 5.1.** $\mathbf{CPM}(\mathcal{C})$ *really is a category. Moreover, it's a DCC.*

*Proof.* Suppose we have $f : \mathsf{double}(A) \to \mathsf{double}(B)$ and $g : \mathsf{double}(B) \to \mathsf{double}(C)$ in $\mathbf{CPM}(\mathcal{C})$. So there is some pure $f' : A \to B \otimes X$ and $g' : B \to C \otimes Y$ in $\mathcal{C}$. Then we compose the morphisms in $\mathbf{CPM}(\mathcal{C})$ as follows:

$g \circ f :=$

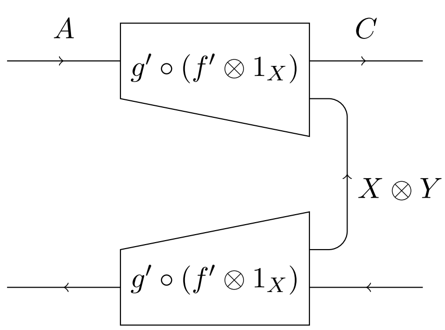



$=$



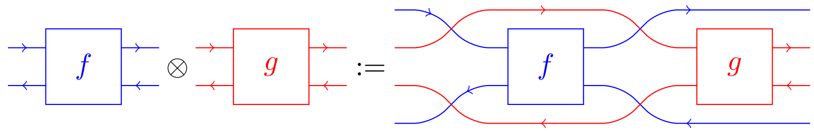

(14)

$=$



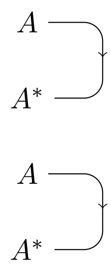

$\in \mathbf{CPM}(\mathcal{C})$ (15)

Which is a morphism of $\mathbf{CPM}(\mathcal{C})$ since $\otimes$ing two discards is a discard. We define $1_{\mathsf{double}(A)} := 1_A \otimes 1_{A^*}$. We define $\otimes$ on $\mathbf{CPM}(\mathcal{C})$ as




(16)

and $I$ as its counterpart in $\mathcal{C}$. We also define the dagger operator as its counterpart in $\mathcal{C}$ and $\epsilon_A : \mathsf{double}(A) \otimes (\mathsf{double}(A))^* \to I$ as $\epsilon \otimes \epsilon$ in $\mathcal{C}$.




(17)

Note that $(\mathsf{double}(A))^* = (A \otimes A^*)^* \cong (A^* \otimes A)$ because one can show that both of these objects are dual to $\mathsf{double}(A)$ in $\mathcal{C}$. One can readily confirm this obeys the laws of a monoidal category and a DCC. $\square$

To see why the CPM construction is a good way to model epistemology, let's look back at our experiments in figures (11) and (12). Consider the box in the Stern-Gerlach experiment (11) labelled 'block'. When an experimenter blocks or discards a part of their experiment, this is the same as simply noting that it will never be used again and so it can be ignored. Is the process of discarding something a physical process? That is, should it considered to be a pure process? We will postulate here that the answer is no, because it is a matter of viewpoint whether or not a part of an experiment can be ignored. Suppose that after the Stern-Gerlach apparatus had split the electrons into beams A and B, these beams were sent to laboratories A and B at opposite ends of the country. An experimenter in A, who doesn't know or care about the work of B, is then justified in modelling her experiment as discarding beam B, and vice versa for experimenter B. The experimenters are not justified in discarding the opposite beams if their experiments depend on the results of the opposite lab. The opposite beam may be discarded even if the two beams are entangled in some way; the results of the beams will be correlated but A has no way of discovering this if she doesn't care about the work of B.

What do we want formally from a discard process? We can use our intuition in the "category of science experiments": producing a pure state and then immediately discarding it should be the equivalent to having never made the state in the first place [10, section 5.2].

**Definition 5.3** (Discard)**.** For any DCC $\mathcal{C}$, define a **discard process** in $\mathbf{CPM}(\mathcal{C})$ as a morphism $\top_A : A \to I$ for each $A \in \mathbf{CPM}(\mathcal{C})$. Such that for any pure state $\psi : I \to A$ in $\mathcal{C}$ where $\psi^\dagger \circ \psi = 1_I$, we have $\top_A \circ \mathsf{double}(\psi) = 1_I$ in $\mathbf{CPM}(\mathcal{C})$. Additionally, $\top_{A \otimes B} = \top_A \otimes \top_B$ for all $A$, $B$ in $\mathbf{CPM}(\mathcal{C})$.

This definition first appears in [19] under the name of $\top$-structure. Using this definition, we can see that one needs $\mathbf{CPM}(\mathcal{C})$ to discuss discards.

**Proposition 5.2.** *In* $\mathbf{FdHilb}$*, for any space* $A$ *with* $\dim(A) > 1$ *there is no process which composes with any state* $\psi : I \to A$ *with* $\psi^\dagger \circ \psi = 1_I$ *to give* $1_I$*. That is, there is no discard process in* $\mathbf{FdHilb}$ *alone.*

*Proof.* Given a space $A$ with $\dim A > 1$ and an ONB $\{|i\rangle\}_i$. We need a map $A \to \mathbb{C}$, but this is $\langle\top| = \sum_i a_i \langle i|$ for some $a_i \in \mathbb{C}$. But $\langle i|i\rangle = 1$ so $\langle\top|i\rangle = 1$ so $a_i = 1$. But then

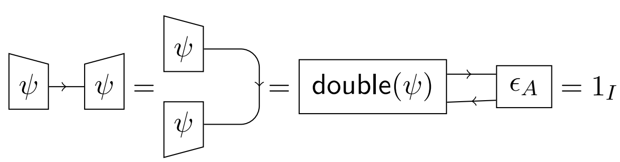
$$\langle\top| \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = \frac{2}{\sqrt{2}} \tag{18}$$

but this should equal 1 if $\langle\top|$ is a discard. □

This isn't too surprising, we already discussed that discarding should not be pure.

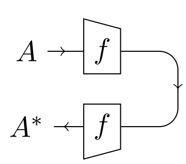
**Proposition 5.3.** $\epsilon_A : A \otimes A^* \to I$ *in* $\mathcal{C}$ *is a discard operation for any object* $\mathsf{double}(A)$ *in* $\mathbf{CPM}(\mathcal{C})$*. If* $\mathcal{C}$ *is* $\mathbf{FdHilb}$*, then this discard is unique.*

*Proof.* Suppose we are given a pure state $\psi : I \to A$ in $\mathcal{C}$ with $\psi^\dagger \circ \psi = 1_I$. Then we have that



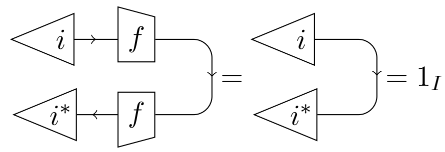
(19)

For uniqueness in $\mathbf{CPM}(\mathbf{FdHilb})$, suppose that one finds another operation $T' : \mathsf{double}(A) \to I$ in $\mathbf{CPM}(\mathbf{FdHilb})$. Then $T'$ has this form for some $f : A \to B$ in $\mathbf{FdHilb}$:

$T'_A =$ 

(20)

But then for an ONB $\{|i\rangle\}_i$, we have;

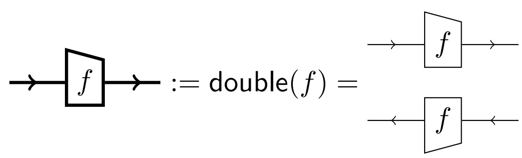



(21)

because $T'_A$ is a discard. So $f$ is an isometry since the ONB spans the space $A$ and so $T'_A = \epsilon_A$. $\square$

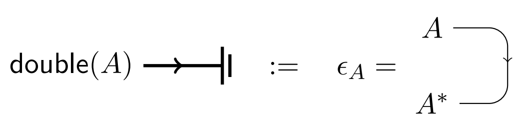
To save some space, we will now adopt a convention for discussing diagrams in $\mathbf{CPM}(\mathcal{C})$.

**Definition 5.4.** Suppose $f$ is some morphism in $\mathcal{C}$. Then we write;



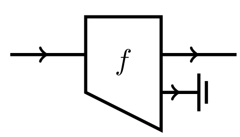

(22)

Also, discarding, $\epsilon_A$, is represented as;




(23)

This means that any morphism in $\mathbf{CPM}(\mathcal{C})$ can be expressed as:



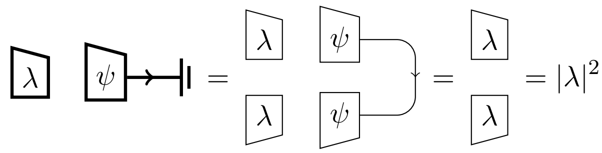

(24)

for some $f$ in $\mathcal{C}$.

So we can choose to write diagrams in $\mathbf{CPM}(\mathcal{C})$ either as 'doubled, thin wire' or 'thick wire' diagrams. Why did we bother constructing this doubling when we could have just added a discard operation to $\mathcal{C}$ to begin with? The answer is that using the doubling construction lets us maintain a clear separation between the category modelling the physics $\mathcal{C}$, and the category modelling both the physics and the epistemology $\mathbf{CPM}(\mathcal{C})$.

Next, we need to define a subcategory of $\mathbf{CPM}(\mathcal{C})$ that contains only the diagrams that represent things that we can build in a laboratory.

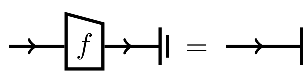
At the moment, in **FdHilb** for $\lambda$ some complex number and $|\psi\rangle$ some pure normalised state, $\lambda \cdot |\psi\rangle$ is a pure state but if we try and discard the state in **CPM**(**FdHilb**):



(25)

Our diagram picks up a factor of $|\lambda|^2$. States that pick up a factor other than 1 can't be implemented in the lab, because we demand that discarding a state is the same as never having the state to start with. So we need to modify **CPM**(**FdHilb**) to prevent these 'bad' states from being allowed. This definition is taken from [10, section 5.2] which is inspired by the work of Chirabella [12].

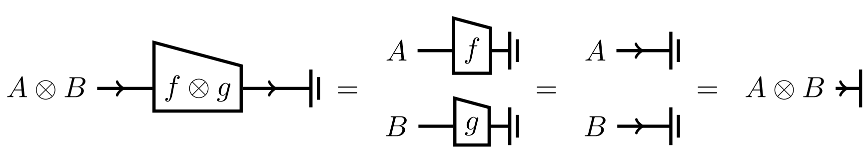
**Definition 5.5** (Causal)**.** Define **Causal**($\mathcal{C}$) to be the category with the same objects **CPM**($\mathcal{C}$) but it only has the morphisms in **CPM**($\mathcal{C}$) that are **causal**. A morphism $f$ is causal when



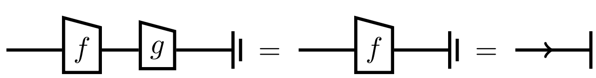
(26)

**Proposition 5.4. Causal**($\mathcal{C}$) *is a monoidal category.*

*Proof.* Since **Causal**($\mathcal{C}$) has a subset of the morphisms in **CPM**($\mathcal{C}$) which is a monoidal category by theorem 5.1, it suffices to note that $I$ and $1_{\mathsf{double}(A)}$ are in **Causal**($\mathcal{C}$) and to show that $\otimes$ and $\circ$ preserve the causal property as shown by (27) and (28).



(27)



(28)

□

**Proposition 5.5.** *For all $A \in$ $\mathsf{double}(\mathcal{C})$, there are no effects $A \to I$ in* **Causal**($\mathcal{C}$) *other than discarding.*

*Proof.* First, we have that $\top_I : I \to I$ commutes with $1_I$ because it is a number. So $\top_I = \top_I \circ 1_I = 1_I$. Now given $x : A \to I$ in **Causal**($\mathcal{C}$) we have $x = 1_I \circ x = \top_I \circ x = \top A$. □

What have we achieved so far? Given a DCC $\mathcal{C}$, we have constructed two categories $\mathbf{CPM}(\mathcal{C})$ and $\mathbf{Causal}(\mathcal{C})$. The claim of this formulation is that all quantum mechanical experiments including epistemology have a corresponding diagram in $\mathbf{Causal}(\mathcal{C})$.

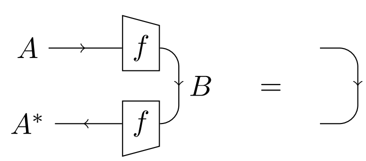
So, in short, $\mathbf{Causal}(\mathcal{C})$ represents your state of knowledge about the experiment. To justify this claim, we will examine what this interpretation says when we set $\mathcal{C}$ to be **FdHilb**. We will see that in this category, the claim corresponds exactly to our orthodox formulation of QM.

## 5.1 CPTP maps from Causal(FdHilb)

Let's investigate what **Causal**(**FdHilb**) looks like. The arguments in this section are based on [10, section 5.2].

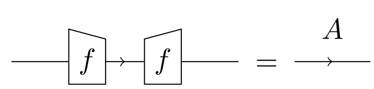
**Proposition 5.6.** *The pure processes in* **Causal**(**FdHilb**) *correspond to isometries.*

*Proof.* A pure causal process $f$ is a morphism $f : A \to B$ in **FdHilb** that obeys the following equation:




(29)

which is equivalent to

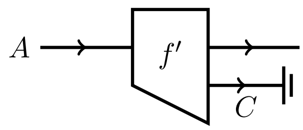


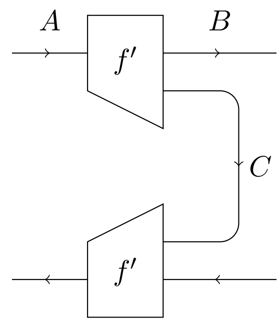

(30)

which is exactly the condition for $f$ to be an isometry. □

**Theorem 5.7.** *The morphisms in* **Causal**(**FdHilb**) *are the completely positive, trace preserving maps in* **FdHilb**.

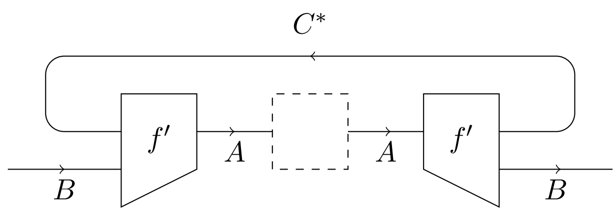
*Proof.* Take $f : \mathsf{double}(A) \to \mathsf{double}(B)$ in **Causal**(**FdHilb**). Then we have for some $f' : A \to B \otimes C$

$f =$ 

$=$ 

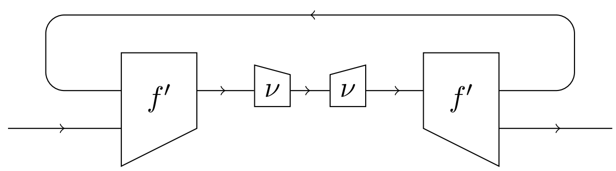
(31)

This can be rearranged to

$f[-] =$ 

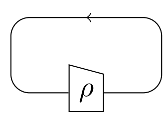
(32)

This is a linear map from $A \to A$ to $B \to B$ in **FdHilb** (the $A \to A$ map is inserted into the dotted box). Suppose that $\rho : A \to A$ in **FdHilb** is positive. This means that $\rho = \nu^\dagger \nu$ for some $\nu : A \to C$. So then

$f[\rho] =$ 

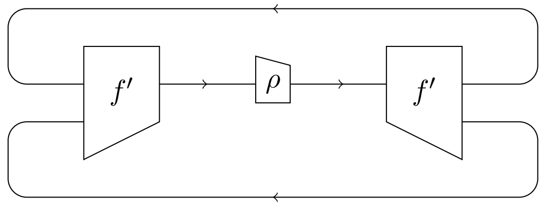
(33)

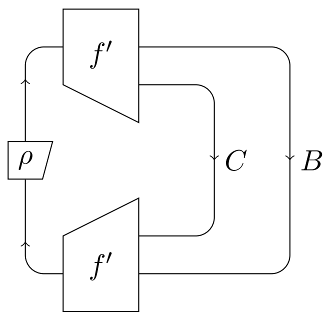
Which is also positive because the diagram is symmetric in the $x$ direction. So $f[-]$ is completely positive. Next, suppose $\rho$ has trace $x$. This means

$x = \mathtt{Tr}_A(\rho) =$ 

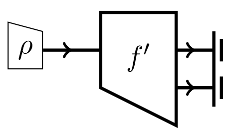
(34)

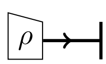
so

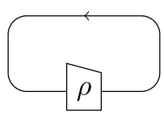
$\mathtt{Tr}_B(f[\rho]) \quad =$ 

$=$ 


(35)

$=$ 

$=$ 

$=$ 

$= \quad x$

So $f$ is trace preserving.

Going in the other direction, we know that all CPTP maps can be written as an isometry $f' : A \to B \otimes C$ where $f[\rho] = \mathtt{Tr}_C(f' \circ f'^{\dagger})$. But this is exactly diagram (32), which can be rearranged to (31) which is a member of **CPM**(**FdHilb**). Causality follows from the isometry of $f'$. □

**Corollary 5.8.** *States in* **Causal**(**FdHilb**) *are the density matrices in* **FdHilb**.

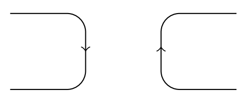
*Proof.* For a causal process $\rho \in \mathbb{C} \to \mathsf{double}(A)$, we have that the map $\rho : (\mathbb{C} \to \mathbb{C}) \to (A \to A)$ is CPTP. But that is isomorphic to $\rho : (A \to A)$ so completely positive just means that $\rho$ is positive and trace preserving means $\mathtt{Tr}\rho = 1$. $\square$

Now we can see the power of the **Causal**($\mathcal{C}$) construction. We have derived the language of mixed states and CPTP maps from the simple intuition that discarding a state should be the same as removing that state from the diagram. This is an improvement over orthodox quantum theory, where mixed states and CPTP maps are simply postulated to be the way of modelling quantum processes.

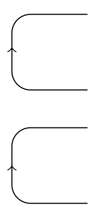
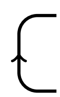
Let's look at some specific examples of some states in **Causal**(**FdHilb**).

An entangled pair of qubits can be represented as a pure density matrix $\rho : Q^2 \to Q^2$ in **FdHilb**, namely

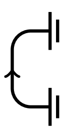
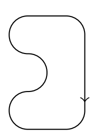
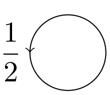
$$\rho := \frac{1}{2}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)(\langle 0| \otimes \langle 0| + \langle 1| \otimes \langle 1|) \qquad (36)$$

$= \frac{1}{2}$ 

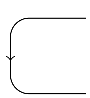
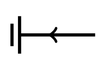
This corresponds to the pure state in **CPM**(**FdHilb**)

$\frac{1}{2}$  $= \frac{1}{2}$  (37)

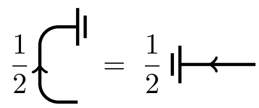
So the maximally entangled state is $\frac{1}{2}\,\mathsf{double}(\eta)$. One can confirm that this is causal:

$\frac{1}{2}$  $= \frac{1}{2}$  $= \frac{1}{2}$  $= \frac{1}{2}\,\mathtt{Tr}(1_Q) \; = \; 1$ (38)

We know that the maximally mixed density matrix of a system is represented as the identity matrix divided by the dimensionality of the system. This means that, as a state in **Causal**(**FdHilb**), the maximally mixed state for the system $A$ is

$\frac{1}{\dim A}$  $= \frac{1}{\dim A}$  (39)

From this, we can see that discarding one of the qubits in an entangled system is the same as having one qubit in the mixed state;



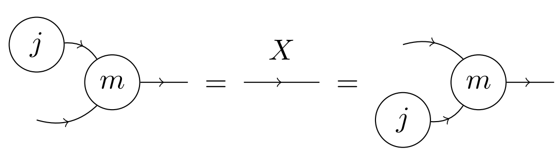

(40)

So throwing away one entangled qubit is the same as knowing nothing about the other. This is a hint at the controversies to come. In most other formulations of QM, we would say that the qubits are entangled whether or not we choose to throw one of them away.

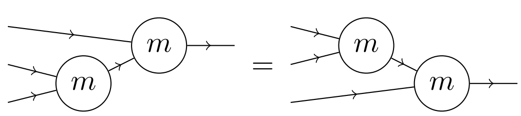
# 6 Classical Information

What does it mean to perform a measurement? From our previous experience with quantum mechanics, this is the process of taking a quantum system and somehow extracting a value from it. This measured value is different from the quantum system because we can copy it. We take this duplicability of the measured information to be our primary motivation behind defining a category including both classical and quantum information. To make this notion precise, we are going to introduce Frobenius algebras. This has a fairly obtuse definition, but with the right intuitions they become very easy to understand. The theorems and definitions in this section follow the work of Coecke and Pavlovic in [20].

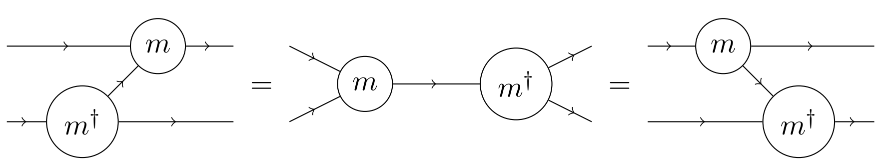
**Definition 6.1** (Frobenius algebra)**.** Given a DCC $\mathcal{C}$, a **dagger Frobenius algebra** is an object $X$ and two morphisms $j : I \to X$, $m : X \otimes X \to X$ which obey the following rules:




(41)

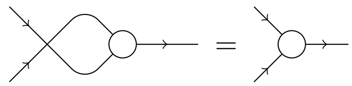



(42)



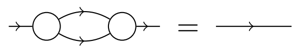

(43)

From now on we can omit the labels from the circles. These circles are often called ‘spiders’. We say the Frobenius algebra is **commutative** when



(44)

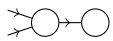
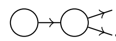
and **special** when

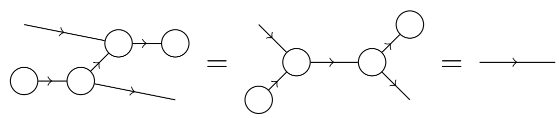


(45)

‘Dagger special commutative Frobenius algebra’ is shortened to †**SCFA**.

**Proposition 6.1.** *If $X$ is a Frobenius algebra, it is self dual.*

*Proof.* We claim that $\epsilon =$  and $\eta =$ . By axioms (43) and (41) we have

$(\epsilon \otimes 1_X) \circ (1_X \otimes \eta) =$  $= 1_X$

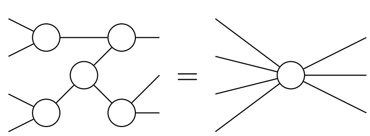
(46)

similar for $(1_X \otimes \epsilon) \circ (\eta \otimes 1_X) = 1_X$. □

So from now on we can drop the little arrows on all of the wires. The motivation behind †**SCFA**’s is that they represent a connection between objects. For instance, the wires used in electronics schematics merely provide a description of which terminals are in electrical contact. The actual topology and routing of the wires doesn’t matter. The electrical wires and the junction form a †**SCFA**.

**Theorem 6.2** (†**SCFA** normal form)**.** *Take $X^{\otimes n}$ to mean $X$ $\otimes$ed with itself $n$ times. Suppose that we are given a morphism $\chi : X^{\otimes n} \to X^{\otimes m}$ which is formed from a connected network of $X$ spiders for $X$ some* †**SCFA**. *All such $\chi$ are equal. This means that we can change our graphical notation of* †**SCFA***’s to ’spider diagrams’ with arbitrary legs, knowing that we can always rewrite it in the canonical form.*

*Proof.* See Lack’s paper [21]. □



(47)

**Definition 6.2** (Classical Object)**.** Given a DCC $\mathcal{C}$, we define a **classical object** as a †**SCFA** in $\mathcal{C}$.

**Proposition 6.3.** *If $X$ is a classical object in $\mathcal{C}$ then* $\mathsf{double}(X)$ *is a classical object in* $\mathbf{CPM}(\mathcal{C})$.

*Proof.* Follows immediately from the fact that the axioms (41, 42, 43) for $\mathsf{double}(X)$ are just the axioms in $X$ $\otimes$ed with its conjugate. $\square$

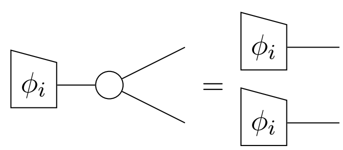
Let's see what a classical object looks like in **FdHilb**:

**Theorem 6.4.** *Suppose we had a space $A$ and an ONB (orthonormal basis) $\{|i\rangle\}_i$ in* **FdHilb**. *Then we can make the classical object;*

$$m = \sum_i |i\rangle \left(\langle i| \otimes \langle i|\right) \qquad j = \sum_i \langle i| \tag{48}$$

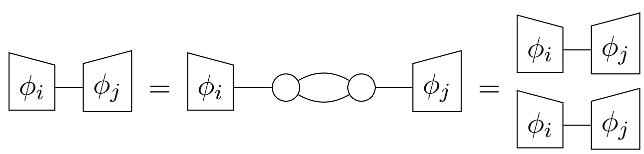
*In fact, each classical object in* **FdHilb** *defines an ONB.*

*Proof.* It is easy to verify that $m, j$ obey the axioms of a †**SCFA**. In the other direction, define the set $\{|\phi_i\rangle\}$ where each $|\phi_i\rangle$ is non-zero and obeys



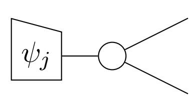

(49)

We claim that this forms an ONB. First, they are orthonormal because



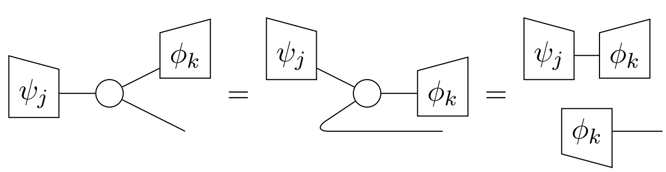

(50)

So $\langle\phi_i|\phi_j\rangle = \langle\phi_i|\phi_j\rangle \cdot \langle\phi_i|\phi_j\rangle$ and therefore $\langle\phi_i|\phi_j\rangle = 1$ or $0$. The positive-definiteness of Hilbert spaces gives $\langle\phi_i|\phi_i\rangle \neq 0$, so $\langle\phi_i|\phi_i\rangle = 1$. But this means that if $\langle\phi_i|\phi_j\rangle = 1$ then $i = j$. So $\langle\phi_i|\phi_j\rangle = \delta_{i,j}$. To show it's a basis, suppose that there is some non-zero, non-empty set $\{|\psi_i\rangle\}$ that is orthonormal to all the $|\phi_i\rangle$ which together form an ONB for the space. Then we can write



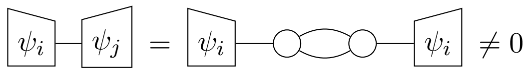

$$= \sum_{p,q} b_{pqj} \, |\psi_p\rangle \, |\psi_q\rangle \tag{51}$$

because if a $|\phi_k\rangle$ component was present, we get

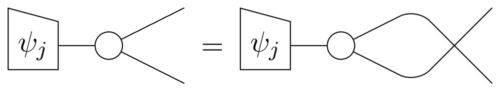



$$= 0 \tag{52}$$

Next, note that;



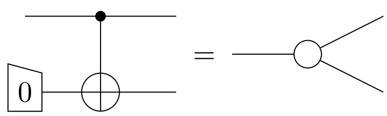
(53)



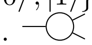
(54)

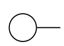
so $b_{pqk}$ is not zero everywhere and is symmetric under any permutation of its indices. This means we can construct an orthonormal set $\{|\chi_i\rangle\}$ out of the $\{|\psi_i\rangle\}$ such that $|\chi_i\rangle$ obeys (49). So $|\chi_i\rangle \in \{|\phi_i\rangle\}_i$ so $\{|\psi_i\rangle\}$ was empty after all. $\square$

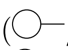
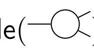
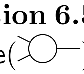
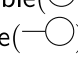
How would one implement a classical object in real life? Suppose we had a system involving 2 qubits, one can quickly verify that $CX$, the quantum controlled NOT gate, is unitary. So it is a pure, causal quantum process. But then applying a $|0\rangle$ state to one of the inputs we have the following causal map.



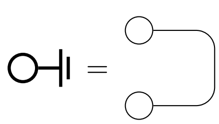
(55)

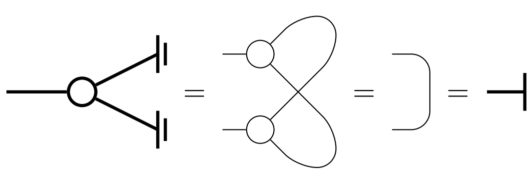
One can then verify that this obeys the laws of a †**SCFA**on a qubit. This is the classical object associated with the $\{|0\rangle, |1\rangle\}$ basis. This gives us an interpretation of classical objects in **FdHilb**.  is the process that takes a state and copies it relative to a certain basis. Note that a classical object does not copy all quantum states but it does copy basis states. Also,

 $= \sum_i |i\rangle$ (56)

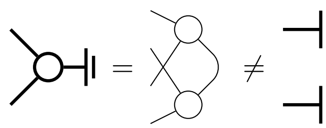
**Proposition 6.5.** *Both* double($/\sqrt{\dim X}$) *and* double() *are in* **Causal**(**FdHilb**) *but* double() *and* double() *are not.*

*Proof.*

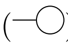
 $= \sum_i \sum_j \langle i|j\rangle = \dim X$ (57)



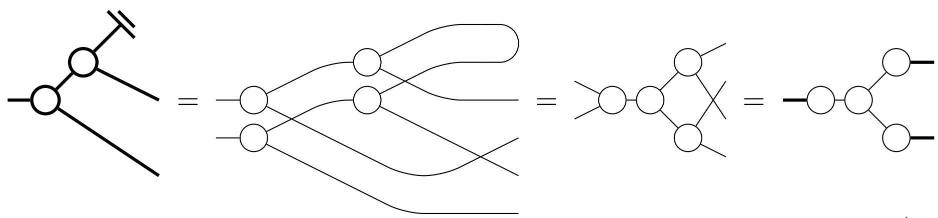
(58)



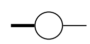
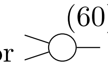
(59)

and double() is an effect, so its not in **Causal**(**FdHilb**) by proposition 5.5. □

Suppose we had a diagram representing some process on a classical object double($X$) in **CPM**($\mathcal{C}$). This looks like a twinned network of wires and spiders. Suppose further that we discard one of the branches somewhere.

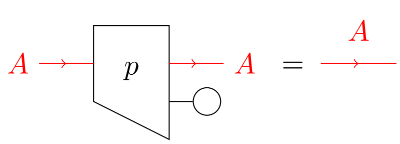


(60)

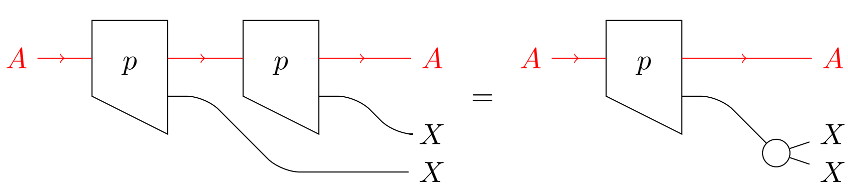
That last diagram involves a new diagram  which is short for  : double($X$) $\to X$. The introduction of a discard into the system causes all the connected double wires to **collapse** into a single wire diagram. We call this collapsed diagram a **classical diagram**.

Non-classical diagrams are very fragile. We only need to discard one of the connected branches for our diagram to become classical. That is, our quantum diagrams only work if we are sure that we are modelling the entire quantum process accurately without discarding any branches of the diagram.

**Definition 6.3** (Projector valued spectrum)**.** Let $\mathcal{C}$ be a DCC and $X$ be a classical object. A **projector valued spectrum** (PVS) of $X$ is an object $A$ and a morphism $p : A \to X \otimes A$ which satisfies the following diagrams;




(61)



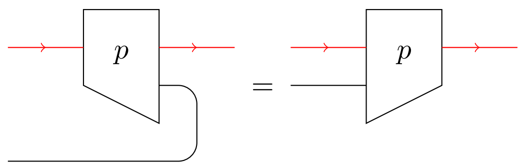

(62)

Categorically minded readers should note that a PVS is an object in the Eilenberg-Moore coalgebra category created by the comonad ($X \otimes -$, $\otimes -$, $\otimes -$), this is also called the category of $X$ comodules.

**Proposition 6.6.** *For all classical objects $X$, the morphism is a PVS of $X$.*

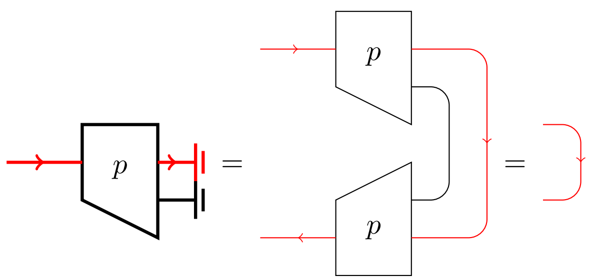
*Proof.* Immediate from (41) and (42). $\square$

**Proposition 6.7.** *Suppose we have a PVS $P$ in $\mathcal{C}$. Then* $\mathsf{double}(P)$ *is in* $\mathbf{Causal}(\mathcal{C})$ *if and only if*



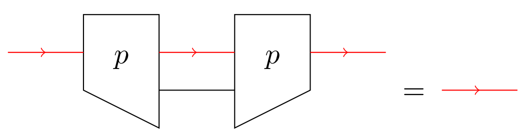
(63)

*Proof.* Suppose $p$ is causal. Then

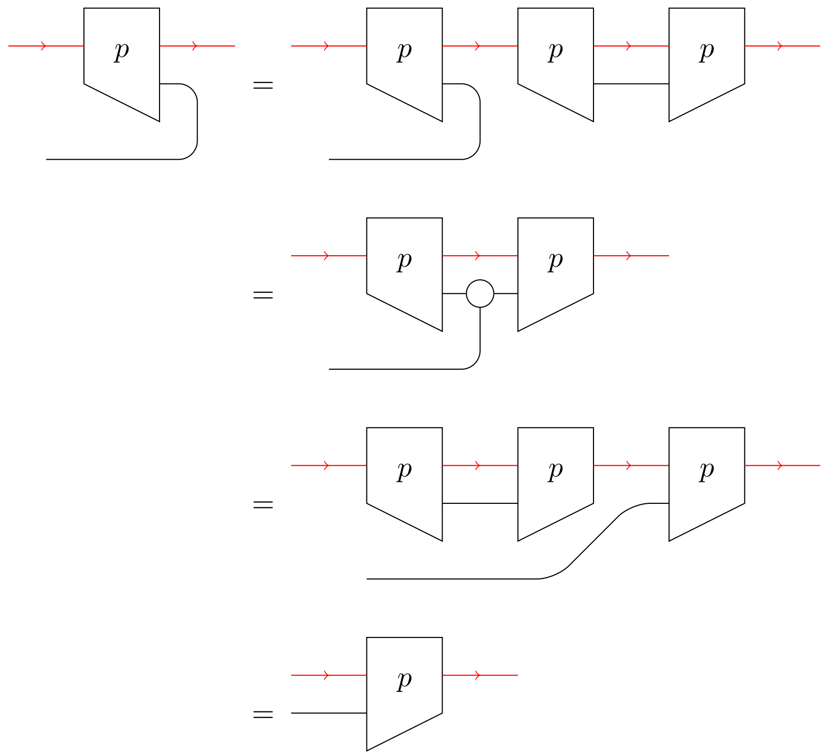


(64)

so

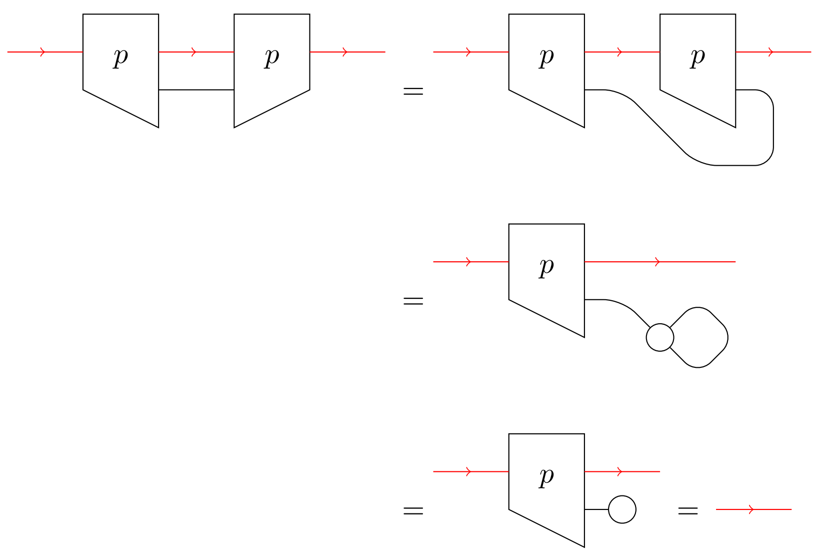


(65)




(66)

Going the other way;




(67)

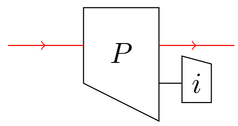
□

**Theorem 6.8.** *Suppose we had a PVS $P$ for* $\dagger$**SCFA** *$X$ in* **FdHilb** *such that* $\mathsf{double}(P)$ *is in* **Causal**(**FdHilb**)*. Then each $P$ corresponds to a complete family of mutually orthogonal projectors. That is, a set*

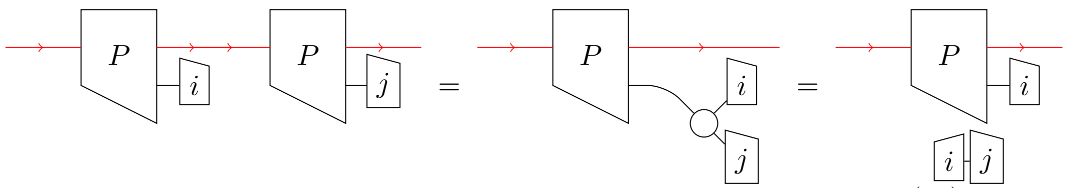
$$\{P_i : A \to A \mid 0 < i < \dim X\} \tag{68}$$

*such that $P_i \circ P_j = \delta_{i,j} P_i$ and $\sum_i P_i = 1$.*

*Proof.* Given a PVS $P$ of $X$, we have by theorem 6.4 an ONB $\{|i\rangle\}_i$. So define

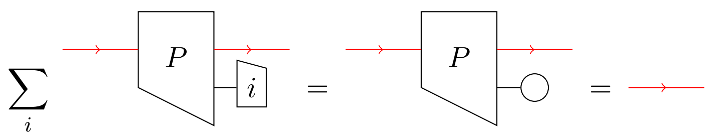
$P_i :=$  (69)

By proposition 6.7 we have (63). Then we have $P_i \circ P_j = \delta_{i,j} P_i$ because

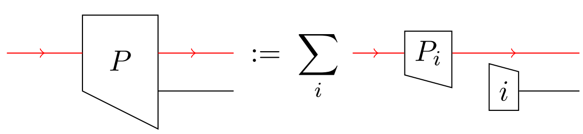


(70)

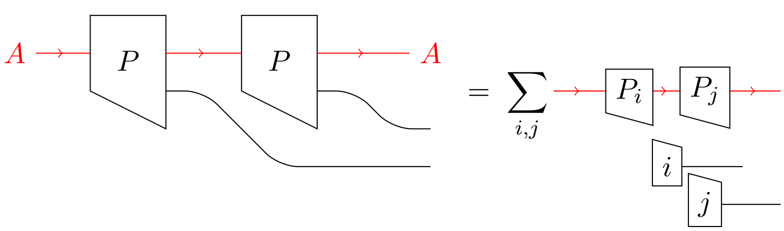
and $\sum_i P_i = 1$ so

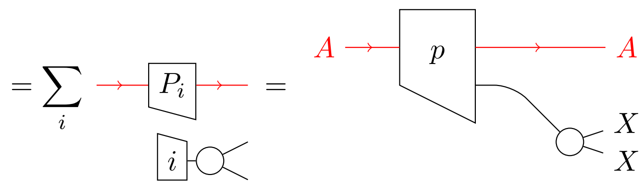
 (71)

So each causal PVS defines a complete family of mutually orthogonal projectors. Going the other way; suppose we had a complete family of $n$ mutually orthogonal projectors $P_i$. Define $X$ as some $n$ dimensional space and define $P$ as

 (72)

We have that $P$ is a PVS:

 (73)



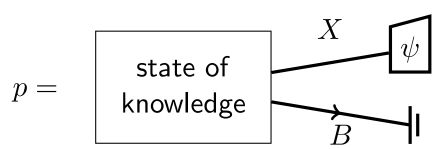
(74)

Causality of double$(P)$ is demonstrated by expanding (65) in a similar way. □

# 7 Alternative postulates of quantum mechanics

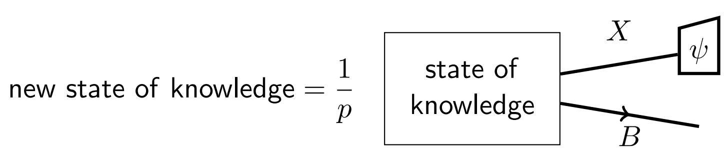
We can now state the postulates of this new formulation of quantum mechanics.

- An experimenter's state-of-knowledge about an experiment in quantum mechanics can be mapped to a state diagram in **Causal**(**FdHilb**).
- Physical processes (a process in reality) are processes $f$ in **FdHilb** such that double$(f)$ is causal.
- The experimenter represents her ignorance of the fate of a particular state with a discard operation.
- The experimenter calculates the probability $p$ of measuring a value $|\psi\rangle$ on a classical object $X$ by calculating the following diagram;



(75)

- If the experimenter learns that a classical object $X$ has value $|\psi\rangle$ (in the ONB associated with $X$), the state-of-knowledge diagram is updated to



(76)

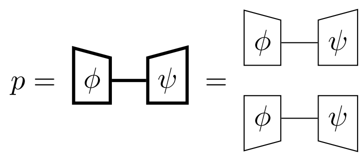
The factor of $1/P$ is needed to ensure the diagram remains causal.

The question of what it means for the experimenter to 'learn' a value is intuitive but philosophically slippery. We will return to the problem later.

**Proposition 7.1.** *If the* state of knowledge *of (75) is in* **Causal**(**FdHilb**) *then* new state of knowledge *of (76) is in* **Causal**(**FdHilb**) *too.*

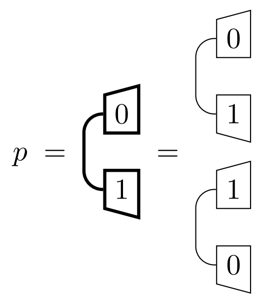
*Proof.* Discarding new state of knowledge gives $\frac{p}{p} = 1$. □

We ought to confirm that the probability calculation actually works. Suppose that we know that a particular quantum state is pure and in $|\phi\rangle$. Then calculating the probability that we will measure it to be $|\psi\rangle$ is




$$= \langle\phi|\psi\rangle \cdot \langle\phi|\psi\rangle^* = |\langle\phi|\psi\rangle|^2 \qquad (77)$$

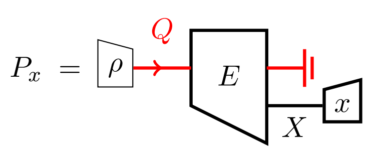
Now suppose that we have two entangled qubits, what is the probability that we measure $|1\rangle$ on one qubit and $|0\rangle$ on the other?




$$= 0 \qquad (78)$$

In general, the probabilities will always be the same as in the orthodox formulation. Here we will prove it for PVM's. Recall that a PVM measurement is defined by a group of complete mutually orthogonal projectors $E_x : Q \to Q$ for $x$ in some finite set with dimension $N$. But this is precisely a causal PVS on a classical $N$ dimensional object by theorem 6.8. Also recall that the probability of measuring $x$ given a initial density matrix $\rho : Q \to Q$ is given by $P_x = \mathtt{Tr}(E_x \circ \rho)$. We also have that the new state after measuring the value $x$ is $E_x \rho E_x / P_x$.

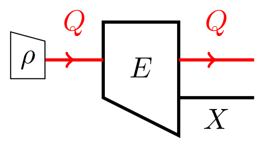
**Theorem 7.2.** *The probability $P_x$ of a measurement of a PVM $E$ is the same as the evaluation of the following probability.*




(79)

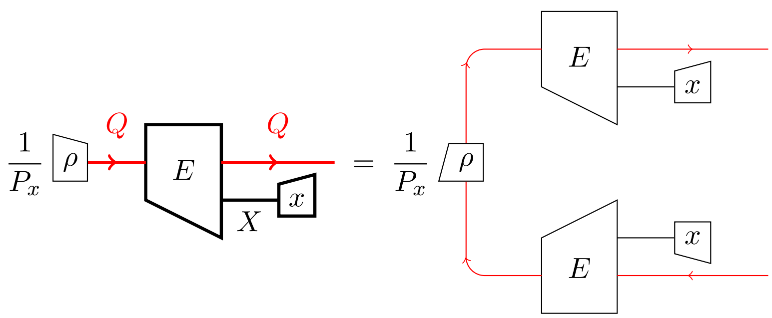
*Which equals $\mathtt{Tr}(E_x \circ \rho)$. Additionally, the new state after the measurement is equal to $E_x \rho E_x / P_x$.*

*Proof.* Our original state-of-knowledge is:




(80)

And after learning that the state of the object is $\langle x|$ we have




(81)

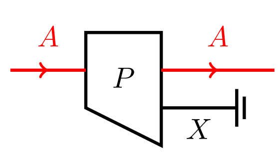
Which is equivalent to $E_x \rho E_x / P_x$. ☐

This is not quite the end of the story for quantum measurement, because we know that sometimes quantum systems behave classically even if the experimenter isn't aware of the result of the measurement.

## 8 Decoherence and observer-less measurement

Decoherence is often sold as that mysterious process when something other than you measures a quantum object, causing the state to collapse to a classical one without you, the experimenter, learning about the value it collapsed to. CQM offers a more precise definition which we draw from [20].

**Definition 8.1** (Decoherence)**.** Given an object $A$ in $\mathbf{CPM}(\mathcal{C})$, a **decoherence process** $A \to A$ is defined as a process that can be written as;



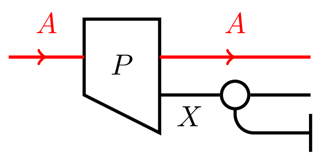

(82)

for some classical object $X$ and a PVS $P : A \to A \otimes X$.

Hence, decoherence is an epistemological phenomenon. It occurs because our state-of-knowledge ignores the fate of the classical object after the PVS occurs. Also note that decoherence is dependent on the classical object.

A more general case of decoherence is quantum measurements. The intuition here is that a measurement is simply the process of making a quantum system interact with a classical object and then discarding a copy of this object [20].

**Definition 8.2** (Observerless measurement)**.** Let $\mathcal{C}$ be a DCC and let $X$ be a classical object in $\mathbf{CPM}(\mathcal{C})$ and $P : A \to A \otimes X$ be a PVS of $X$. Then define a **P-measurement** in $\mathbf{CPM}(\mathcal{C})$ as

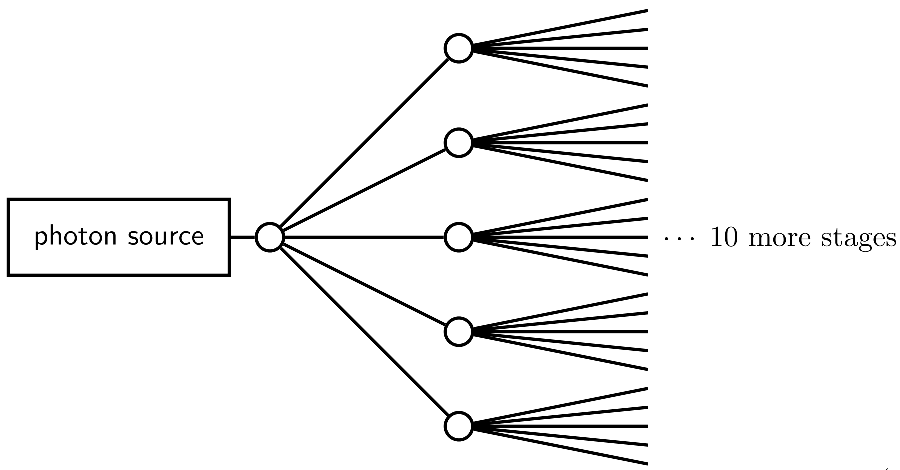



(83)

That is, a measurement of a quantum system is a projection followed by a copy, followed by a discard on one of the copies. Let's interpret this informally and then work through some examples formally.

Note that these collapses don't occur in the physics (that is, the world 'out there') because it only happens when an epistemological discard is present. This means that the collapse happens in our *state-of-knowledge* about the experiment. This concept is critical to the quantum Bayesian way of thinking. The diagrams we have been reasoning with are in $\mathbf{Causal}(\mathcal{C})$. This category was defined to encapsulate our state-of-knowledge of an experimental setup. Once a discard is introduced, our state of knowledge about the experiment is not complete because we don't know what happens to the state we discarded. We already agreed that discarding is not a physical phenomenon, so the collapsing must be epistemological. This is controversial because it means that two experimenters with different knowledge about an experiment may disagree about whether the system is behaving in a classical or quantum fashion.

So, to summarise, a measurement is defined as a process where a quantum state is copied in such a way that you, the experimenter, lose track of some of the copies. In most physics experiments there are processes that involve copying a value millions of times. For instance, in a photomultiplier tube, the information about the presence of a photon is copied into a cascade of millions of electrons. We only need to fail to understand the fate of a single electron to result in a discard of that photon in our state-of-knowledge diagram. Once that discard is present, the information about the detection of the photon becomes classical. This photomultiplier can be seen informally as the following diagram where each wire is the quantum information representing "has a photon been detected?"




(84)

Our state of knowledge only needs to discard one of these $5^{12}$ electrons carrying the information about the photon detection to get a wavefunction collapse as depicted in diagram (60).

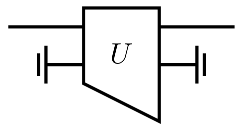
Another important way decoherence can occur is when a quantum state associates with a large, chaotic process with a lot of uncertainty. A quantitative sketch of how this might occur is as follows. Suppose that we are storing the value of a qubit $Q$ in some system which creates a small electric field if in the $|1\rangle$ state. Next, suppose there is some qubit $S$ whose energy levels depend on the strength of this electric field. Let us say that $S$, which initially we do not know the state of, is brought into the range of $Q$'s electric field for a small time $t$ and then $S$ is moved away and is discarded. The Hamiltonian when both of the qubits are together is;

$$H := \begin{bmatrix} E & 0 & 0 & 0 \\ 0 & E & 0 & 0 \\ 0 & 0 & (E-\epsilon) & 0 \\ 0 & 0 & 0 & (E+\epsilon) \end{bmatrix} : Q \otimes S \to Q \otimes S \qquad (85)$$

So the unitary transformation corresponding to $t$ seconds together $U = \exp(itH/\hbar)$. The process $P_{t,\epsilon} : \mathsf{double}(Q) \to \mathsf{double}(Q)$ in $\mathbf{Causal}(\mathbf{FdHilb})$

representing the entire encounter is

$$P_{t,\epsilon} := \frac{1}{2}$$




$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{86}$$

Where $c := \cos(t\epsilon/\hbar)$. If many $P$ processes happened, we would represent this as a chain of $P$ processes. Since the $c$ terms always have magnitude $\leq 1$ then as more and more $P$ processes occur, the two middle diagonal terms will approach zero. When this happens our process is the same as decoherence. If $\epsilon$ is large, then decoherence may occur after one or two $P$ processes. If $\epsilon t \ll \hbar$ then decoherence will happen slowly. Suppose that $t$ and $\epsilon$ are fixed and we let $P_{t,\epsilon}$ processes happen at some frequency $f$. We can then show that the qubit decays to a decohered classical bit with time constant of order

$$\tau \propto \frac{1}{ft^2\epsilon^2} \tag{87}$$

This number is often called the **relaxation time** of the qubit.

When might such a succession of $P$ processes occur in real life? A clear example would be if $Q$ was surrounded by a gas of polar molecules but the basic idea is quite general. The point is that even though each $P$ process only makes a tiny change to our state-of-knowledge of the qubit, over a huge number of processes a measurement occurs even though we can't point to the single qubit which was discarded to cause the collapse. Note also that the collapse happened with reference to a certain basis of the qubit. This is because we understood the way in which $Q$ was influencing $S$ was basis dependent.

There is still one annoying issue with quantum Bayesianism. In a 'true' measurement something extra happens: you, the experimenter, learn about whether an instrument display - which notifies you if a photon was detected - is on or off. If you learn the value of the classical object then decoherence *must* have occurred. This is because the value was copied into your brain which is certainly not something you understand well enough to accurately track the propagation of individual quantum mechanical processes.

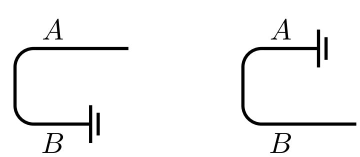
So what does this exploration tell us philosophically? The main message is that we can explain the phenomena of wave function collapse, decoherence and the link between classical and quantum information as being an epistemological phenomenon. They result from the fact that our model of a given experiment can't account for the whole universe so we have to discard information. I will argue that this points strongly to the interpretation of quantum Bayesianism. We will now illustrate quantum Bayesianism further with some examples.

## 9 Entanglement and Teleportation

We first consider the situation where two qubits are entangled and then are moved a large distance away from each other. Alice accompanies qubit A and Bob accompanies qubit B.

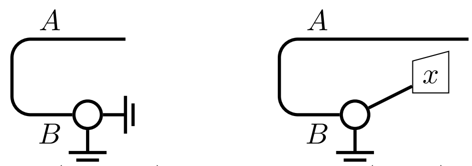
In the Copenhagen interpretation of QM, we would model the two qubits as being in the $|0\rangle|0\rangle + |1\rangle|1\rangle$ state. Then, when Alice measures her qubit, this state suddenly collapses. This is often called 'spooky action at a distance'. At first it seems to violate the key postulate of relativity that no information can travel faster than the speed of light. But while the collapse of the state does indeed travel faster than light, it is impossible for Bob to learn whether Alice caused the state to collapse or not until they meet and compare results later. This is why the action is spooky; the physics can influence faster than the speed of light but conspires to do so in such a way that we can never know this.

Quantum Bayesianism [1] offers an alternative explanation that keeps the physics within the universe's speed limit. The spookiness is resolved simply by noting that a wavefunction collapse is purely epistemological. This means that Alice and Bob can have different states-of-knowledge regarding the same experiment about wavefunction collapses. Suppose first that the two qubits are entangled and then Alice and Bob depart never to communicate again. Then Alice's and Bob's state-of-knowledge about the experiment are respectively;

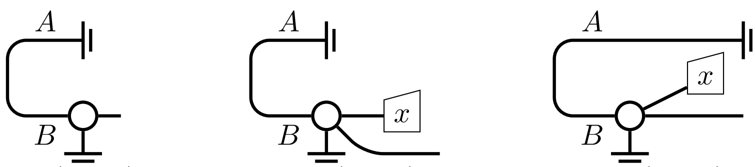



Alice Bob (88)

Now suppose that the friends agree that after separation at time 0, Bob will measure in the computational basis at time 1 and send the result to Alice which is received at time 2. Suppose that $x$ is measured by Bob. Then the states of knowledge at the various times are proportional to the following diagrams.



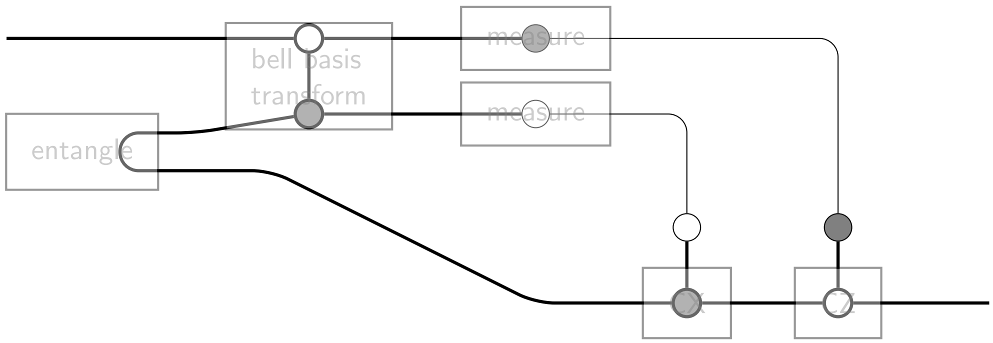

Alice ($t = 0, 1$) Alice ($t = 2$); (89)




Bob ($t = 0$) Bob ($t = 1$) Bob ($t = 2$) (90)

Note that Bob's state-of-knowledge of Alice's qubit changes instantaneously at time 1 at the moment that the result of the measurement is learnt. But the abrupt change is *epistemological* - it happens in Bob's head! - so we don't have the spookiness. Consider the classical analogue: suppose Alice and Bob both receive an envelope, and they know the envelopes contain the same number. Then, when they move far apart, Bob can learn the contents of Alice's envelope by simply looking inside his. It would be absurd to call Bob's learning 'action at a distance'. The only difference in the quantum case is that the correlation between the results of Alice and Bob can't be explained using predetermined variables.

Now let's see how one can use CQM and quantum Bayesianism to understand quantum teleportation. The process is outlined as follows; Alice and Bob again have a pair of entangled qubits. Alice takes a qubit $X$ that she wishes to teleport, and measures the $X \otimes A$ state in the Bell basis. These two bits are then sent to Bob classically. Bob receives these and uses them to perform a couple of unitary transformations on his qubit. All of this is represented by diagram (12). Diagram (12) translates to the following state-of-knowledge diagram



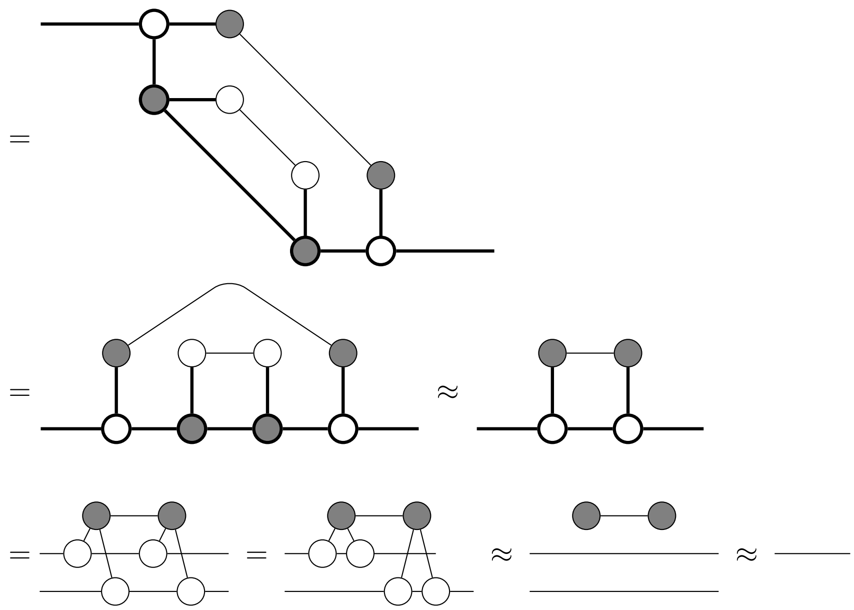

(91)

Here, the white and grey spiders are the Frobenius algebras associated with the $|0\rangle, |1\rangle$ basis and $|+\rangle, |-\rangle$ basis respectively. One can confirm that each component of the diagram is causal and that they are indeed the processes that we claim they are. We will use one additional fact of grey and white spiders which can be easily confirmed.

(92)

Where $\approx$ means that the two diagrams are equal up to a number. This is a result of the two bases being **strongly complementary**. For much more on these multi-spider diagrams, see [10, sections 7,8] and [22, 23]. Using this fact, we can show that our state-of-knowledge diagram simplifies;

(91) =



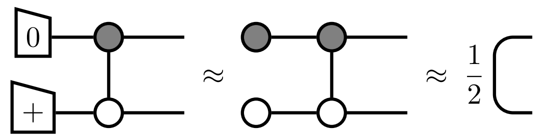
(93)

Here we can really see the power of graphical reasoning. Once one knows the rules of manipulating the diagrams, producing proofs becomes easy and intuitive. So, the process of quantum teleportation is epistemologically the same as simply moving the state $X$ from Alice to Bob. That is, while the real-world setups of quantum teleportation and translating a quantum state are very different, they have exactly the same outcomes.

## 10 Schrödinger's Cat

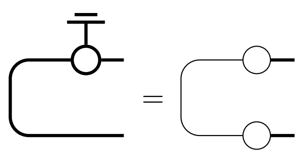
To finish off this essay, let's turn quantum Bayesianism to analyse a fun, cornerstone example. I believe that the work in this section is original, but unfortunately I haven't had time to do a review of the literature on the analysis of Schrödingers Cat. The possibility that the analysis below has not been published elsewhere cannot be ruled out.

The formulation of quantum mechanics we have developed has sad consequences for the famous Schrödinger's Cat experiment. Before we tackle the quantum process representing a full-blown cat, consider a simple 'Schrödinger's Qubit' experiment. Here, our 'cat' qubit is initially in the alive $|1\rangle$ state. We will model our poison as another qubit which is in the superposition of $|0\rangle$ and $|1\rangle$ where the 0 state represents no poison and 1 means that poison is present. The idea of this experiment is for the cat qubit to flip depending on the status of the other qubit. So our experiment is modelled by a controlled NOT gate. But this arrangement is just equal to the cat qubit and the poison qubit being entangled.




(94)

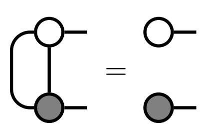
Suppose we have this box in front of us and we know that we are going to open it in 10 minutes and measure the cat qubit in the 1,0 basis. Then our state of knowledge for the experiment looks like this:



(95)

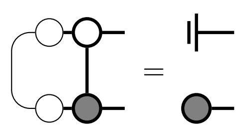
This analysis shows that this state of knowledge is *the same as* the state of knowledge that the two qubits are only classically correlated. There is no 'quantum weirdness' to see here. So Schrödinger's cat is exactly the same as a classical cat.

In what way was the above experiment quantum at all? The answer is that if instead of 'opening the box' we had decided to make the qubits interact with another quantum process, we could get a different answer to its classical partner. For instance, suppose that we piped the two qubits through another CNOT before opening the box. In the quantum case:



(96)

Whereas in the classical case:



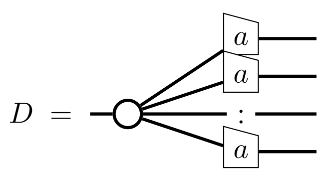
(97)

So we are not yet done in disproving Schrödinger's cat. We must also show that even before we decide to open the box, our state of knowledge is still (at least approximately) equal to its classical counterpart. To show this, we will use the fact that we have great uncertainty about the quantum state of a cat before we put it into the box. A cat is too fluffy to consider mathematically so we shall instead imagine that the cat system is modelled as a huge number of qubits.

Consider a situation where we know that the state of a qubit is copied to many locations, but there is a small chance that it is copied to some qubit that we are not aware of, and is thus discarded. In our example of the cat, we know that each cell in the cat's body will contain the information of the cat's demise. But are you confident enough in your cellular biology to know all of the subsystems of this cell which one could also extract this information from? Even if we hired a team of biologists, how confident can we be that our model is keeping track of every place that the information may be stored?

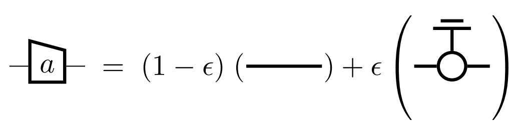
**Theorem 10.1.** *Take $X$ to be a classical object in* **CPM**(**FdHilb**)*. Suppose that we have a system of $n$ qubits and a single qubit is $X$-copied to each qubit. But for each of the $n$ qubits, there is a small chance $\epsilon \ll 1$ that the value of the qubit is copied again and discarded. In this case, decoherence occurs when $\epsilon n \gg 1$.*

*Proof.* The diagram representing the above situation is




(98)

where




(99)

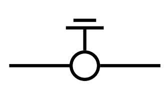
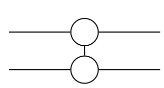
But as a matrix, this is

$$
\begin{aligned}
D &= \sum_{ij=00}^{11} \left( (1-\epsilon) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \epsilon \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right)^{\otimes n} (|ij\rangle)^{\otimes n} \langle ij| \\
&= \sum_{ij=00}^{11} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & (1-\epsilon) & 0 & 0 \\ 0 & 0 & (1-\epsilon) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{\otimes n} (|ij\rangle)^{\otimes n} \langle ij| \\
&= \sum_{ij=00}^{11} \left( \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & (1-\epsilon) & 0 & 0 \\ 0 & 0 & (1-\epsilon) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} |ij\rangle \right)^{\otimes n} \langle ij| \\
&= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}^{\otimes n} \langle 00| + \begin{bmatrix} 0 \\ (1-\epsilon) \\ 0 \\ 0 \end{bmatrix}^{\otimes n} \langle 01| + \begin{bmatrix} 0 \\ 0 \\ (1-\epsilon) \\ 0 \end{bmatrix}^{\otimes n} \langle 10| + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}^{\otimes n} \langle 11| \\
&= \quad |0000\cdots 00\rangle \langle 00| + (1-\epsilon)^n |0101\cdots 01\rangle \langle 01| \\
&\quad + (1-\epsilon)^n |1010\cdots 10\rangle \langle 10| + |1111\cdots 11\rangle \langle 11|
\end{aligned}
\tag{100}
$$

Consider the $\langle 01|$ and $\langle 10|$ terms which are multiplied by $(1-\epsilon)^n$. Taking the log of this expression and remembering that $\epsilon \ll 1$ we get

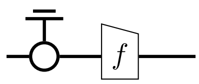
$$\ln (1-\epsilon)^n = n \ln (1-\epsilon) \approx -n\epsilon \tag{101}$$

So if $n\epsilon \gg 1$, these terms vanish. But, this vanishing implies that decoherence occurs. This is because a decoherence process is given by

 $=$ 

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{102}$$

So a process $f$ preceded by decoherence has zero entries in the 2nd and 3rd columns.



$$= \begin{bmatrix} f_{11} & f_{12} & f_{13} & f_{14} \\ f_{21} & f_{22} & f_{23} & f_{24} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \circ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_{11} & 0 & 0 & f_{14} \\ f_{21} & 0 & 0 & f_{24} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{103}$$

Which is exactly what occurs in our process $D$ if $n\epsilon \gg 1$. $\square$

Let's put some numbers into this theorem. There are of order $10^9$ cells in a cat's body. Suppose we wish to model the cat in such a way that it does not collapse into a classical cat. In order to do this, we would have to model each cell in such a way that we assign a probability of much less than $10^{-9}$ to the chance that the information of the cats demise is copied to a subsystem of the cell in a way that we did not anticipate. This is an extreme level of confidence in one's ability to understand the inner workings of a cat. So all realistic attempts to get the cat to enter a superposition of being alive and dead inside an isolated box will not work.

If the analysis above is correct, Schrödinger's cat is *not* in a superposition of being both dead and alive because we are unable to model the cat accurately enough to prevent a wavefunction collapse.

# References


[1] C. A. Fuchs, "Qbism, the perimeter of quantum bayesianism," *arXiv preprint arXiv:1003.5209*, 2010.

[2] C. A. Fuchs, N. D. Mermin, and R. Schack, "An introduction to qbism with an application to the locality of quantum mechanics," *American Journal of Physics*, vol. 82, no. 8, pp. 749–754, 2014.

[3] E. T. Jaynes, *Probability theory: the logic of science.* Cambridge university press, 2003.

[4] R. Jeffrey, *Subjective probability: The real thing.* Cambridge University Press, 2004.

[5] S. Abramsky and B. Coecke, "A categorical semantics of quantum protocols," in *Logic in Computer Science, 2004. Proceedings of the 19th Annual IEEE Symposium on*, pp. 415–425, IEEE, 2004.

[6] S. Abramsky and B. Coecke, "Categorical quantum mechanics," *Handbook of quantum logic and quantum structures: quantum logic*, pp. 261–324, 2009.

[7] S. Mac Lane, *Categories for the working mathematician*, vol. 5. Springer Science & Business Media, 1978.

[8] A. Joyal and R. Street, "The geometry of tensor calculus, i," *Advances in Mathematics*, vol. 88, no. 1, pp. 55–112, 1991.

[9] P. Selinger, "A survey of graphical languages for monoidal categories," *ArXiv e-prints*, Aug. 2009.

[10] B. Coecke and A. Kissinger, "Picturing quantum processes. a first course in quantum theory and diagrammatic reasoning," 2016. to be published later in 2016.

[11] G. Chiribella, G. M. DAriano, and P. Perinotti, "Probabilistic theories with purification," *Physical Review A*, vol. 81, no. 6, p. 062348, 2010.

[12] G. Chiribella, G. M. DAriano, and P. Perinotti, "Informational derivation of quantum theory," *Physical Review A*, vol. 84, no. 1, p. 012311, 2011.

[13] P. Selinger, "Finite dimensional hilbert spaces are complete for dagger compact closed categories," *Electronic Notes in Theoretical Computer Science*, vol. 270, no. 1, pp. 113–119, 2011.

[14] R. Penrose, "Applications of negative dimensional tensors," *Combinatorial mathematics and its applications*, vol. 221244, 1971.

[15] J. H. Larkin and H. A. Simon, "Why a diagram is (sometimes) worth ten thousand words," *Cognitive science*, vol. 11, no. 1, pp. 65–100, 1987.

[16] S. Card, J. Mackinlay, and B. Shneiderman, *Readings in Information Visualization: Using Vision to Think*. Interactive Technologies Series, Morgan Kaufmann Publishers, 1999.

[17] P. Selinger, "Dagger compact closed categories and completely positive maps," *Electronic Notes in Theoretical Computer Science*, vol. 170, pp. 139–163, 2007.

[18] R. Piedeleu, D. Kartsaklis, B. Coecke, and M. Sadrzadeh, "Open system categorical quantum semantics in natural language processing," *CoRR*, vol. abs/1502.00831, 2015.

[19] B. Coecke, "Axiomatic description of mixed states from selinger's cpm-construction," *Electronic Notes in Theoretical Computer Science*, vol. 210, pp. 3–13, 2008.

[20] B. Coecke and D. Pavlovic, "Quantum measurements without sums," *arXiv preprint quant-ph/0608035*, 2006.

[21] S. Lack, "Composing props," *Theory and Applications of Categories*, vol. 13, no. 9, pp. 147–163, 2004.

[22] B. Coecke and R. Duncan, "Interacting quantum observables: categorical algebra and diagrammatics," *New Journal of Physics*, vol. 13, no. 4, p. 043016, 2011.

[23] A. Kissinger, "Pictures of processes: automated graph rewriting for monoidal categories and applications to quantum computing," *arXiv preprint arXiv:1203.0202*, 2012.