Inverses, disintegrations, and Bayesian inversion in quantum Markov categories

Arthur PARZYGNAT

Institut des Hautes Études Scientifiques
35, route de Chartres
91440 – Bures-sur-Yvette (France)

Janvier 2020

IHES/M/20/01

Inverses, disintegrations, and Bayesian inversion in quantum Markov categories

Arthur J. Parzygnat

January 20, 2020

Abstract

We analyze three successively more general notions of reversibility and statistical inference: ordinary inverses, disintegrations, and Bayesian inferences. We provide purely categorical definitions of these notions and show how each one is a strictly special instance of the latter in the cases of classical and quantum probability. This provides a categorical foundation for Bayesian inference as a generalization of reversing a process. To properly formulate these ideas, we develop quantum Markov categories by extending recent work of Cho–Jacobs and Fritz on classical Markov categories. We unify Cho–Jacobs' categorical notion of almost everywhere (a.e.) equivalence in a way that is compatible with Parzygnat–Russo's C^*-algebraic a.e. equivalence in quantum probability.

Contents

2010 Mathematics Subject Classification. 62F15, 46L53 (Primary); 81R15, 18D10, 81P45 (Secondary)

Key words and phrases. Quantum probability; conditional probability; monad; Kleisli category; quantum information; categorical quantum mechanics; reversibility; Bayes; Kadison–Schwarz; statistical inference

1 Introduction and outline

In Gromov's lectures on entropy, he emphasized that concepts in mathematics should frequently be revisited due to our constantly growing and changing perspectives, which may provide new insight on old subjects [12]. Probability theory is no exception, and a dramatic change in viewpoint on the structural foundations of probability theory has gained enormous momentum recently [6], [14], [10]. However, most of the guiding examples towards this perspective have come from classical probability theory. Here, we would like to continue our investigation of quantum disintegrations by extending our work [27] to define and incorporate quantum Bayesian inference in abstract probability theory. We will define and analyze the properties of, and relationships between, inverses, disintegrations (also known as regular conditional probabilities or optimal hypotheses), and Bayesian inferences in the general context of reversing dynamics in quantum Markov categories, which are also introduced in this paper. This context is broad enough to include classical and quantum probability.[1]

More specifically, we show that invertible maps always have disintegrations and we classify which deterministic maps are invertible in terms of disintegrations. We then prove disintegrations are only possible for deterministic maps and disintegrations are automatically Bayesian inferences. This shows that Bayesian inference is the most general of these three notions of reversibility in the classical and quantum setting, i.e.

$$\text{Invertible} \quad \Rightarrow \quad \text{Disintegrable} \quad \Rightarrow \quad \text{Bayesian invertible.}$$

In the process of introducing disintegrations, one enlarges their original category to include probabilistic morphisms that optimally reverse certain deterministic dynamics.[2] Hence, one now has new morphisms describing *stochastic* dynamics. In this work, we show that re-using the notion of a disintegration is not sufficient to reverse these processes optimally. More precisely, if a stochastic morphism has a disintegration, then the original stochastic morphism is necessarily essentially deterministic. Bayesian inference, the third notion of reversibility that we will examine, correctly captures an appropriate reversal procedure that reduces to the disintegration case when the original dynamics is deterministic. Although some of these results hold generally, we prove these claims in our two main categories of interest: the first is the category of finite sets and functions/stochastic maps (conditional probabilities), while the second is the category of finite-dimensional unital C^*-algebras and unital *-homomorphisms/completely positive unital maps (quantum operations).

[1] Most of our results are stated in the finite-dimensional setting purely for simplicity. Nevertheless, many of the results also hold for von Neumann algebras, though we have not explicitly checked if any continuity conditions (such as normality) are required.

[2] Although this is reminiscent of what one does in the localization of a category with respect to a class of morphisms, we have not made any explicit connection. It would be interesting to see the relationship, if one exists. In our setup, one begins with a category of deterministic processes and uses a monad to construct a Kleisli category, whose new morphisms are thought of as describing *stochastic* dynamics. For classical (quantum) systems, this categorical procedure takes us from evolution described by functions (*-homomorphisms) on the algebras of observables to evolution described by Markov kernels [17], [11] (completely positive unital maps [33]).

The notion of a.e. equivalence[3] in classical probability theory plays an important role in uniqueness properties. In [27], Russo and the author introduced the notion of a.e. equivalence for maps between C^*-algebras equipped with states (or more generally positive functionals) to determine the uniqueness of disintegrations. The definition is simple, intuitive, and is motivated by the Gelfand–Naimark–Segal (GNS) construction. In [3], Cho and Jacobs introduced a categorical formulation of a.e. equivalence valid for any (commutative) Markov category. In this paper, we will show that these two notions agree for $*$-preserving morphisms[4] in the quantum Markov category of von Neumann algebras. This notion of a.e. equivalence also plays an essential role in determining the uniqueness of quantum Bayesian inverses. Many of the important properties of disintegrations, Bayesian inverses, and their relationships to each other discussed here will be used in forthcoming work on a quantum Bayes' theorem [28], [26]. Although, the topic of reversibility in quantum mechanics has been studied in great depth in the literature (a small selection of references include [29], [2], [23], [19], and [20]), the categorical approach we take here seems novel. The quantum Markov categories we define enable us to reason probabilistically via diagrammatic techniques as a form of two-dimensional algebra, similar to the growing subject of categorical quantum mechanics [5], [13].

The outline of this paper is as follows. In Section 3, we define quantum Markov categories and provide the two main examples used in this work: finite sets with stochastic maps and finite-dimensional C^*-algebras with completely positive unital (CPU) maps. Technically, the latter is modified to include all the morphisms needed to make it a quantum Markov category. In Section 4, we adapt Fritz' definition of a positive Markov category (cf. [10, Definition 11.22]) to the quantum setting. In Theorem 4.5, we prove that the category of CPU maps forms a positive subcategory of the quantum Markov category of linear and conjugate linear maps on finite-dimensional C^*-algebras. We then prove the surprising result that ordinary positivity (as opposed to complete positivity) in the quantum setting is not enough to satisfy Fritz' categorical definition of positivity. Section 5 reviews a.e. equivalence and proves several new results: Theorem 5.16 shows that the notion of a.e. equivalence via GNS introduced in [27, Definition 3.16] coincides with one of the two definitions of Cho–Jacobs a.e. equivalence [3, Definition 5.1]. Section 6 defines disintegrations and Bayesian inference in quantum Markov categories. Proposition 6.16 shows that every $*$-preserving morphism is a Bayesian inverse of its Bayesian inverse and Theorem 6.27 shows that a Bayesian inverse of a deterministic morphism is a disintegration. Section 7 contains statements that were proven explicitly for finite sets and stochastic maps for which we did not find diagrammatic proofs. Section 8 does the same but for CPU maps on finite-dimensional C^*-algebras (and sometimes von Neumann algebras). An interesting result here is Theorem 8.3, which shows that if a CPU map between two von Neumann algebras has a disintegration, then the map is a.e. deterministic. In addition, Theorem 8.27 proves that all disintegrations are Bayesian inverses.

[3]The a.e. here stands for almost everywhere and comes from measure theory. Probability theorists might instead use a.s., which stands for almost surely.

[4]In a quantum Markov category, there is $\mathbb{Z}_2$-grading and an involution morphism $*$ for every object. The notion of a $*$-preserving morphism isolates an important symmetry satisfied in classical systems that need not hold in quantum systems.

2 What is Bayes' theorem?

To provide a setting for our results, we would first like to illustrate that Bayes' theorem can be described purely diagrammatically [9], [6], [4], [3], [10].[5] We will presently illustrate it in the case of finite sets and stochastic maps (for the reader unfamiliar with the notation, we will briefly review it after the statement of the theorem).

Theorem 2.1 (Bayes' theorem). *Let* X *and* Y *be finite sets, let* $\{\bullet\} \overset{p}{\rightsquigarrow} X$ *be a probability measure, and let* $X \overset{f}{\rightsquigarrow} Y$ *be a stochastic map. Then there exists a stochastic map* $Y \overset{g}{\rightsquigarrow} X$ *such that*[6]

$$\begin{tikzcd} Y \arrow[d, "\Delta_Y"'] & \{\bullet\} \arrow[l, squiggly, "q"'] \arrow[r, squiggly, "p"] \arrow[d, phantom, "="] & X \arrow[d, "\Delta_X"] \\ Y \times Y \arrow[r, squiggly, "g \times \mathrm{id}_Y"'] & X \times Y & X \times X \arrow[l, squiggly, "\mathrm{id}_X \times f"] \end{tikzcd}, \tag{2.2}$$

where $\{\bullet\} \overset{q}{\rightsquigarrow} Y$ *is given by* $q := f \circ p$. *Furthermore, for any other* g' *satisfying this condition,* $g \underset{q}{=} g'$.

We quickly recall some notation to explain the theorem (see [10], [27], and [24] for a more leisurely introduction).[7] If X and Y are finite sets, a ***stochastic map*** $X \overset{f}{\rightsquigarrow} Y$ is an assignment sending $x \in X$ to a probability measure f_x on Y. The value of this probability measure on $y \in Y$ will be denoted by f_{yx}. Stochastic maps are drawn with squiggly arrows to distinguish them from deterministic maps (stochastic maps assigning Dirac delta measures), which are drawn with straight arrows $\rightarrow$. Such straight arrows correspond to functions. A single element set will be denoted by $\{\bullet\}$. A stochastic map $\{\bullet\} \overset{p}{\rightsquigarrow} X$ is precisely a probability measure on X. Stochastic maps $X \overset{f}{\rightsquigarrow} Y \overset{g}{\rightsquigarrow} Z$ can be composed via the Chapman–Kolmogorov equation

$$(g \circ f)_{zx} := \sum_{y \in Y} g_{zy} f_{yx}. \tag{2.3}$$

Given $X \overset{f}{\rightsquigarrow} Y$ and $X' \overset{f'}{\rightsquigarrow} Y'$, the ***product*** $X \times X' \overset{f \times f'}{\rightsquigarrow} Y \times Y'$ is defined by the product of probability measures

$$(f \times f')_{(y,y')(x,x')} := f_{yx} f_{y'x'}. \tag{2.4}$$

Given

$$\begin{tikzcd} \{\bullet\} \arrow[r, squiggly, "p"] & X \arrow[r, squiggly, shift left, "f"] \arrow[r, squiggly, shift right, "h"'] & Y \end{tikzcd}, \tag{2.5}$$

[5] In these references, Bayes' theorem is formulated as a bijection between joint distributions and conditionals. Our emphasis is on the process of inference from conditionals, which will be used more in the non-commutative setting. Why this is so will be explained in [28]. To the best of our knowledge, the first reference that explicitly draws the diagram (2.2) is Fong's thesis [9] (see the section "Further Directions"), though it is formulated using string diagrams. Here, we have elevated this diagram to encapsulate what the *statement* of Bayes' theorem is.

[6] The equals sign in this diagram indicates that the diagram commutes. The notation is meant to be consistent with higher categorical notation. Namely, we think of this equality as the identity 2-cell. We will not comment on higher categorical generalizations in this paper.

[7] The reader may also enjoy the short introductory video lectures available at `https://www.youtube.com/playlist?list=PLSx1kJDjrLRQksb7H9fqRE8GVMJdkX-4A`.

f is p-***a.e. equivalent to*** h, written $f \underset{p}{=} h$, whenever

$$p\{x \in X \ : \ f_{yx} \neq h_{yx} \text{ for some } y \in Y\} = 0, \tag{2.6}$$

i.e. the set on which f and h differ is a set of p-measure zero. Finally, the map $X \xrightarrow{\Delta_X} X \times X$ is determined by the function $\Delta_X(x) := (x, x)$ for all $x \in X$.

With all this notation explained, the reader can now verify that the diagram (2.2) in Bayes' theorem reads

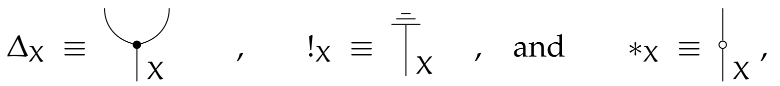
$$g_{xy} q_y = f_{yx} p_x \tag{2.7}$$

for all values of $x \in X$ and $y \in Y$. This is Bayes' rule for point events.[8] The case of Bayes' rule for more general events is a simple consequence of this rule. The morphism g is called the ***Bayesian inference*** associated to (f, p, q).

3 Quantum Markov categories

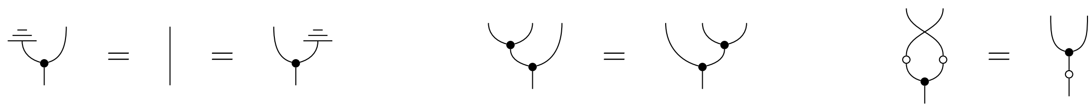
We begin by defining our main categories of study. In what follows, let $\mathbb{Z}_2 = \{0, 1\}$ be the abelian group, where 0 is the identity and $1 + 1 := 0$. Let $\mathbb{B}\mathbb{Z}_2$ be the one object category whose set of morphisms equals $\mathbb{Z}_2$ with composition as addition modulo 2.

Definition 3.1. A category $\mathcal{C}$ equipped with a functor $\mathcal{C} \to \mathbb{B}\mathbb{Z}_2$ is called a $\mathbb{Z}_2$-***graded category***. Morphisms in $\mathcal{C}$ that get sent to 0 are called ***even*** and morphisms that get sent to 1 are called ***odd***. A collection of morphisms that is either all even or all odd is said to be ***homogeneous***. A ***quantum Markov category*** is a $\mathbb{Z}_2$-graded symmetric monoidal category $(\mathcal{C}, \otimes, I)$ together with a family of morphisms $\Delta_X : X \rightsquigarrow X \times X$, $!_X : X \rightsquigarrow I$, and $*_X : X \rightsquigarrow X$, all depicted in string diagram notation as



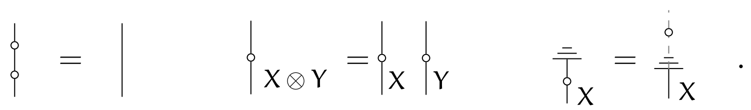
(3.2)

for all objects X in $\mathcal{C}$. These morphisms are required to satisfy the following conditions[9]



(3.3)

(3.4)

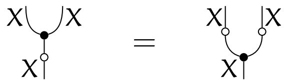


(3.5)

[8]If we set $P(x|y) := g_{xy}$, $P(y) := q_y$, $P(y|x) := f_{yx}$, and $P(x) := p_x$, this equation reads $P(x|y)P(y) = P(y|x)P(x)$ in more standard (albeit abusive) notation.

[9]These conditions will be examined in the case of C^*-algebras in Example 3.12 to assist the reader in deciphering such string diagrams.

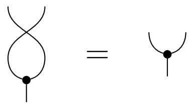
In addition, id_X, Δ_X, and $!_X$ are declared to be even for all X. The collection of involutions $*_X$ is declared to be homogeneous. The map Δ_X is sometimes called ***copy*** or ***duplicate*** and the map $!_X$ is sometimes called ***delete*** or ***ground***. If, in addition,



$\forall\, X,$ (3.6)

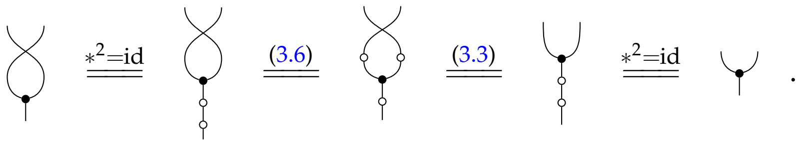
then $\mathcal{C}$ is said to be a ***classical Markov category***.[10]

Remark 3.7. Note that we have dropped the condition that grounding is natural for every morphism (cf. [10, Definition 2.1]) working more closely with Cho–Jacobs' version (cf. [3, Definition 2.2]). Also notice that we have not required $*_I$ to be the identity (the latter of which is always even). This is potentially awkward for string diagram computations because we now have to keep track of this whenever we ever pull $*_X$ through $!_X$ as in the last identity in (3.5). Fortunately, this will never show up in any of the string-diagrammatic computations that will follow. We have also replaced the axiom



(3.8)

from Fritz' definition of a Markov category with a version that interacts with the $*$ operation. The reason is to include categories of quantum probability (cf. Example 3.12 below). Note that the axiom (3.8) is a consequence of the axioms of a quantum Markov category and (3.6). This follows from



(3.9)

The terminology 'Markov category' was first used by Fritz [10]. The terminology 'CD category' was used earlier by Cho–Jacobs, which is also where the axioms were first provided [3]. We prefer the terminology 'Markov category' because this sounds more appropriate for our generalization to the non-commutative context.[11]

Remark 3.10. The choice of a functor $\mathcal{C} \to \mathbb{B}\mathbb{Z}_2$ means that the composite of two morphisms of parities p_1 and p_2 is of parity $(p_1 + p_2) \mod 2$. If $*$ is odd, then pre- or post-composing with $*$ sets up two bijections $\mathcal{C}^{\text{even}}(X, Y) \to \mathcal{C}^{\text{odd}}(X, Y)$.

Example 3.11. One of our main examples of a quantum Markov category is **FinStoch**. An object of **FinStoch** is a finite set. A morphism from X to Y is a Markov kernel/stochastic map/conditional probability from X to Y. Such a morphism assigns to each element $x \in X$ a probability measure on Y. Composition is defined by the Chapman–Kolmogorov equation

[10] This is automatic if $*_X = \mathrm{id}_X$ for all X. Otherwise, $*_X$ is declared to be odd.

[11] It is known in quantum mechanics that the operations of copy (C) and discard (D) are not quantum operations. Hence, if we called our non-commutative analogues 'non-commutative CD categories' or 'quantum CD categories,' this might cause some alarm in the quantum information community.

(i.e. summing over all intermediaries). The tensor product is the cartesian product of sets and the product of Markov kernels for morphisms. The tensor unit is the single element set, often denoted by $\{\bullet\}$. The maps Δ_X and $!_X$ are given by $\Delta_X(x) := (x,x)$ and $!_X(x) = \bullet$ for all $x \in X$. **FinStoch** is also a classical Markov category upon setting $*_X = \mathrm{id}_X$ because axiom (3.8) holds. See Section 2 above, [10, Example 2.5], and [27, Section 2.1] for more details. One can also drop the condition that a morphism sends each point to a probability measure and instead associate to each point a signed measure. The resulting category is also a quantum Markov category (see Example 11.27 in [10] but drop the condition that the total measure must be 1).

Example 3.12. Our second (and motivating) main example of a quantum Markov category (that is not a classical Markov category) is **fdC*-AlgU**$^{\mathrm{op}}$. The objects here are finite-dimensional unital C^*-algebras (henceforth, all C^*-algebras will be assumed unital). Every such finite-dimensional C^*-algebra is *-isomorphic to a finite direct sum of (square) matrix algebras [8, Theorem 5.5]. A matrix algebra will be written as $\mathcal{M}_n(\mathbb{C})$ indicating the C^*-algebra of complex $n \times n$ matrices. On occasion, the shorthand $\mathcal{M}_n$ may be used in place of $\mathcal{M}_n(\mathbb{C})$. A morphism from $\mathcal{A}$ to $\mathcal{B}$ in **fdC*-AlgU**$^{\mathrm{op}}$ is *either* a linear *or conjugate-linear* unital map $\mathcal{B} \rightsquigarrow \mathcal{A}$. Notice that the function goes backwards because of the superscript $^{\mathrm{op}}$ (in the physics literature, this convention is known as the *Heisenberg picture*). The tensor product is the tensor product of finite-dimensional C^*-algebras. For example,

$$\left(\bigoplus_{x\in X}\mathcal{M}_{m_x}(\mathbb{C})\right)\otimes\left(\bigoplus_{y\in Y}\mathcal{M}_{n_y}(\mathbb{C})\right)=\bigoplus_{x,y}\mathcal{M}_{m_x}(\mathbb{C})\otimes\mathcal{M}_{n_y}(\mathbb{C}), \tag{3.13}$$

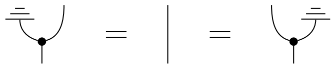
where X and Y are finite sets labelling the matrix factors. A similar situation holds for the tensor product of morphisms. The $*$ operation is the involution on C^*-algebras. Notice that if $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ is linear (conjugate linear), then $F \circ *$ is conjugate linear (linear) since $(F\circ *)(\lambda b) = F(\overline{\lambda} b^*) = \overline{\lambda} F(b^*) = \overline{\lambda}(F \circ *)(b)$ (and similarly if F is conjugate linear). We will ignore associators and unitors in what follows. This is permissible thanks to Mac Lane's coherence theorem [22]. We define the copy map $\Delta_{\mathcal{A}}$ from $\mathcal{A}$ to $\mathcal{A}\otimes\mathcal{A}$ in **fdC*-AlgU**$^{\mathrm{op}}$ to be the multiplication map determined on elementary tensors by

$$\begin{aligned} \mathcal{A}\otimes\mathcal{A} &\overset{\mu_{\mathcal{A}}}{\rightsquigarrow} \mathcal{A} \\ A\otimes B &\longmapsto AB. \end{aligned} \tag{3.14}$$

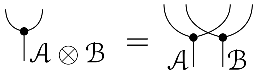
in **fdC*-AlgU**. The map $\mu_{\mathcal{A}}$ is linear and unital, but it is not a *-homomorphism unless $\mathcal{A}$ is commutative. In fact, μ_A is not even positive in general (cf. Example 3.19). Nevertheless, it is coherent with the involution $*$ (in the sense of the last identity in (3.3)) because $(ab)^* = b^*a^*$ for all $a, b \in \mathcal{A}$. Finally, the discard map $!_{\mathcal{A}} : \mathcal{A} \to \mathbb{C}$ in **fdC*-AlgU**$^{\mathrm{op}}$ is defined to be the unit inclusion map

$$\begin{aligned} \mathbb{C} &\to \mathcal{A} \\ \lambda &\mapsto \lambda 1_{\mathcal{A}} \end{aligned} \tag{3.15}$$

in **fdC*-AlgU**. Here are some of the conditions of a quantum Markov category and their corresponding expressions in terms of these morphisms:

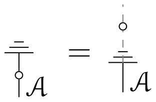


$$\iff \qquad 1_{\mathcal{A}}A = A = A1_{\mathcal{A}} \qquad \forall\, A \in \mathcal{A}, \tag{3.16}$$



$$\iff (A \otimes B)(A' \otimes B') = (AA') \otimes (BB') \quad \forall\, A, A' \in \mathcal{A},\ B, B' \in \mathcal{B}, \tag{3.17}$$

and



$$\iff \qquad (\lambda 1_{\mathcal{A}})^* = \bar{\lambda} 1_{\mathcal{A}} \qquad \forall\, \lambda \in \mathbb{C}. \tag{3.18}$$

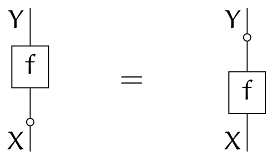
One can check that the rest of the axioms of a quantum Markov category are satisfied for **fdC*-AlgU**$^{\text{op}}$. In fact, the larger category where we drop the unit-preserving assumption on the morphisms is also a quantum Markov category. In this paper, we will denote this latter category by **fdC*-Alg**$^{\text{op}}$. We will be lax with our notation and from now on not distinguish between the category **fdC*-AlgU** and its opposite. When we refer to **fdC*-AlgU** as a quantum Markov category, we will always mean its opposite. In all the string diagrams that appear, the only difference is that we will compose from the top to the bottom of the page (rather than from the bottom to the top).

The following provides an example of an important subcategory of **fdC*-AlgU** that is *not* a quantum Markov category. Nevertheless, it is the main category of interest here and the fact that it embeds into a quantum Markov category is crucial for the theorems that will follow for Bayesian inference and disintegrations.

Example 3.19. Let **fdC*-AlgCPU** be the subcategory of **fdC*-AlgU** consisting of the same objects as **fdC*-AlgU** but whose morphisms are (linear) completely positive unital (CPU) maps. This is not a quantum Markov category because there is no CPU map $\mathcal{A} \otimes \mathcal{A} \rightsquigarrow \mathcal{A}$ satisfying the conditions of Definition 3.1. In fact, the no-cloning (no-broadcasting) theorem states that a CPU map $\mu_{\mathcal{A}} : \mathcal{A} \otimes \mathcal{A} \rightsquigarrow \mathcal{A}$ satisfying the first condition in (3.3), i.e. $\mu_{\mathcal{A}}(1_{\mathcal{A}} \otimes A) = A = \mu_{\mathcal{A}}(A \otimes 1_{\mathcal{A}})$ for all $A \in \mathcal{A}$, exists if and only if $\mathcal{A}$ is commutative (cf. [21, Theorem 6]). The reader should be able to reconstruct a simple proof of this claim from the Multiplication Theorem, which will be stated below in Lemma 4.8.

We now introduce a few properties that we wish to distinguish for certain morphisms in quantum Markov categories. The first is the notion of a $*$-preserving morphism.

Definition 3.20. Let $\mathcal{C}$ be a quantum Markov category. A morphism $X \overset{f}{\rightsquigarrow} Y$ in $\mathcal{C}$ is said to be ***$*$-preserving*** iff f is natural with respect to $*$,[12] meaning

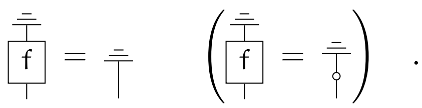


$$, \qquad \text{i.e.} \quad f \circ *_X = *_Y \circ f. \tag{3.21}$$

[12]This word 'natural' is meant in the categorical sense. The assignment $*$ assigns to each object X a morphism $*_X$. This assignment is natural (in the sense of natural transformations) precisely for morphisms that are $*$-preserving.

Remark 3.22. In a quantum Markov category $\mathcal{C}$, copy Δ is $*$-preserving if and only if $*$ is the identity. The collection of all objects and $*$-preserving morphisms of $\mathcal{C}$ form a subcategory of $\mathcal{C}$.

Definition 3.23. Let $\mathcal{C}$ be a quantum Markov category. An even (odd) morphism $X \overset{f}{\rightsquigarrow} Y$ in $\mathcal{C}$ is called ***causal***[13] iff the composite $X \overset{f}{\rightsquigarrow} Y \overset{!_Y}{\rightsquigarrow} I$ is equal to $X \overset{!_X}{\rightsquigarrow} I$ ($X \overset{*_X}{\rightsquigarrow} X \overset{!_X}{\rightsquigarrow} I$). In pictures,

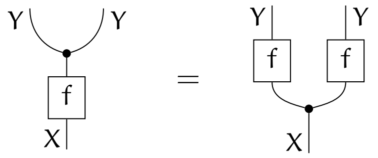


(3.24)

Example 3.25. A morphism in any of the categories of finite sets together with morphisms that associate to each point a signed measure is causal iff the total measure associated to each point is 1. A morphism in any of the categories of finite-dimensional C^*-algebras we have introduced is causal if and only if it is unital.

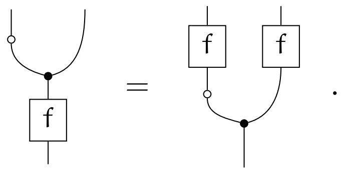
Remark 3.26. It follows from the axioms in Definition 3.1 that id_X, Δ_X, $!_X$, and $*_X$ are automatically causal for all X.

Definition 3.27. A morphism $f : X \rightsquigarrow Y$ in a quantum Markov category is called ***deterministic*** iff f is $*$-preserving, causal, and natural with respect to Δ, meaning



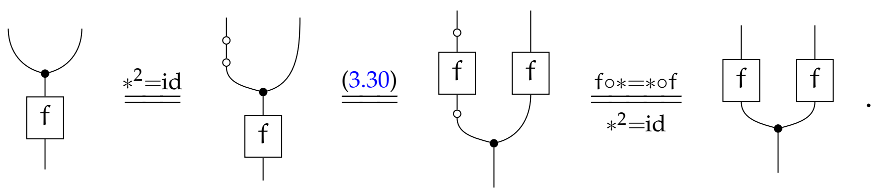
, i.e. $\Delta_Y \circ f = (f \otimes f) \circ \Delta_X$. (3.28)

Remark 3.29. In a quantum Markov category, if $f : X \rightsquigarrow Y$ is causal, then f is both $*$-preserving and deterministic if and only if



(3.30)

If f is deterministic, (3.30) easily follows. For the converse, grounding the upper right string in (3.30) shows f is $*$-preserving. Determinism of f then follows from $*^2 = \mathrm{id}$ and the fact that f is $*$-preserving and satisfies this condition:



(3.31)

One can also show that the tensor product of two deterministic maps is deterministic. This follows from naturality of the braiding, the definition of determinism, the third identity in (3.4), and the second identity in (3.5).

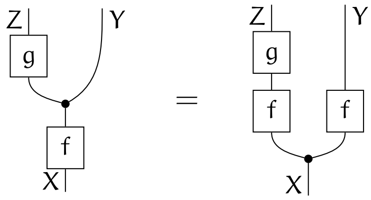
[13] This definition of causal is consistent with Cho–Jacobs [3] and the school on categorical quantum mechanics [31]. It is merely just naturality (in the sense of natural transformations) with respect to the assignment that sends each X to the morphism $!_X$ [10, Equation (2.5) in Definition 2.1].

Example 3.32. In **FinStoch**, deterministic maps correspond to functions, assignments where the measures associated to points are Dirac measures [24, Theorems 2.82 and 2.85]. In $\mathbf{fdC^*\text{-}AlgU}^{\mathrm{op}}$, deterministic maps correspond to *-homomorphisms. Indeed, if $f : \mathcal{B} \rightsquigarrow \mathcal{A}$ is a linear unital map of C^*-algebras, then condition (3.30) says $f(B^*B') = f(B)^*f(B')$ for all $B, B' \in \mathcal{B}$.

4 2-Positive subcategories

The following definition of positivity is due to Fritz [10, Definition 11.22]. However, based on Example 4.13, we have decided to use the terminology '2-positivity' instead.

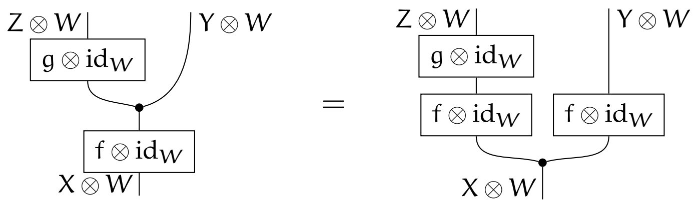
Definition 4.1. Let $\mathcal{M}$ be a quantum Markov category. A subcategory $\mathcal{C} \subseteq \mathcal{M}$ is said to be ***2-positive*** in $\mathcal{M}$ iff for every pair of composable morphisms $X \overset{f}{\rightsquigarrow} Y \overset{g}{\rightsquigarrow} Z$ in $\mathcal{C}$ such that $g \circ f$ is deterministic, the equality



(4.2)

must also hold.

Remark 4.3. Let $X \overset{f}{\rightsquigarrow} Y \overset{g}{\rightsquigarrow} Z$ be a composable pair of morphisms in a 2-positive subcategory $\mathcal{C}$ of a quantum Markov category $\mathcal{M}$ such that $g \circ f$ is deterministic. Let W be any object of $\mathcal{M}$. Then $(g \otimes \mathrm{id}_W) \circ (f \otimes \mathrm{id}_W) = (g \circ f) \otimes \mathrm{id}_W$ is deterministic by Remark 3.29 and



. (4.4)

The fact that **FinStoch** is a 2-positive category was proved in [10, Example 11.25] (in fact, this was proved for the larger category of Markov kernels between measurable spaces). Here, we prove a non-commutative version of this result.

Theorem 4.5. *The category* $\mathbf{fdC^*\text{-}AlgCPU}$ *is a 2-positive subcategory of* $\mathbf{fdC^*\text{-}AlgU}$.

To prove this theorem, we recall two important results regarding multiplicative properties of CPU maps.

Lemma 4.6 (The Kadison–Schwarz inequality)**.** *Let $\mathcal{B} \overset{\varphi}{\rightsquigarrow} \mathcal{A}$ be a CPU map between C^*-algebras. Then*

$$\varphi(A)^*\varphi(A) \leqslant \varphi(A^*A) \qquad \forall\, A \in \mathcal{A}. \tag{4.7}$$

Proof. See [21, Proposition 6] (or [16] and [7] for the original references). ∎

Lemma 4.8 (The Multiplication Theorem)**.** *Let* $\mathcal{B} \overset{\varphi}{\rightsquigarrow} \mathcal{A}$ *be a CPU map between* C^**-algebras. Suppose that* $\varphi(B^*B) = \varphi(B)^*\varphi(B)$ *for some* $B \in \mathcal{B}$*. Then*

$$\varphi(B^*C) = \varphi(B)^*\varphi(C) \quad \textit{and} \quad \varphi(C^*B) = \varphi(C)^*\varphi(B) \qquad \forall\, C \in \mathcal{B}. \tag{4.9}$$

Proof of Lemma 4.8. See [21, Theorem 4] for a proof or the more general result that we will prove later (Lemma 8.28). ■

Proof of Theorem 4.5. Let $\mathcal{C} \overset{G}{\rightsquigarrow} \mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ be a pair of composable CPU maps of C^*-algebras such that the composite $F \circ G$ is a *-homomorphism. Then,

$$\begin{aligned} F\big(G(C)^*G(C)\big) &\leqslant F\big(G(C^*C)\big) \quad \text{by Kadison–Schwarz for } G \\ &= F\big(G(C)\big)^*F\big(G(C)\big) \quad \text{since } F \circ G \text{ is deterministic} \\ &\leqslant F\big(G(C)^*G(C)\big) \quad \text{by Kadison–Schwarz for } F \end{aligned} \tag{4.10}$$

holds for all $C \in \mathcal{C}$. Thus, all inequalities become equalities. In particular,

$$F\big(G(C)^*G(C)\big) = F\big(G(C)\big)^*F\big(G(C)\big) \qquad \forall\, C \in \mathcal{C}. \tag{4.11}$$

By the Multiplicative Theorem (Lemma 4.8), this implies

$$F\big(G(C)^*B\big) = F\big(G(C)\big)^*F(B) \qquad \forall\, C \in \mathcal{C},\ B \in \mathcal{B}. \tag{4.12}$$

Since F and G are $*$-preserving and $*$ is an involution, this reproduces condition (4.2). ■

Example 4.13. The subcategory of all 2-positive unital maps between finite-dimensional C^*-algebras is also a 2-positive subcategory of **fdC*-AlgU**. This is because all of the lemmas used to prove Theorem 4.5 also hold for 2-positive unital maps. Somewhat surprisingly, however, the subcategory of finite-dimensional C^*-algebras together with positive unital maps is *not* a 2-positive subcategory of **fdC*-AlgU**. To see this, take $\mathcal{A} = \mathcal{B} = \mathcal{M}_n(\mathbb{C})$ (with $n \geqslant 2$) and set $f := T =: g$, where T is the map that takes the transpose of matrices. This map is known to be positive and unital, but it is not 2-positive. Furthermore, $g \circ f = T^2 = \mathrm{id}$, which is deterministic. Nevertheless, we have

$$(A^TB)^T \neq AB^T \qquad \forall\, A, B \in \mathcal{A}. \tag{4.14}$$

This prompts the following question: is the subcategory of 2-positive unital maps the largest 2-positive subcategory of **fdC*-AlgU** that contains all CPU maps? This would support our choice of using the terminology '2-positive.' We will not answer this question here, but if this is the case, then Fritz' definition of a positive subcategory seems to capture not quite positivity, but some categorical notion of 2-positivity.

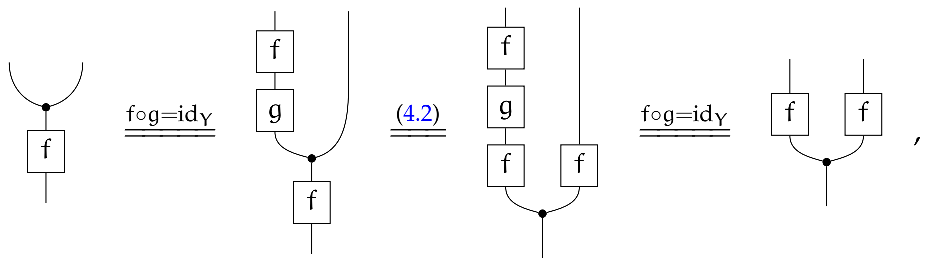
Example 4.15. Based on the fact that the category of finite sets together with morphisms assigning signed measures to points embeds fully and faithfully into the (opposite of the) category of finite-dimensional C^*-algebras together with linear maps, the latter is *not* a 2-positive subcategory of itself. This follows immediately from the fact that $\mathbf{FinStoch}_{\pm}$, as defined in [10, Example 11.27], is not positive.

Remark 4.16. Interestingly, if we consider the category **fdC*-AlgCPU** together with all morphisms generated by the (conjugate-linear) involutions, the 2-positivity condition is also lost. Indeed, setting $g = f = *$ for some fixed (strictly non-commutativite) C*-algebra, then $f \circ g$ is deterministic (it equals the identity), but the identity (4.2) does not hold.

2-positive subcategories of quantum Markov categories have several useful properties [10, Remark 11.28].

Lemma 4.17. *Let $\mathcal{C}$ be a 2-positive subcategory of a quantum Markov category $\mathcal{M}$. Then every morphism in $\mathcal{C}$ that has an inverse in $\mathcal{C}$ is deterministic.*

Proof. Let $X \overset{f}{\rightsquigarrow} Y$ be a morphism in $\mathcal{C}$ with inverse $Y \overset{g}{\rightsquigarrow} X$ in $\mathcal{C}$. Then



(4.18)

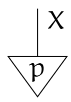
where 2-positivity applies because $g \circ f = \mathrm{id}_X$ is deterministic. ■

Corollary 4.19. *Every invertible morphism in* **FinStoch** *or* **fdC*-AlgCPU** *is deterministic.*

Proof. Combine Lemma 4.17 with Theorem 4.5. ■

Remark 4.20. Note that the transpose map is a positive unital map with a positive unital inverse (itself). Nevertheless, it is clearly not deterministic. This is consistent with our earlier observations that positive unital maps do not form a 2-positive subcategory of **fdC*-AlgU**.

Definition 4.21. Let $\mathcal{M}$ be a quantum Markov category and let $\mathcal{C}$ be a 2-positive subcategory of $\mathcal{M}$. A ***state*** on an object X in $\mathcal{C}$ is a causal morphism $I \overset{p}{\rightsquigarrow} X$ in $\mathcal{C}$. Such a state will be drawn in string-diagrammatic notation as



. (4.22)

Similarly, if Θ and Y are also in $\mathcal{C}$, a morphism $\Theta \overset{q}{\rightsquigarrow} Y$ in $\mathcal{M}$ is ***2-positive*** if it is in $\mathcal{C}$.

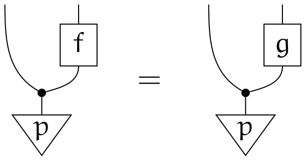
The preceding definition suffices for our main two examples **FinStoch** and **fdC*-AlgCPU**. Indeed, every positive unital functional $\mathcal{A} \rightsquigarrow \mathbb{C}$ on a C*-algebra $\mathcal{A}$ is automatically CP (cf. [32, Theorem 3]). Hence, the states we are considering coincide with the usual states on C*-algebras.

Convention 4.23. In everything that follows, for a given quantum Markov category, we will always work with a 2-positive subcategory unless otherwise stated. This means that for any quantum Markov category discussed, we will implicitly choose a 2-positive subcategory and all 2-positive morphisms and states will be from that subcategory. When working with finite-dimensional C*-algebras, the 2-positive subcategory that we will always pick is **fdC*-AlgCPU**.

5 Almost everywhere equivalence

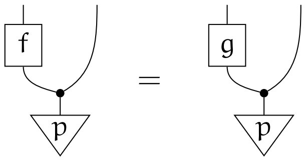
The following definition is based on the insightful observation of Cho and Jacobs that a.e. equivalence has a diagrammatic formulation [3]. However, we distinguish two versions of their definition to isolate the one most suitable for the quantum Markov categories we will work with.

Definition 5.1. Let X and Y be objects in a 2-positive subcategory $\mathcal{C}$ of a quantum Markov category $\mathcal{M}$, let $I \overset{p}{\rightsquigarrow} X$ be a state on X, and let $f, g : X \rightsquigarrow Y$ be morphisms in $\mathcal{M}$. The morphism f is said to be ***right*** p***-a.e. equivalent to*** g iff



(5.2)

Dually, f is said to be ***left*** p***-a.e. equivalent to*** g iff



(5.3)

When f is both right and left p-a.e. equivalent to g, we will say f is p***-a.e. equivalent to*** g, and the notation $f \underset{p}{=} g$ will be used.

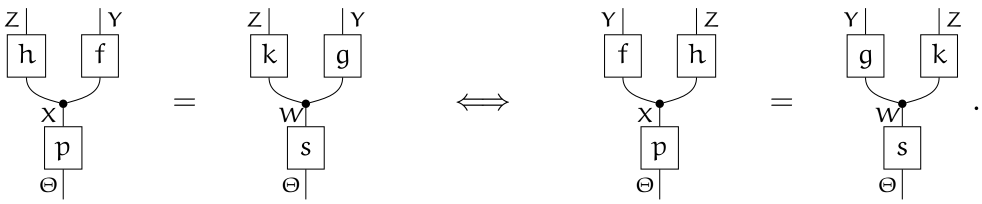
Remark 5.4. One can also replace the state $I \overset{p}{\rightsquigarrow} X$ with an arbitrary 2-positive morphism $\Theta \overset{p}{\rightsquigarrow} X$, as done by Fritz [10, Definition 13.1], to obtain more general notions of p-a.e. equivalence. We will occasionally, but rarely, use this. Note that we demand p to be in $\mathcal{C}$ as opposed to $\mathcal{M}$ to avoid developing an abstract theory of Jordan decompositions in this language.

The fact that we have two notions of a.e. equivalence may seem strange. We will see that for $*$*-preserving* morphisms, the two notions of a.e. equivalence are themselves equivalent. However, if the morphisms are *not* $*$-preserving, which can happen in the quantum setting (cf. Remark 5.26, Proposition 5.42, and Remark 6.10), there are instances where the notions are actually inequivalent.

Proposition 5.5. *Let*

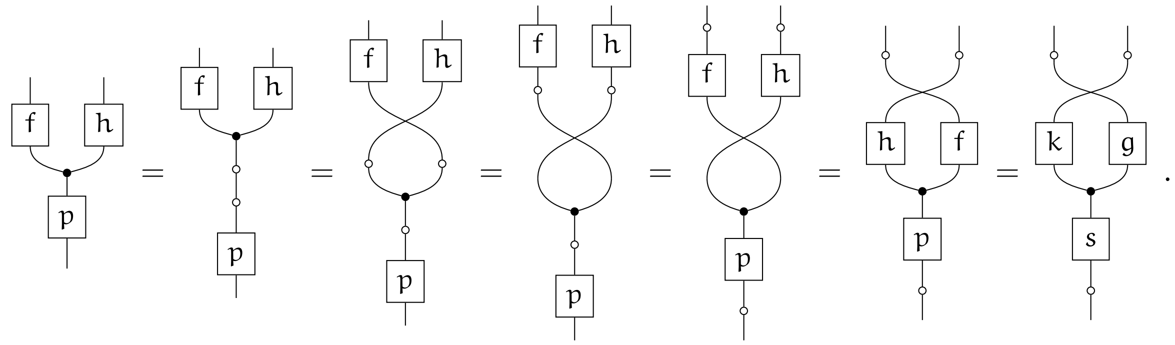
$$\begin{tikzcd} X \arrow[rr, squiggly, "f"] \arrow[dd, squiggly, "h"'] & & Y \\ & \Theta \arrow[ul, squiggly, "p"'] \arrow[dr, squiggly, "s"] & \\ Z & & W \arrow[ll, squiggly, "k"] \arrow[uu, squiggly, "g"'] \end{tikzcd} \tag{5.6}$$

be a (not necessarily commuting) diagram of $*$-preserving morphisms in a quantum Markov category. Then



(5.7)

Proof. Assume the left-hand-side of (5.7) holds. Then



(5.8)

One then rewinds the steps with f replaced by g, h replaced by k, and p replaced by s. A completely analogous argument holds if the right-hand-side of (5.7) is assumed. ∎

Corollary 5.9. *In a quantum Markov category (using the same notation as in Definition 5.1), f is right p-a.e. equivalent to g if and only if f is left p-a.e. equivalent to g provided f, p, and g are $*$-preserving. In particular, if f is left (or right) p-a.e. equivalent to g, then f is p-a.e. equivalent to g.*

Proof. This follows from Proposition 5.5 in the special case described by the diagram

$$\begin{tikzcd} X \arrow[rr, squiggly, "f"] \arrow[dd, "\mathrm{id}_X"'] & & Y \\ & I \arrow[ul, squiggly, "p"'] \arrow[dr, squiggly, "p"] & \\ X & & X \arrow[ll, "\mathrm{id}_X"] \arrow[uu, squiggly, "g"'] \end{tikzcd} . \tag{5.10}$$

∎

Remark 5.11. The $*$-preserving condition in Proposition 5.5, and hence Corollary 5.9, is crucial. Here is a counter-example in the category of finite-dimensional C^*-algebras and unital linear maps. Let $\mathcal{A}$ and $\mathcal{B}$ both be $\mathcal{M}_2(\mathbb{C})$ and let $\omega = \mathrm{tr}(\rho\,\cdot\,)$, where $\rho = \left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right]$. Let $F = \mathrm{id}_{\mathcal{M}_2(\mathbb{C})}$ and set

$$\begin{aligned} \mathcal{M}_2(\mathbb{C}) &\overset{F'}{\rightsquigarrow} \mathcal{M}_2(\mathbb{C}) \\ \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} &\longmapsto \begin{bmatrix} b_{11} & b_{12} \\ 0 & b_{22} \end{bmatrix} . \end{aligned} \tag{5.12}$$

Then, one can easily check that $\omega = \omega \circ F = \omega \circ F'$ and

$$\omega\big(F(B)A\big) = \omega\big(F'(B)A\big) \quad \forall\, A, B, \quad \text{while} \quad \omega\big(AF(B)\big) \neq \omega\big(AF'(B)\big) \quad \forall\, A, B. \tag{5.13}$$

This is because F' is not $*$-preserving. However, if ρ happens to commute with the images of F and F', then the two notions agree and all the expressions in (5.13) are equal (due to the cyclicity of trace). Therefore, one can view the difference of these two notions of a.e. equivalence as being related to the non-commutativity present in the quantum setting.

Our notion of a.e. equivalence for morphisms of C^*-algebras from [27] was motivated by the GNS construction and had little to do with diagrammatic reasoning. Amazingly, our notion coincides with the categorical Definition 5.1 due to Cho and Jacobs [3] *when* the morphisms in question are $*$-preserving. However, there are subtle differences when the morphisms are merely linear (this difference will be important for the notions of a.e. determinism and Bayesian inference).

Lemma 5.14. *Let $\mathcal{A}$ be a C^*-algebra, let $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ be a state, let P_ω be its support, and set $P_\omega^\perp := 1_{\mathcal{A}} - P_\omega$. Then*

$$\omega(A) = \omega(P_\omega A) = \omega(AP_\omega) = \omega(P_\omega AP_\omega) \qquad \forall\, A \in \mathcal{A}. \tag{5.15}$$

In particular, $\omega(P_\omega^\perp A) = 0$ and $\omega(AP_\omega^\perp) = 0$ for all $A \in \mathcal{A}$.

Proof. See Section 1.14 of Sakai [30]. ■

Theorem 5.16. *Let $\mathcal{A}$ and $\mathcal{B}$ be finite-dimensional C^*-algebras (or more generally von Neumann algebras), let $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ be a state on $\mathcal{A}$ with corresponding support P_ω, and let $F, G : \mathcal{B} \rightsquigarrow \mathcal{A}$ be linear maps. Consider the following four conditions.*

(a) *F is left ω-a.e. equivalent to G in the sense of Definition 5.1.*

(b) *F is right ω-a.e. equivalent to G in the sense of Definition 5.1.*

(c) *$F(B)P_\omega = G(B)P_\omega$ for all $B \in \mathcal{B}$.*

(d) *$\omega\Big(\big(F(B) - G(B)\big)^*\big(F(B) - G(B)\big)\Big) = 0$ for all $B \in \mathcal{B}$, i.e. $F(B) - G(B)$ is in the null space $\mathcal{N}_\omega := \{A \in \mathcal{A} \,:\, \omega(A^*A) = 0\}$ of ω for all $B \in \mathcal{B}$.*

Then the following facts hold.

i. *Conditions (b), (c), and (d) are equivalent.*

ii. *If F and G are $*$-preserving, then all conditions are equivalent.*

Proof.

i. The equivalence between conditions (c) and (d) is not difficult to show if one recalls the identity $\mathcal{N}_\omega = \mathcal{A}P_\omega^\perp$ (for a proof anyway, see Lemma 3.42 in [27]). To see that (b) is equivalent to (c), first suppose (b) holds. This means

$$\omega\big(AF(B)\big) = \omega\big(AG(B)\big) \qquad \forall\, A \in \mathcal{A},\ B \in \mathcal{B}. \tag{5.17}$$

By linearity of ω, this is equivalent to

$$\omega\Big(A\big(F(B) - G(B)\big)\Big) = 0 \qquad \forall\, A \in \mathcal{A},\ B \in \mathcal{B}. \tag{5.18}$$

In particular, one can set $A := \big(F(B) - G(B)\big)^*$. This immediately gives condition (d), and hence (c). Now, suppose (c) holds. Then

$$\omega\big(AF(B)\big) \xlongequal{\text{Lemma 5.14}} \omega\big(AF(B)P_\omega\big) = \omega\big(AG(B)P_\omega\big) \xlongequal{\text{Lemma 5.14}} \omega\big(AG(B)\big) \tag{5.19}$$

for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. This proves (c) implies (b). Thus, the last three conditions have been shown to be equivalent.

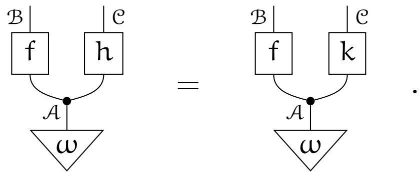
ii. This follows from the previous steps and Corollary 5.9 (ω is $*$-preserving because it is a state), which proves (a) is equivalent to (b).

$\blacksquare$

Remark 5.20. One of the convenient properties of condition (c) in Theorem 5.16 is that it is linear and involves only a single variable, as opposed to the definition of right a.e. equivalence from (5.2), which involves two variable inputs. More generally, we have the following result. Given a diagram

$$
\begin{tikzcd}
\mathcal{A} \arrow[dr, squiggly, "\omega"] & & \mathcal{B} \arrow[ll, squiggly, "f"'] \arrow[dd, squiggly, "f"] \\
& \mathbb{C} & \\
\mathcal{C} \arrow[uu, squiggly, "h"] \arrow[rr, squiggly, "k"'] & & \mathcal{A} \arrow[ul, squiggly, "\omega"']
\end{tikzcd}
\tag{5.21}
$$

in **fdC*-AlgCPU**, if $h(C)P_\omega = k(C)P_\omega$ for all $C \in \mathcal{C}$, then



(5.22)

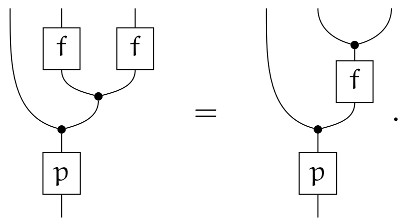
However, the converse is not true in general. A simple example is given by the following. Set $\mathcal{A} := \mathcal{M}_2(\mathbb{C})$, $\mathcal{B} := \mathbb{C}$, $\mathcal{C} := \mathcal{M}_2(\mathbb{C})$, and $\omega := \frac{1}{2}\mathrm{tr}$. Also, set $f :=!_{\mathcal{M}_2}$, $h := \mathrm{id}$, and $k := \mathrm{Ad}_{\left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]}$. Then $P_\omega = \mathbb{1}_2$ and the equality (5.22) holds, but $h(C)P_\omega = h(C) \neq k(C) = k(C)P_\omega$.

We have two reasonable notions of being deterministic almost everywhere. Although we mostly work with states, we include the generalizations for 2-positive morphisms for future reference (cf. Remark 5.4).

Definition 5.23. Let $\Theta \overset{p}{\rightsquigarrow} X$ be 2-positive and let $X \overset{f}{\rightsquigarrow} Y$ be a morphism in a quantum Markov category. The morphism f is p***-a.e. equivalent to a deterministic morphism*** iff there exists a deterministic morphism $X \overset{g}{\rightsquigarrow} Y$ such that $f \underset{p}{=} g$.

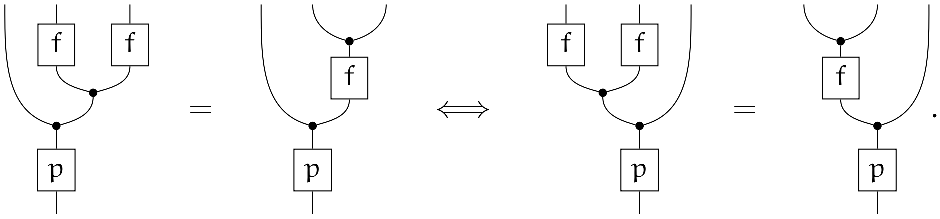
This definition is a-priori different from the following notion, introduced recently by Fritz [10, Definition 13.10].

Definition 5.24. Let $\Theta \overset{p}{\rightsquigarrow} X$ be 2-positive and let $X \overset{f}{\rightsquigarrow} Y$ be a morphism in a quantum Markov category. Then f is p***-a.e. deterministic*** iff

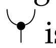


(5.25)

Remark 5.26. In a classical Markov category,



(5.27)

However, this is *not* generally true in a quantum Markov category *even if* f *is* $*$*-preserving*. This is exactly because  is not necessarily $*$-preserving. This has important consequences for a.e. determinism even in the category of finite-dimensional C^*-algebras. In the non-commutative context, (5.25) might be more appropriately called *right* p-a.e. deterministic. However, since we will mostly use (5.25), we prefer to drop the word 'right.'

Example 5.28. Using the same notation as in Definition 5.24 but in the category **FinStoch**, a stochastic map f is p-a.e. deterministic if and only if

$$f_{yx}f_{y'x}p_{x\theta} = \delta_{y'y}f_{yx}p_{x\theta} \qquad \forall\, \theta \in \Theta,\ x \in X,\ y, y' \in Y. \tag{5.29}$$

What this entails will be spelled out in more detail in Proposition 5.41.

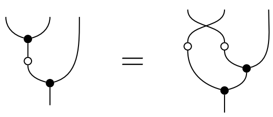
Example 5.30. In the quantum Markov category **fdC*-AlgU**, given CPU maps $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ and a state $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$, F is ω-a.e. deterministic if and only if

$$F(B_1)F(B_2)P_\omega = F(B_1B_2)P_\omega \qquad \forall\, B_1, B_2 \in \mathcal{B}. \tag{5.31}$$

This claim follows from Theorem 5.16. This equation gives us a reasonable notion of a.e. determinism that provides interesting consequences (cf. Theorem 8.3 for example).

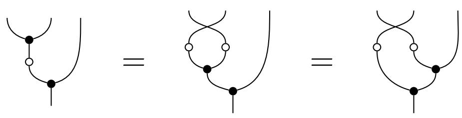
We will now examine in a bit more detail the two notions of almost everywhere determinism introduced. We will first prove a fact regarding the interaction between the involution $*$ and the associativity of the copy morphism Δ.

Lemma 5.32. *In a quantum Markov category,*



(5.33)

Proof of Lemma 5.32.

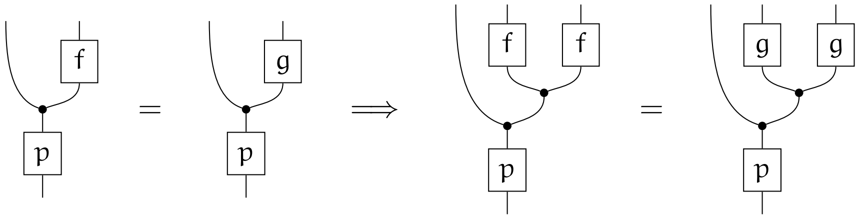


(5.34)

■

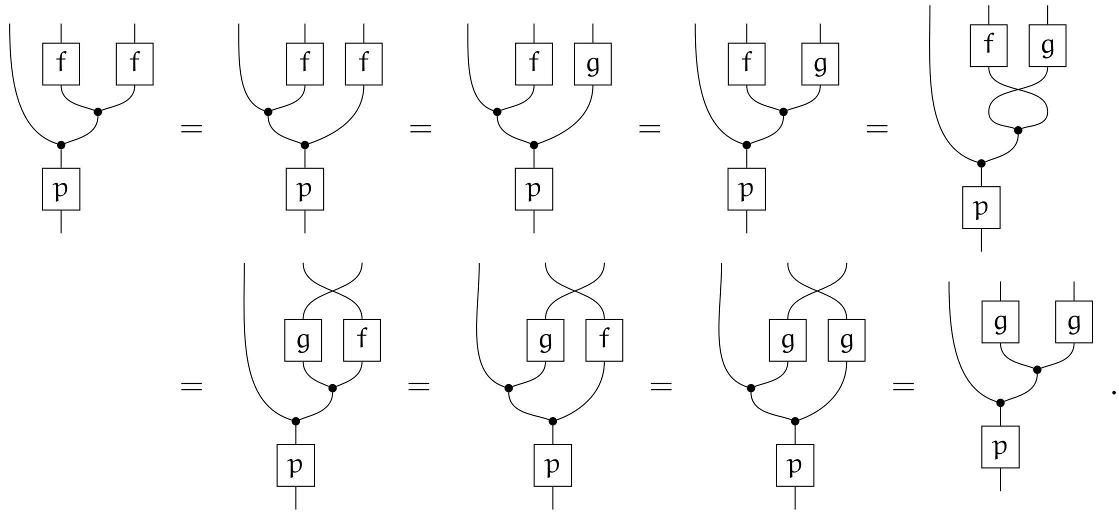
Example 5.35. In any of the categories of finite-dimensional C^*-algebras, Lemma 5.32 says $(AB)^*C = B^*(A^*C)$ for all A, B, C in any fixed C^*-algebra.

Lemma 5.36. *In a classical Markov category,*



(5.37)

Proof. This follows from

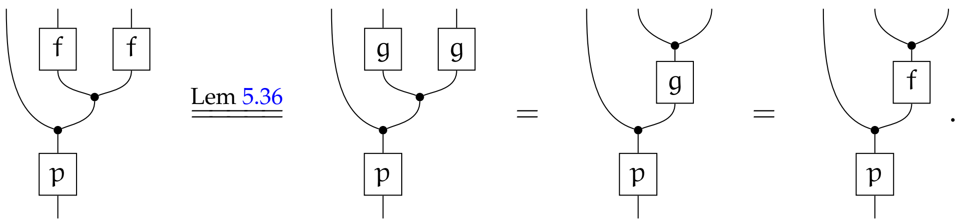


(5.38)

■

Proposition 5.39. *Let* $\Theta \overset{p}{\rightsquigarrow} X$ *be a 2-positive morphism in a classical Markov category. If* $X \overset{f}{\rightsquigarrow} Y$ *is* p*-a.e. equivalent to a deterministic map* $X \overset{g}{\rightsquigarrow} Y$*, then* f *is* p*-a.e. deterministic.*

Proof. This follows from



(5.40)

■

Proposition 5.41. *Given* $\Theta \overset{p}{\rightsquigarrow} X \overset{f}{\rightsquigarrow} Y$ *in* **FinStoch**, f *is* p*-a.e. deterministic if and only if* f *is* p*-a.e. equivalent to a deterministic map.*

Proof. We have already proved the reverse direction in greater generality for an arbitrary classical Markov category in Proposition 5.39. Hence, assume f is p-a.e. deterministic. Fix $x \in X$ and $y, y' \in Y$. By Example 5.28, $f_{yx}f_{y'x} = \delta_{yy'}f_{yx}$ if there exists a θ such that $p_{x\theta} > 0$. In this case, $f_{yx}f_{y'x} = 0$ when $y \neq y'$ and $(f_{yx})^2 = f_{yx}$. This means $f_{yx} \in \{0, 1\}$. Since f_x is a probability measure, this implies there exists a unique y such that $f_{yx} = 1$. Hence, for such x, set $g_{yx} := f_{yx}$. Now, if x is such that $p_{x\theta} = 0$ for all θ, then set g_x to be any (unit) point measure. Then g is deterministic and $f \underset{p}{=} g$. ■

One might try to use Lemma 5.32 in an attempt to prove Lemma 5.36 in an arbitrary quantum Markov category (provided the morphisms are $*$-preserving). However, the claim turns out to be false. In fact, in the category of finite-dimensional C^*-algebras and CPU maps, if a CPU map is a.e. deterministic, it is not necessary equivalent to a deterministic map (see Remarks 8.19 and 8.21 for a counter-example).

Proposition 5.42. *Lemma 5.36 does not generally hold in a quantum Markov category, even if all morphisms are $*$-preserving.*

Proof. We will supply a simple counter-example in the subcategory **fdC*-AlgCPU** of the quantum Markov category **fdC*-AlgU**. Set $\mathcal{B} \equiv \mathcal{A} := \mathcal{M}_2(\mathbb{C})$ and set $\rho := \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$. Let $\omega := \operatorname{tr}(\rho\,\cdot\,)$ be its associated state, and let P_ω denote its support (in this case, $\rho = P_\omega$). For any $\lambda \in (0,1)$, set

$$
\begin{aligned}
F &:= \lambda \mathrm{id} + (1-\lambda)\mathrm{Ad}_{P_\omega} + (1-\lambda)\mathrm{Ad}_{\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]} \circ \mathrm{Ad}_{P_\omega} \quad \text{and} \\
G &:= \lambda \mathrm{id} + (1-\lambda)\mathrm{Ad}_{P_\omega} + (1-\lambda)\mathrm{Ad}_{P_\omega^\perp},
\end{aligned}
\tag{5.43}
$$

which explicitly shows that F and G are CP. In terms of their action on matrices, these maps are given by

$$
F\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) := \begin{bmatrix} a & \lambda b \\ \lambda c & (1-\lambda)a + \lambda d \end{bmatrix} \quad \text{and} \quad G\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) := \begin{bmatrix} a & \lambda b \\ \lambda c & d \end{bmatrix}, \tag{5.44}
$$

from which it easily follows F and G are unital, and therefore CPU. These maps are ω-a.e. equivalent because multiplying both expressions by P_ω on the right gives the same result (we are freely using the equivalent notions of a.e. equivalence from Theorem 5.16 because F and G are CP and hence $*$-preserving). Using these formulas, we find

$$
F\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)^2 P_\omega - G\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)^2 P_\omega = \begin{bmatrix} 0 & 0 \\ \lambda c\big((1-\lambda)a + (\lambda - 1)d\big) & 0 \end{bmatrix}, \tag{5.45}
$$

which is non-zero in general. Therefore, the equality on the right-hand-side of (5.37) does not hold in **fdC*-AlgCPU**. ■

Remark 5.46. Although we have proved a.e. determinism does not imply a.e. equivalence to a deterministic map for CPU maps, we have not made any claims regarding whether being a.e. equivalent to a *-homomorphism implies a.e. determinism. We leave this question open.

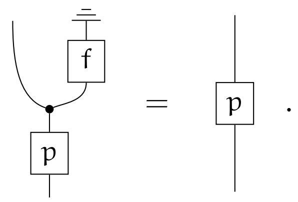
Remark 5.47. Let $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ be a CPU map, let $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ be a state, and let $\xi := \omega \circ F$ be the pullback state. Knowing what a CPU map does on $P_\xi \mathcal{B} P_\xi$ does not uniquely determine the map on $\mathcal{B} P_\xi$. It is also not enough to know the value of that CPU map followed by Ad_{P_ω}. For example, for a CPU map $\mathcal{M}_n(\mathbb{C}) \overset{F}{\rightsquigarrow} \mathcal{M}_n(\mathbb{C})$ and the density matrix $\rho = e_1 e_1^*$ (e_1 is the first standard unit vector of $\mathbb{C}^n$) with associated state $\omega = \operatorname{tr}(\rho\,\cdot\,)$, if $\mathrm{Ad}_{P_\omega} \circ F$ is ω-a.e. equivalent to Ad_{P_ω}, then F is not necessarily equal to the identity. However, if F is ω-a.e. equivalent to the identity, then it *is* equal to the identity. The latter result was proved in [27, Theorem 3.67].

For the former statements, consider the $n = 2$ case take the CPU map (which is even a *-isomorphism)

$$F\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) := \mathrm{Ad}_{\left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a & -c \\ -b & d \end{bmatrix}. \tag{5.48}$$

Then $\omega = \omega \circ F$, $\mathrm{Ad}_{P_\omega} \circ F = \mathrm{Ad}_{P_\omega}$, and $F \circ \mathrm{Ad}_{P_\omega} = \mathrm{Ad}_{P_\omega}$, but $F \neq \mathrm{id}_{\mathcal{M}_n}$. Thus, one should keep in mind that there is a good deal of information about $F(B)$ in $F(B)P_\omega$, which would be lost if one only worked with $P_\omega F(B) P_\omega$. This remark will be important in subsequent work [26] when we compare and contrast our definition of quantum Bayesian inference to the Bayesian inference of Leifer [18].

Definition 5.49. Let $\Theta \overset{p}{\rightsquigarrow} X \overset{f}{\rightsquigarrow} Y$ be a composable pair of morphisms in a quantum Markov category where p is 2-positive. The morphism f is said to be p-***a.e. causal*** iff



(5.50)

Example 5.51. Given a state $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ in **fdC*-Alg** (recall the end of Example 3.12 for our definition of **fdC*-Alg**), a positive map $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ is ω-a.e. causal if and only if $F(1_\mathcal{B})P_\omega = P_\omega$ by Theorem 5.16 and the first axiom in (3.3). Since $P_\omega^\perp F(1_\mathcal{B})P_\omega = P_\omega^\perp P_\omega = 0$ and $P_\omega F(1_\mathcal{B})P_\omega^\perp = (P_\omega^\perp F(1_\mathcal{B})P_\omega)^* = 0$ (because F is $*$-preserving), we conclude

$$F(1_\mathcal{B}) = P_\omega F(1_\mathcal{B}) P_\omega + P_\omega^\perp F(1_\mathcal{B}) P_\omega^\perp = P_\omega + \mathrm{Ad}_{P_\omega^\perp}\big(F(1_\mathcal{B})\big). \tag{5.52}$$

This guarantees $F(1_\mathcal{B}) \geqslant P_\omega$ since F is a positive map. In the case $\mathcal{A} = \mathbb{C}^X$ and $\mathcal{B} = \mathbb{C}^Y$, the state ω corresponds to a probability measure $p : I \rightsquigarrow X$, and this condition means that the corresponding map $f : X \rightsquigarrow Y$ associates to each $x \in X \setminus N_p$ a probability measure on Y. However, it can assign *any* (possibly signed) measure on Y to each element of N_p. Indeed, condition (5.50) provides us with the equation

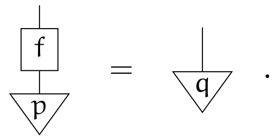
$$p_x = \sum_{y \in Y} f_{yx} p_x = p_x \sum_{y \in Y} f_{yx} \qquad \forall\, x \in X. \tag{5.53}$$

When $x \in X \setminus N_p$, this gives the constraint $\sum_{y \in Y} f_{yx} = 1$, but when $x \in N_p$, this gives no condition.

6 Abstract disintegrations and Bayesian inversion

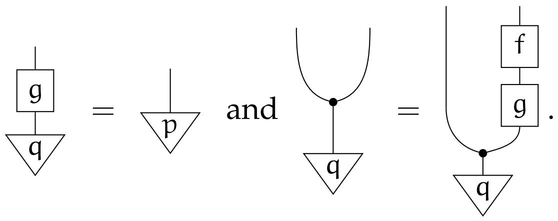
Definition 6.1. Let $\mathcal{M}$ be a quantum Markov category and let $\mathcal{C}$ be a 2-positive subcategory of $\mathcal{M}$ that contains all the objects of $\mathcal{M}$. Given states $I \overset{p}{\rightsquigarrow} X$ and $I \overset{q}{\rightsquigarrow} Y$ (which are in $\mathcal{C}$), a causal morphism $X \overset{f}{\rightsquigarrow} Y$ in $\mathcal{M}$ is said to be ***state-preserving*** iff

$$\begin{tikzcd} & I \arrow[dl, rightsquigarrow, "p"'] \arrow[dr, rightsquigarrow, "q"] \arrow[d, phantom, "="] & \\ X \arrow[rr, rightsquigarrow, "f"'] & & Y \end{tikzcd} \quad \text{i.e.} \tag{6.2}$$



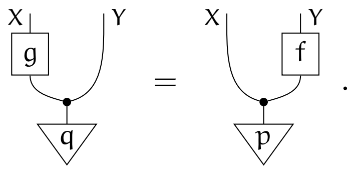
Such data will be denoted by (f, p, q). A ***disintegration*** of (f, p, q) is a causal morphism $Y \overset{g}{\rightsquigarrow} X$ such that

$$\begin{tikzcd} & I \arrow[dl, rightsquigarrow, "p"'] \arrow[dr, rightsquigarrow, "q"] \arrow[d, phantom, "="] & \\ X & & Y \arrow[ll, rightsquigarrow, "g"] \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} & X \arrow[dl, rightsquigarrow, "f"'] \arrow[d, phantom, "=_q"] & \\ Y & & Y \arrow[ll, "\mathrm{id}_Y"] \arrow[ul, rightsquigarrow, "g"'] \end{tikzcd}, \quad \text{i.e.} \tag{6.3}$$



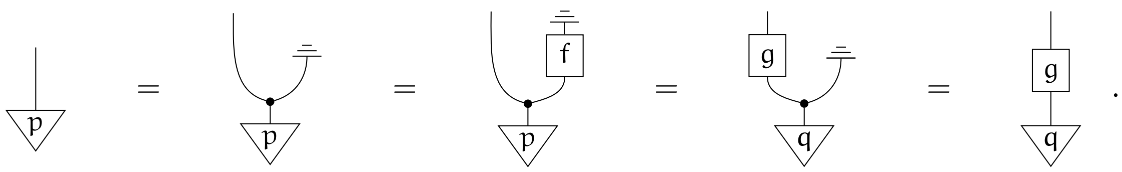
A disintegration is also called a ***regular conditional probability*** and an ***optimal hypothesis***. A ***Bayesian inverse*** of (or a ***Bayesian inference*** for) (f, p, q) is a causal morphism $Y \overset{g}{\rightsquigarrow} X$ such that[14]

$$\begin{tikzcd} Y \arrow[d, "\Delta_Y"'] & I \arrow[l, rightsquigarrow, "q"'] \arrow[r, rightsquigarrow, "p"] \arrow[d, phantom, "="] & X \arrow[d, "\Delta_X"] \\ Y \times Y \arrow[r, rightsquigarrow, "g \times \mathrm{id}_Y"'] & X \times Y & X \times X \arrow[l, rightsquigarrow, "\mathrm{id}_X \times f"] \end{tikzcd}, \quad \text{i.e.} \tag{6.4}$$



This condition will be referred to as ***the Bayes condition***. It will often (though not always) be assumed that the morphisms f and g are $*$-preserving and belong to $\mathcal{C}$.

Remark 6.5. It is a consequence of causality of f that g preserves states in the definition of a Bayesian inverse. Indeed,



(6.6)

However, causality is not necessary for this. It is enough that f is p-a.e. causal. Similarly, if g is not assumed to be causal, then g is still q-a.e. causal if g, q, and p are $*$-preserving. Once we learn that disintegrations are Bayesian inverses, this will also imply the same for disintegrations. Nevertheless, we will not dwell on a.e. causality and mostly work with causal morphisms since these physically correspond to probability-preserving processes.

Remark 6.7. The definition of a disintegration in Definition 6.1 differs from the one introduced in [27] in that we are no longer assuming $X \overset{f}{\rightsquigarrow} Y$ is deterministic. The reason for this is to

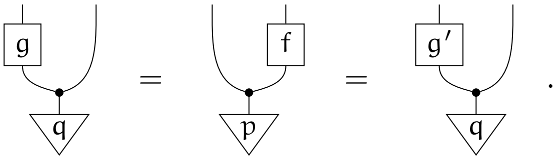
[14] In other words, one can 'slide' either stochastic map over Δ but this swaps the probability measures and the stochastic maps.

apriori allow the possibility for more morphisms to have disintegrations. However, it turns out that a morphism $X \overset{f}{\rightsquigarrow} Y$ together with a state $I \overset{p}{\rightsquigarrow} X$ in **FinStoch** has a disintegration if and only if f is p-a.e. deterministic (cf. Theorem 7.8). In the more general setting of finite-dimensional C^*-algebras, we prove a similar result: if f has a CPU disintegration, then f is p-a.e. deterministic (cf. Theorem 8.3).

The following Lemma shows that Bayesian inverses are a.e. unique whenever they are $*$-preserving.

Lemma 6.8. *Let $I \overset{p}{\rightsquigarrow} X \overset{f}{\rightsquigarrow} Y$ be a composable pair of morphisms in a quantum Markov category with p and $q := f \circ p$ $*$-preserving states and f causal. If $g, g' : Y \rightsquigarrow X$ are two Bayesian inverses of (f, p, q) such that both g and g' are $*$-preserving, then $g \underset{q}{=} g'$.*

Proof. By assumption,



(6.9)

Applying Corollary 5.9 gives the required result. ∎

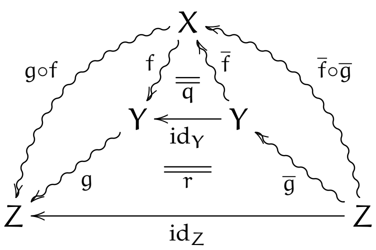
Without assuming g and g' are $*$-preserving in Lemma 6.8, we can only conclude that g is *left* q-a.e. equivalent to g'. Since our convention of a.e. equivalence in the non-commutative setting is *right* a.e. equivalence, this does not agree with our convention (unless g and g' are $*$-preserving or satisfy some other specialized condition).

Remark 6.10. In the category **fdC*-AlgU**, if two Bayesian inverses are not $*$-preserving, then they need not be a.e. equivalent. Explicit examples are provided in [28].

The following proposition shows that disintegrations are functorial in appropriate quantum Markov categories.

Proposition 6.11. *Let $\mathcal{C} \subseteq \mathcal{M}$ be a 2-positive subcategory of a quantum Markov category where composing state-preserving a.e. equivalence classes of morphisms in $\mathcal{C}$ is well-defined. If $\overline{g} : Z \rightsquigarrow Y$ and $\overline{f} : Y \rightsquigarrow X$ are disintegrations of $(X \overset{f}{\rightsquigarrow} Y, I \overset{p}{\rightsquigarrow} X, q := f \circ p)$ and $(Y \overset{g}{\rightsquigarrow} Z, I \overset{q}{\rightsquigarrow} Y, r := g \circ q)$, respectively, then $\overline{f} \circ \overline{g}$ is a disintegration of $(g \circ f, p, r)$.*

Proof. The probability-preserving condition is immediate. The composite $\overline{f} \circ \overline{g}$ is causal because the composite of causal morphisms is causal. The second condition follows from



, (6.12)

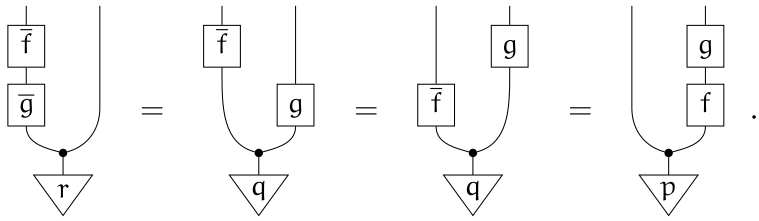
since composing a.e. equivalence classes of morphisms is well-defined. ∎

Remark 6.13. The composition of a.e. equivalence classes of morphisms is well-defined in our main two examples, namely in **FinStoch** and **fdC*-AlgCPU**. See Proposition 3.106 in [27] for precise details and a proof.

Bayesian inversion, on the other hand, is functorial in *any* quantum Markov category.

Proposition 6.14. *If* $Z \overset{\bar{g}}{\rightsquigarrow} Y$ *and* $Y \overset{\bar{f}}{\rightsquigarrow} X$ *are Bayesian inverses of* $(X \overset{f}{\rightsquigarrow} Y, I \overset{p}{\rightsquigarrow} X, q := f \circ p)$ *and* $(Y \overset{g}{\rightsquigarrow} Z, I \overset{q}{\rightsquigarrow} Y, r := g \circ q)$*, respectively, then* $\bar{f} \circ \bar{g}$ *is a Bayesian inverse of* $(g \circ f, p)$.

Proof. As before $\bar{f} \circ \bar{g}$ is causal. Secondly, the calculation



(6.15)

proves the Bayes condition. ■

Proposition 6.16. *Given causal* $*$*-preserving morphisms*

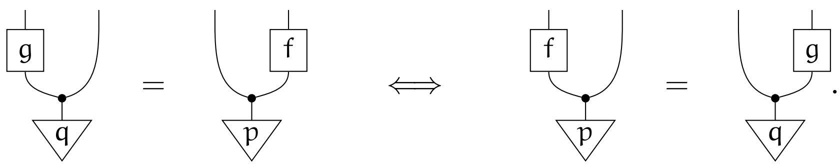
$$\begin{tikzcd} & I \arrow[dl, squiggly, "p"'] \arrow[dr, squiggly, "q"] & \\ X \arrow[rr, squiggly, shift left, "f"] & & Y \arrow[ll, squiggly, shift left, "g"] \end{tikzcd} \tag{6.17}$$

(with p *and* q *states) in a quantum Markov category, if* $q = f \circ p$ *and* g *is a Bayesian inverse of* (f, p, q)*, then* f *is a Bayesian inverse of* (g, q, p).

Proof. This is an immediate consequence of Proposition 5.5 applied to the diagram

$$\begin{tikzcd} X \arrow[rr, squiggly, "f"] \arrow[dd, "\mathrm{id}_X"'] & & Y \\ & I \arrow[ul, squiggly, "p"] \arrow[dr, squiggly, "q"'] & \\ X & & Y \arrow[ll, squiggly, "g"] \arrow[uu, "\mathrm{id}_Y"'] \end{tikzcd} , \tag{6.18}$$

which entails



(6.19)

The implication towards the right proves the proposition. ■

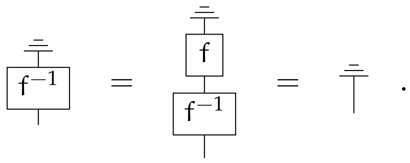
Remark 6.20. The $*$-preserving assumption for the Bayesian inverses in Proposition 6.16 is crucial. This is due to Remark 5.11.

Proposition 6.21. *Let* $I \overset{p}{\rightsquigarrow} X$ *be a state, let* $X \overset{f}{\rightsquigarrow} Y$ *be causal and 2-positive and set* $q := f \circ p$.

i. If f *is invertible, then* f^{-1} *is a disintegration of* (f, p, q).

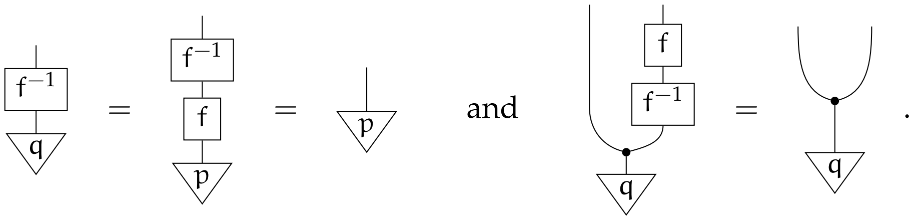
ii. If f *is invertible with a 2-positive inverse* f^{-1}, *then* f^{-1} *is a Bayesian inverse of* (f, p, q).

Proof. If f is causal, then f^{-1} is causal as well because



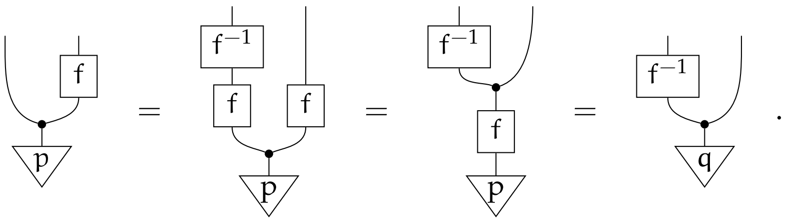
(6.22)

i. The disintegration conditions for f^{-1} follow from



(6.23)

ii. The Bayes condition for f^{-1} follows from 2-positivity since $f^{-1} \circ f = \mathrm{id}_X$ is deterministic:



(6.24)

■

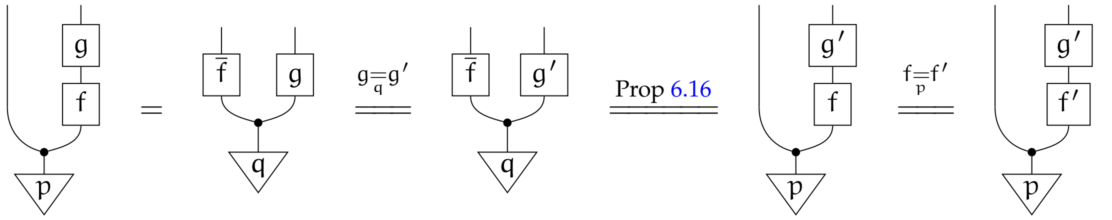
The following remark explains that Bayesian inversion defines a dagger functor on a.e. equivalence classes. This is a generalization of Remark 13.9 in [10] to the non-commutative setting.[15] Note, however, that we have not described *when* Bayesian inverses exist, so this is not a substantial generalization of Fritz' result other than the fact that we can now include quantum probability. Conditions for the *existence* of quantum Bayesian inverses will be addressed in forthcoming work [28].

Remark 6.25. Let $\mathcal{C} \subseteq \mathcal{M}$ be a 2-positive subcategory of a quantum Markov category containing only $*$-preserving morphisms.[16] Let $I_{/\mathcal{C}}$ be the category whose objects are pairs $(X, I \overset{p}{\rightsquigarrow} X)$ (with X in $\mathcal{C}$ and p a state) and a morphism from $(X, I \overset{p}{\rightsquigarrow} X)$ to $(Y, I \overset{q}{\rightsquigarrow} Y)$ is a morphism

[15]The proof in [10, Proposition 13.8 and Remark 13.9] uses the notion of a causal subcategory (cf. [10, Definition 11.30]) of a Markov category. However, since we have not checked if **fdC*-AlgCPU** is causal in this sense, our argument will be slightly different.

[16]We have not checked if 2-positive morphisms are necessarily $*$-preserving. They are in our main two examples **FinStoch** and **fdC*-AlgCPU**.

$X \overset{f}{\rightsquigarrow} Y$ in $\mathcal{C}$ such that $f \circ p = q$. Such morphisms are called ***state-preserving***. Let $\mathfrak{B}\mathrm{I}_{/\mathcal{C}}$ be the subcategory of $\mathrm{I}_{/\mathcal{C}}$ consisting of the same objects as $\mathrm{I}_{/\mathcal{C}}$ but whose morphisms consist of Bayesian invertible morphisms (whose Bayesian inverses are also in $\mathcal{C}$). Consider two a.e. equivalent pairs of composable morphisms $f, f' : (X, I \overset{p}{\rightsquigarrow} X) \to (Y, I \overset{q}{\rightsquigarrow} Y)$ and $g, g' : (Y, I \overset{q}{\rightsquigarrow} Y) \to (Z, I \overset{r}{\rightsquigarrow} Z)$, i.e. $f \underset{p}{=} f'$ and $g \underset{q}{=} g'$, in $\mathfrak{B}\mathrm{I}_{/\mathcal{C}}$. Then $g \circ f \underset{p}{=} g' \circ f'$ follows from



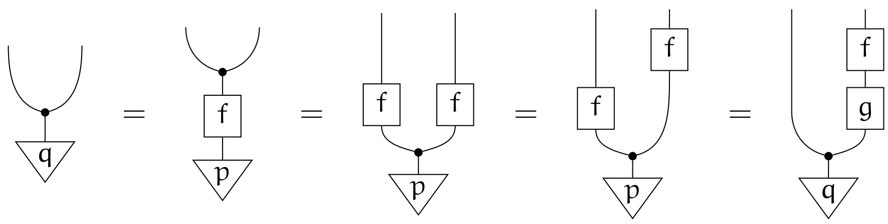
, (6.26)

where we have used a Bayesian inverse $\bar{f}$ for (f, p, q) in the first equality. One can also show if f is p-.a.e. equivalent to f' and f has $\bar{f}$ as a Bayesian inverse, then $\bar{f}$ is also a Bayesian inverse of f'. It is even easier to check that the identity is a Bayesian inverse of the identity for any states. Hence, taking equivalence classes of morphisms in $\mathfrak{B}\mathrm{I}_{/\mathcal{C}}$ defines a category, which will be denoted by $\mathfrak{B}^{ae}\mathrm{I}_{/\mathcal{C}}$. It consists of a.e. equivalence classes of causal $*$-preserving 2-positive Bayesian invertible morphisms. These facts together with Propositions 6.14 and 6.16 say that Bayesian inversion defines a dagger functor on $\mathfrak{B}^{ae}\mathrm{I}_{/\mathcal{C}}$. This is to be contrasted with the notion of having a disintegration. Even in categories where composition of state-preserving a.e. classes is well-defined so that disintegrations are compositional when they exist, if f has a disintegration g, it is almost never the case that f is a disintegration of g. More on this will be explained in examples later in this work when we discover that having a disintegration of (f, p, q) imposes constraints on f.

We now state a theorem that provides our first indication of how disintegrations are related to Bayesian inverses. We will see many theorems of this sort throughout this work. However, the following theorem distinguishes itself in that we have a direct proof internal in the language of string diagrams.

Theorem 6.27. *Let $I \overset{p}{\rightsquigarrow} X$ be a state on X, let $X \xrightarrow{f} Y$ be a deterministic map, and set $q := f \circ p$. If f has a Bayesian inverse $Y \overset{g}{\rightsquigarrow} X$, then g is a disintegration of (f, p, q).*

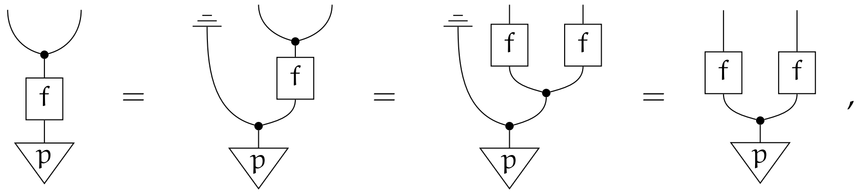
Proof. The state-preserving condition of a disintegration follows from Remark 6.5. The other condition of a disintegration follows from the simple string diagram calculation:



. (6.28)

The fact that f is deterministic was used in the second equality. ■

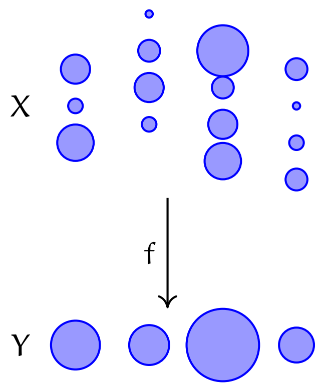
Remark 6.29. The conclusion of Theorem 6.27 remains true if f is merely p-a.e. deterministic and p-a.e. causal. The proof is given by filling in the second equality in (6.28) with



(6.30)

and then completing the calculation just as in (6.28).

7 Classical Bayesian inference



In what follows, we gather some facts about inverses, disintegrations and Bayesian inference in **FinStoch**. Many of the results here are generalized in Section 8 in the quantum setting, and the reader interested in the quantum-mechanical side may feel free to skip immediately to that section. There are two main results of interest here. Theorem 7.8 states that if a disintegration exists, then the original map is a.e. deterministic. Theorem 7.11 states that disintegrations are Bayesian inverses. Recall the notation that given a probability measure $\{\bullet\} \overset{q}{\rightsquigarrow} Y$ on a finite set Y, we call (Y, q) a ***finite probability space*** and we let $N_q \subseteq Y$ denote the ***null space*** of q, i.e. $N_q := \{y \in Y \ : \ q_y = 0\}$. It may be helpful to visualize a probability-preserving map $(X, p) \xrightarrow{f} (Y, q)$ in terms of combining water droplets as in the figure on the right [12], [27, Section 2.2]. A disintegration of (f, p, q) then splits the water droplets back into the form above.

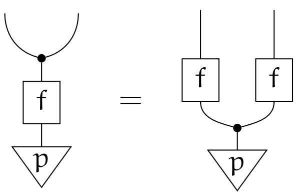
Theorem 7.1. *Let (X, p) and (Y, q) be finite probability spaces. Let $f : X \to Y$ be a measure-preserving function. Then there exists a disintegration $g : Y \rightsquigarrow X$ of (f, p, q). Moreover, g is unique q-a.e. and a formula for such a (representative of a) disintegration is given by*

$$g_{xy} := \begin{cases} p_x \delta_{yf(x)}/q_y & \text{if } y \in Y \setminus N_q \\ 1/|X| & \text{if } y \in N_q \end{cases}. \tag{7.2}$$

Proof. See Section 2.2 in [27] for details. ∎

However, disintegrations of *stochastic maps* (as opposed to functions) do not always exist. We will show why in Theorem 7.8. Before we explain this, we will prove some interesting facts about disintegrations. The first few facts establish how disintegrations generalize inverses.

Lemma 7.3. *A stochastic map $f : X \rightsquigarrow Y$ is deterministic if and only if*



(7.4)

for all states $p : \{\bullet\} \rightsquigarrow X$.

Proof. The forward implication follows immediately from the definition of f being deterministic. For the reverse implication, fix $x \in X$ and let p be the Dirac measure at x. The assumption then reads $(\Delta_Y \circ f \circ p)_{(y,y')} = f_{yx}\delta_{yy'} = f_{yx}f_{y'x} = \big((f \times f) \circ \Delta_X \circ p\big)_{(y,y')}$ for all $y, y' \in Y$. Setting $y' = y$ gives $f_{yx} \in \{0,1\}$. Since x was arbitrary and $\sum_{y \in Y} f_{yx} = 1$, f is deterministic. ■

Proposition 7.5. *A deterministic map* $X \xrightarrow{f} Y$ *is invertible if and only if both of the following facts hold.*

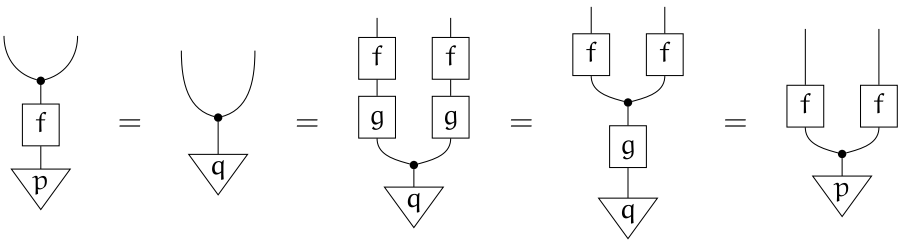
i. For every probability measure $\{\bullet\} \overset{q}{\rightsquigarrow} Y$*, there exists a probability measure* $\{\bullet\} \overset{p}{\rightsquigarrow} X$ *such that* $q = f \circ p$.

ii. For every probability measure $\{\bullet\} \overset{p}{\rightsquigarrow} X$*, f has a deterministic disintegration of* $(f, p, q := f \circ p)$.

Proof. If f is invertible, then $p := f^{-1} \circ q$ satisfies $q = f \circ p$. Furthermore, f^{-1} is a deterministic disintegration. Conversely, suppose f satisfies the two conditions. The first condition implies f is surjective by setting $q := \delta_y$ for various $y \in Y$. The second condition implies f is injective by setting $p_x := \frac{1}{|X|}$ for all $x \in X$. In more detail, let g be a deterministic disintegration of (f, p, q). If f were not injective, then there exists a $y \in Y$ such that $f^{-1}(\{y\})$ contains more than a single element. Since $q_y > 0$ and $p_x > 0$ for all $x \in f^{-1}(\{y\})$, it must be that $1 > g_{xy} > 0$ for those same values of x. This contradicts the fact that g is deterministic. ■

Corollary 7.6. *A stochastic map* $X \overset{f}{\rightsquigarrow} Y$ *is invertible if and only if both of the conditions in Proposition 7.5 hold.*

Proof. Suppose f is invertible. Then f is deterministic by Corollary 4.19. Hence, the forward implication in Proposition 7.5 applies. Conversely, suppose the two conditions of Proposition 7.5 hold. Let $p : \{\bullet\} \rightsquigarrow X$ be an arbitrary state and let g be a deterministic disintegration of $(f, p, q := f \circ p)$. Then,



(7.7)

(the second equality follows from two applications of the equation $f \circ g \underset{q}{=} \mathrm{id}_Y$ and the interchange law). Since p was arbitrary, f is deterministic by Lemma 7.3. Hence, f is invertible by the reverse implication in Proposition 7.5. ■

Theorem 7.8. *Let* $\{\bullet\} \overset{p}{\rightsquigarrow} X$ *be a probability measure on* X*, let* $X \overset{f}{\rightsquigarrow} Y$ *be a stochastic map, and set* $q := f \circ p$*. If there exists a disintegration of* (f, p, q)*, then f is p-a.e. equivalent to a deterministic map.*

Proof. Let $Y \overset{g}{\rightsquigarrow} X$ be such a disintegration and let $x_0 \in X \setminus N_p$. Suppose to the contrary that there exist distinct $y_0, y_0' \in Y$ such that $f_{y_0x_0}, f_{y_0'x_0} > 0$. Then $q_{y_0}, q_{y_0'} > 0$ because $q = f \circ p$. Let

$\tilde{Y} := \{\tilde{y}_0, \tilde{y}_0'\}$ be a two element set and define $Y \xrightarrow{\pi} \tilde{Y}$ by

$$Y \ni y \mapsto \pi(y) := \begin{cases} \tilde{y}_0 & \text{if } y = y_0 \\ \tilde{y}_0' & \text{otherwise} \end{cases}. \tag{7.9}$$

Let $\tilde{q} := \pi \circ q$ so that $\tilde{q}_{\tilde{y}_0}, \tilde{q}_{\tilde{y}_0'} > 0$. Since π is deterministic, a disintegration $\tilde{Y} \overset{\overline{\pi}}{\rightsquigarrow} Y$ of $(\pi, q, \tilde{q})$ exists by Theorem 7.1. By Proposition 6.11, $h := g \circ \overline{\pi}$ is a disintegration of $(\tilde{f} := \pi \circ f, p, \tilde{q})$. Hence, $0 = \delta_{\tilde{y}_0\tilde{y}_0'} = \sum_{x\in X} \tilde{f}_{\tilde{y}_0 x} h_{x\tilde{y}_0'}$. This implies $\tilde{f}_{\tilde{y}_0 x} h_{x\tilde{y}_0'} = 0$ for all $x \in X$. Since $\tilde{f}_{\tilde{y}_0 x_0} > 0$, this entails $h_{x_0\tilde{y}_0'} = 0$. A similar calculation swapping $\tilde{y}_0$ with $\tilde{y}_0'$ yields $h_{x_0\tilde{y}_0} = 0$. However, since h is a disintegration, $p = h \circ \tilde{q}$ so that $p_{x_0} = h_{x_0\tilde{y}_0}\tilde{q}_{\tilde{y}_0} + h_{x_0\tilde{y}_0'}\tilde{q}_{\tilde{y}_0'} = 0$, a contradiction since it was assumed that $p_{x_0} > 0$. ∎

This tells us that disintegrations are only possible for maps that are deterministic (almost everywhere). So although we can use disintegrations to reverse deterministic maps, we cannot use them to reverse stochastic maps in general. Therefore, one might ask if there is any reasonable way to reverse stochastic maps. For this, we have Bayes' theorem (cf. Theorem 2.1). Proving Bayes' theorem is entirely straightforward—a formula for a Bayesian inverse g of (f, p, q) is given by

$$g_{xy} := \begin{cases} p_x f_{yx}/q_y & \text{if } q_y > 0 \\ 1/|X| & \text{otherwise} \end{cases}. \tag{7.10}$$

We now describe how disintegrations are special kinds of Bayesian inverses.

Theorem 7.11. *Let $\{\bullet\} \overset{p}{\rightsquigarrow} X$ be a probability measure on X, let $X \overset{f}{\rightsquigarrow} Y$ be a stochastic map, and set $q := f \circ p$. If there exists a disintegration $Y \overset{g}{\rightsquigarrow} X$ of (f, p, q), then g is a Bayesian inverse of f.*

Proof. The goal is to prove Bayes' diagram commutes, i.e. $f_{yx}p_x = g_{xy}q_y$ for all $x \in X$ and $y \in Y$. If $p_x = 0$, then $0 = \sum_{y\in Y} g_{xy}q_y$ (since $p = g \circ q$) so that $g_{xy}q_y = 0$ for all $y \in Y$. Thus, the diagram commutes for all $x \in N_p$ and $y \in Y$. To see that it also commutes when $p_x > 0$, it suffices to assume f is deterministic by Theorem 7.8. In this case, we have

$$f_{yx}p_x \overset{\text{Thm 7.8}}{=\!=\!=} \delta_{yf(x)}p_x \overset{\text{Thm 7.1}}{=\!=\!=} g_{xy}q_y. \tag{7.12}$$

∎

In conclusion, we have learned the following facts regarding disintegrations and Bayesian inference in **FinStoch**.

Corollary 7.13. *Let X and Y be finite sets, let $\{\bullet\} \overset{p}{\rightsquigarrow} X$ be a probability measure on X, let $X \overset{f}{\rightsquigarrow} Y$ be a stochastic map, and set $q := f \circ p$.*

i. A Bayesian inference for (f, p, q) always exists and is q-a.e. unique.

ii. The map f is p-a.e. equivalent to a deterministic map (or equivalently p-a.e. deterministic by Proposition 5.41) if and only if a Bayesian inference of (f, p, q) is a disintegration of (f, p, q).

In summary, not every deterministic map is invertible, but every measure-preserving deterministic map has a disintegration in the enlarged category including stochastic maps. In this enlarged category, not every measure-preserving stochastic map has a disintegration, but every such map has a Bayesian inverse and Bayesian inverses reduce to disintegrations if and only if the original maps are a.e. deterministic. How much of this remains true in the quantum setting? What new insight does this perspective offer us in the quantum setting? What other categories admit such structure and properties? The rest of this paper is dedicated to answering the first question. The second question is the subject of a forthcoming paper [26] (partial answers in the classical setting are provided in Jacobs' recent work [15]). The last question has not yet been explored by the author.

8 Quantum Bayesian inference

In what follows, we will use the conventions and terminology of [27], much of which was reviewed in Example 3.12, Theorem 5.16, and elsewhere in this paper. Although disintegrations were defined more generally in Definition 6.1, we set the notation here.

Definition 8.1. Let $(\mathcal{A}, \omega)$ and $(\mathcal{B}, \xi)$ be finite-dimensional C^*-algebras equipped with states. Let $F : \mathcal{B} \rightsquigarrow \mathcal{A}$ be a CPU state-preserving map, i.e. $\omega \circ F = \xi$. A ***disintegration*** of ω over ξ consistent with F is a CPU map $G : \mathcal{A} \rightsquigarrow \mathcal{B}$ such that

$$
\begin{tikzcd}
\mathcal{A} \arrow[rr, rightsquigarrow, "G"] \arrow[dr, rightsquigarrow, "\omega"'] & {} \arrow[d, phantom, "="] & \mathcal{B} \arrow[dl, rightsquigarrow, "\xi"] \\
& \mathbb{C} &
\end{tikzcd}
\quad \text{and} \quad
\begin{tikzcd}
\mathcal{B} \arrow[rr, "\mathrm{id}_{\mathcal{B}}"] \arrow[dr, rightsquigarrow, "F"'] & {} \arrow[d, phantom, "\underset{\xi}{=}"] & \mathcal{B} \\
& \mathcal{A} \arrow[ur, rightsquigarrow, "G"'] &
\end{tikzcd}
, \tag{8.2}
$$

the latter diagram signifying commutativity ξ-a.e.

We now state and prove a theorem that says if a quantum operation has a state-preserving left-inverse, then the quantum operation is deterministic almost surely. This is a non-commutative generalization of Theorem 7.8.

Theorem 8.3. *Let $\mathcal{A}$ and $\mathcal{B}$ be finite-dimensional C^*-algebras, let $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ be a state on $\mathcal{A}$, let $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ be a CPU map, and set $\xi := \omega \circ F$. If there exists a disintegration of (F, ω, ξ), then F is ω-a.e. deterministic.*

The following proof will be broken up into a series of three Lemmas, which will be useful in their own right. Our proof of Theorem 8.3 is completely inspired by (and closely follows) the proof of Theorem 6.38 in Attal's notes [1].

Lemma 8.4. *Let $\mathcal{A}$ and $\mathcal{B}$ be finite-dimensional C^*-algebras, let $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ be a state on $\mathcal{A}$, let $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ be a CPU map, set $\xi := \omega \circ F$, and suppose there exists a CPU $\mathcal{A} \overset{G}{\rightsquigarrow} \mathcal{B}$ such that $G \circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{B}}$. Then*

$$
P_{\xi} B^* B P_{\xi} = P_{\xi} G\big(F(B)^* F(B)\big) P_{\xi} \qquad \forall\, B \in \mathcal{B}. \tag{8.5}
$$

Proof of Lemma 8.4. For any $B \in \mathcal{B}$, we have

$$G\big(F(B)\big)P_\xi = BP_\xi \tag{8.6}$$

by the condition $G \circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{B}}$ (cf. Theorem 5.16). Hence,

$$\begin{aligned} P_\xi B^*BP_\xi &= P_\xi G\big(F(B^*B)\big)P_\xi \quad \text{by (8.6)} \\ &\geqslant P_\xi G\big(F(B)^*F(B)\big)P_\xi \quad \text{by Kadison–Schwarz for } F \\ &\geqslant P_\xi G\big(F(B)\big)^* G\big(F(B)\big)P_\xi \quad \text{by Kadison–Schwarz for } G \\ &= \Big(G\big(F(B)\big)P_\xi\Big)^* G\big(F(B)\big)P_\xi \\ &= (BP_\xi)^*BP_\xi \quad \text{by (8.6)} \\ &= P_\xi B^*BP_\xi . \end{aligned} \tag{8.7}$$

Therefore, all intermediate equalities become equalities. ∎

Lemma 8.8. *Let $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ be a state and let $\mathcal{N}_\omega \subseteq \mathcal{A}$ be its associated null space (see Theorem 5.16 item (d) for terminology). If $A \geqslant 0$ and $\omega(A) = 0$, then $A \in \mathcal{N}_\omega$.*

Proof of Lemma 8.8. Write A as $A = D^*D$. Then $D \in \mathcal{N}_\omega$ by assumption. Since $\mathcal{N}_\omega$ is a left ideal in $\mathcal{A}$ (cf. [25, Construction 3.11]), $D^*D \in \mathcal{N}_\omega$. ∎

Lemma 8.9. *Let $\mathcal{A}$ and $\mathcal{B}$ be finite-dimensional C^*-algebras, let $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ be a state on $\mathcal{A}$, let $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ be a CPU map. Then*

$$F(B^*B)P_\omega = F(B)^*F(B)P_\omega \qquad \forall\, B \in \mathcal{B} \tag{8.10}$$

if and only if F is ω-a.e. deterministic, i.e.

$$F(B^*C)P_\omega = F(B)^*F(C)P_\omega \qquad \forall\, B, C \in \mathcal{B}. \tag{8.11}$$

Proof of Lemma 8.9. The only non-trivial direction is the forward one. Fix $B, C \in \mathcal{B}$. On the one hand, we have

$$\begin{aligned} F\Big((B+C)^*(B+C)\Big)P_\omega &= \Big(F(B^*B) + F(B^*C) + F(C^*B) + F(C^*C)\Big)P_\omega \\ &= \Big(F(B)^*F(B) + F(C)^*F(C)\Big)P_\omega + \Big(F(B^*C) + F(C^*B)\Big)P_\omega \end{aligned} \tag{8.12}$$

by (8.10). On the other hand, we have

$$\begin{aligned} F\Big((B+C)^*(B+C)\Big)P_\omega &= F(B+C)^*F(B+C)P_\omega \quad \text{by (8.10)} \\ &= \Big(F(B)^*F(B) + F(C)^*F(C)\Big)P_\omega + \Big(F(B)^*F(C) + F(C)^*F(B)\Big)P_\omega . \end{aligned} \tag{8.13}$$

Equating (8.12) with (8.13) gives

$$\Big(F(B^*C) + F(C^*B)\Big)P_\omega = \Big(F(B)^*F(C) + F(C)^*F(B)\Big)P_\omega . \tag{8.14}$$

Since this is valid for all B and C, replacing C with iC gives

$$i\Big(F(B^*C) - F(C^*B)\Big)P_\omega = i\Big(F(B)^*F(C) - F(C)^*F(B)\Big)P_\omega. \tag{8.15}$$

Dividing out by i and adding these two results gives

$$2F(B^*C)P_\omega = 2F(B)^*F(C)P_\omega. \tag{8.16}$$

Cancelling the 2 completes the proof. ■

Proof of Theorem 8.3. Let G be a disintegration of (F,ω,ξ) and fix $B \in \mathcal{B}$. By the Kadison–Schwarz inequality for F, we have

$$F(B^*B) - F(B)^*F(B) \geqslant 0. \tag{8.17}$$

Therefore,

$$\begin{aligned}
\omega\big(F(B^*B) - F(B)^*F(B)\big) &= \xi\Big(G\big(F(B^*B)\big) - G\big(F(B)^*F(B)\big)\Big) \quad \text{since } \omega = \xi \circ F \\
&= \xi\Big(P_\xi G\big(F(B^*B)\big)P_\xi - P_\xi G\big(F(B)^*F(B)\big)P_\xi\Big) \quad \text{by Lemma 5.14} \\
&= \xi(0) \quad \text{by Lemma 8.4 since } G \circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{B}} \\
&= 0.
\end{aligned} \tag{8.18}$$

By (8.17) and Lemma 8.8, $F(B^*B) - F(B)^*F(B) \in \mathcal{N}_\omega$. Since B was arbitrary, Lemma 8.9 implies F is ω-a.e. deterministic. ■

Remark 8.19. Our proof of Theorem 8.3 (and the three lemmas used) works for von Neumann algebras. Attal's Theorem 6.38 in [1] is similar in flavor. It states that if a CPU map $F : \mathcal{B}(\mathcal{H}) \rightsquigarrow \mathcal{B}(\mathcal{H})$ has a CPU left inverse G, then F is a *-homomorphism. In his theorem, $\mathcal{H}$ can be infinite-dimensional. If one allows a different codomain for F, then this claim is false. Indeed, a simple example, even in finite dimensions, is

$$\begin{aligned}
\mathcal{M}_n(\mathbb{C}) &\overset{F}{\rightsquigarrow} \mathcal{M}_m(\mathbb{C}) \\
A &\longmapsto \begin{bmatrix} A & 0 \\ 0 & \frac{1}{n}\mathrm{tr}(A)\mathbb{1}_{m-n} \end{bmatrix}
\end{aligned} \tag{8.20}$$

supposing $n < m$. A CPU left inverse of F is $G = \mathrm{Ad}_{[\mathbb{1}_n \; 0]}$, where the 0 is of size $(m-n) \times n$. It is therefore interesting that merely adding a state-preserving assumption to Attal's theorem guarantees the *-homomorphism claim almost surely regardless of the domain and codomain (and is even valid for von Neumann algebras).

Remark 8.21. Given the assumptions in Theorem 8.3, it is generally *not* true that F is ω-a.e. equivalent to a deterministic morphism. This is because there are disintegrations of CPU maps of the form $\mathcal{M}_n(\mathbb{C}) \overset{F}{\rightsquigarrow} \mathcal{M}_m(\mathbb{C})$, where m is not a multiple of n. Indeed, the example in Remark 8.19 provides such an instance if one equips $\mathcal{M}_n(\mathbb{C})$ with a density matrix of the form $\left[\begin{smallmatrix} \rho & 0 \\ 0 & 0 \end{smallmatrix}\right]$, where ρ is of size $n \times n$.

Proposition 8.22. *Let $\mathcal{A}$ and $\mathcal{B}$ be finite-dimensional C^*-algebras. A CPU map $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ is a *-isomorphism if and only if both of the following facts hold.*

i. For every state $\mathcal{B} \overset{\xi}{\rightsquigarrow} \mathbb{C}$, there exists a state $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ such that $\xi = \omega \circ F$.

ii. For every state $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$, a deterministic disintegration of $(F, \omega, \xi := \omega \circ F)$ exists.

A large part of the following proof uses [21, Theorem 5] and its proof.

Proof of Proposition 8.22. If F is a *-isomorphism, the two properties immediately follow (set $\omega := \xi \circ F^{-1}$ and $G := F^{-1}$). In the other direction, temporarily let ξ be any faithful state on $\mathcal{B}$ and let ω be any state on $\mathcal{A}$ such that $\xi = \omega \circ F$. Since a disintegration G exists, faithfulness of ξ guarantees $G \circ F = \mathrm{id}_{\mathcal{B}}$. Hence, F is injective. Now that injectivity of F has been established, let ω be *any* faithful state on $\mathcal{A}$ and set $\xi := \omega \circ F$. Suppose $B \in \mathcal{B}$ satisfies $\xi(B^*B) = 0$. Then $\xi(B^*B) = \omega\big(F(B^*B)\big) = 0$. Since $F(B^*B) \geqslant 0$ and since ω is faithful, $F(B^*B) = 0$ by Lemma 8.8. Since F is injective, $B^*B = 0$, which entails $B = 0$. Hence, ξ is also faithful. Let G be a deterministic disintegration of (F, ω, ξ). Then $G \circ F = \mathrm{id}_{\mathcal{B}}$ since ξ is faithful. Thus,

$$B^*B = G\big(F(B^*B)\big) \geqslant G\big(F(B)^*F(B)\big) = G\big(F(B)\big)^*G\big(F(B)\big) = B^*B \qquad \forall\, B \in \mathcal{B} \tag{8.23}$$

by the Kadison–Schwarz inequality for F. Hence, all terms in (8.23) are equal. Therefore,

$$\omega\big(F(B^*B) - F(B)^*F(B)\big) = \xi\Big(G\big(F(B^*B)\big) - G\big(F(B)^*F(B)\big)\Big) = \xi(0) = 0 \qquad \forall\, B \in \mathcal{B} \tag{8.24}$$

because $\omega = \xi \circ G$. Since $F(B^*B) - F(B)^*F(B) \geqslant 0$ by Kadison–Schwarz for F and since ω is faithful, $F(B^*B) = F(B)^*F(B)$ for all $B \in \mathcal{B}$. By the Multiplication Theorem (Lemma 4.8), F is a unital *-homomorphism. Note that G is also injective because if $G(A) = 0$ then $\omega(A^*A) = \xi\big(G(A^*A)\big) = \xi\big(G(A)^*G(A)\big) = 0$, which entails $A = 0$ by faithfulness of ω. Since both F and G are injective, finite-dimensionality of $\mathcal{A}$ and $\mathcal{B}$ imply they have the same dimension. Hence, $F^{-1} = G$ and so F is a *-isomorphism. ■

We now move to Bayesian inference in quantum mechanics. Although we have formulated the definition of a Bayesian inverse generally in Definition 6.1, we restate it here using the notation of C^*-algebras.

Definition 8.25. Let $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ be a CPU map, let $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ be a state, and set $\xi := \omega \circ F$. A ***Bayesian inverse*** of F is a CPU map $\mathcal{A} \overset{G}{\rightsquigarrow} \mathcal{B}$ such that

$$\begin{tikzcd}
\mathcal{B} \otimes \mathcal{B} \arrow[d, squiggly, "\mu_{\mathcal{B}}"'] & \mathcal{A} \otimes \mathcal{B} \arrow[l, squiggly, "G \otimes \mathrm{id}_{\mathcal{B}}"'] \arrow[r, squiggly, "\mathrm{id}_{\mathcal{A}} \otimes F"] \arrow[d, phantom, "="] & \mathcal{A} \otimes \mathcal{A} \arrow[d, squiggly, "\mu_{\mathcal{A}}"] \\
\mathcal{B} \arrow[r, squiggly, "\xi"'] & \mathbb{C} & \mathcal{A} \arrow[l, squiggly, "\omega"]
\end{tikzcd} . \tag{8.26}$$

In the following theorem, we show that all disintegrations are Bayesian inverses in the category **fdC*-AlgCPU**. Interestingly, the proof is general enough to avoid using an explicit formula for disintegrations.

Theorem 8.27. *Let $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ be a state on $\mathcal{A}$, let $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ be a CPU map, and set $\xi := \omega \circ F$. If there exists a disintegration $\mathcal{A} \overset{G}{\rightsquigarrow} \mathcal{B}$ of (F, ω, ξ), then G is a Bayesian inverse of F.*

The proof of this theorem will use the following lemma.

Lemma 8.28. *Let $\mathcal{B} \overset{\varphi}{\rightsquigarrow} \mathcal{A}$ be a CPU map between C^*-algebras and let $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ be a state. Suppose that $P_\omega \varphi(B^*B) P_\omega = P_\omega \varphi(B)^* \varphi(B) P_\omega$ for some $B \in \mathcal{B}$. Then*

$$P_\omega \varphi(B^*C) P_\omega = P_\omega \varphi(B)^* \varphi(C) P_\omega \quad \textit{and} \quad P_\omega \varphi(C^*B) P_\omega = P_\omega \varphi(C)^* \varphi(B) P_\omega \quad \forall\, C \in \mathcal{B}. \tag{8.29}$$

Note that this lemma is not a consequence of, nor does it imply, Lemma 8.9 due to how the supports are placed in the expressions. The proof of Lemma 8.28 below follows exactly the same steps as the proof of Theorem 4 in [21].

Proof of Lemma 8.28. Fix $C \in \mathcal{B}$ and $\lambda > 0$. Then

$$\begin{aligned} P_\omega \varphi\big((B+\lambda C)^*(B+\lambda C)\big) P_\omega \geqslant P_\omega \varphi(B^*)\varphi(B) P_\omega + \lambda P_\omega \big(\varphi(B^*)\varphi(C) + \varphi(C^*)\varphi(B)\big) P_\omega \\ + \lambda^2 P_\omega \varphi(C^*)\varphi(C) P_\omega \end{aligned} \tag{8.30}$$

by Kadison–Schwarz for φ. On the other hand,

$$\begin{aligned} P_\omega \varphi\big((B+\lambda C)^*(B+\lambda C)\big) P_\omega = P_\omega \varphi(B^*)\varphi(B) P_\omega + \lambda \big(\varphi(B^*C) + \varphi(C^*B)\big) P_\omega \\ + \lambda^2 P_\omega \varphi(C^*C) P_\omega \end{aligned} \tag{8.31}$$

by assumption. Combining the two results, dividing by λ, and taking the limit $\lambda \to 0$ gives

$$P_\omega \big(\varphi(B^*C) + \varphi(C^*B)\big) P_\omega \geqslant P_\omega \big(\varphi(B^*)\varphi(C) + \varphi(C^*)\varphi(C)\big) P_\omega. \tag{8.32}$$

Replacing B by iB and C by $-iC$ gives the reverse inequality. Hence,

$$P_\omega \big(\varphi(B^*C) + \varphi(C^*B)\big) P_\omega = P_\omega \big(\varphi(B^*)\varphi(C) + \varphi(C^*)\varphi(C)\big) P_\omega. \tag{8.33}$$

Now, replacing C by iC and adding/subtracting, the resulting terms entail (8.29). Since C was arbitrary, the lemma has been proved. ∎

Proof of Theorem 8.27. Since G is a disintegration of (F, ω, ξ), we have $G \circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{B}}$. By Lemma 8.4, more specifically (8.7),

$$P_\xi G\big(F(B)^*\big) G\big(F(B)\big) P_\xi = P_\xi G\big(F(B)^* F(B)\big) P_\xi \qquad \forall\, B \in \mathcal{B}. \tag{8.34}$$

Therefore, since $*$ is an involution,

$$P_\xi G\big(AF(B)\big) P_\xi \overset{\text{Lem } 8.28}{=\!=\!=} P_\xi G(A) G\big(F(B)\big) P_\xi \overset{G\circ F \underset{\xi}{=} \mathrm{id}_{\mathcal{B}}}{=\!=\!=\!=} P_\xi G(A) B P_\xi \qquad \forall\, A \in \mathcal{A},\ B \in \mathcal{B}. \tag{8.35}$$

This implies

$$\omega\big(AF(B)\big) \overset{\omega = \xi\circ G}{=\!=\!=} \xi\Big(G\big(AF(B)\big)\Big) = \xi\Big(P_\xi G\big(AF(B)\big) P_\xi\Big) \overset{(8.35)}{=\!=\!=} \xi\big(P_\xi G(A) B P_\xi\big) = \xi\big(G(A)B\big) \tag{8.36}$$

for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. This proves the theorem. ∎

The previous theorem has a converse, which we have already proved but we state it for the special case for finite-dimensional C^*-algebras.

Theorem 8.37. *Let* $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ *be a state on* $\mathcal{A}$*, let* $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ *be a* ω*-a.e. deterministic map, and set* $\xi := \omega \circ F$*. If* $\mathcal{A} \overset{G}{\rightsquigarrow} \mathcal{B}$ *is a Bayesian inverse of* F*, then* G *is a disintegration of* (F, ω, ξ)*.*

Proof. This is Theorem 6.27 and Remark 6.29 in the context of finite-dimensional C^*-algebras and CPU maps. ■

In conclusion, we have learned the following facts regarding disintegrations and Bayesian inference in **fdC*-AlgCPU**.

Corollary 8.38. *Let* $\mathcal{A}$ *and* $\mathcal{B}$ *be finite-dimensional* C^**-algebras, let* $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ *be a state on* $\mathcal{A}$*, let* $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ *be a CPU map, and set* $\xi := \omega \circ F$*.*

i. *If a Bayesian inference for* (F, ω, ξ) *exists, it is* ξ*-a.e. unique.*

ii. *Suppose a Bayesian inference* G *for* (F, ω, ξ) *exists. Then the map* F *is* ω*-a.e. deterministic if and only if* G *is a disintegration of* (F, ω, ξ)*.*

iii. *If* (F, ω, ξ) *has a disintegration* G*, then* F *is* ω*-a.e. deterministic and* G *is a Bayesian inference for* (F, ω, ξ)*.*

iv. *If* F *is* ω*-a.e. deterministic, then a Bayesian inference exists if and only if a disintegration exists.*

v. *A Bayesian inference for* (F, ω, ξ) *need not exist.*

Hence, not every deterministic map is invertible nor does every deterministic map equipped with a state have a disintegration (cf. [27, Section 5.2]). Nevertheless, there are more deterministic morphisms admitting disintegrations than inverses. In this enlarged category of C^*-algebras and (state-preserving) CPU maps, more morphisms have Bayesian inverses. Therefore, we have partially answered our first question addressed at the end of Section 7.

Acknowledgements

I gratefully acknowledge support from the Simons Center for Geometry and Physics, Stony Brook University and for the opportunity to participate in the workshop "Operator Algebras and Quantum Physics" in June 2019. Specifically, I thank Luca Giorgetti for encouraging me to pursue what happens to the notion of a disintegration when the morphism you are disintegrating is stochastic as opposed to deterministic. I thank Tobias Fritz for answering many of my questions on Markov categories and I thank Kenta Cho for informing me of Fritz' work on Markov categories. I also thank Kenta Cho, Chris Heunen, and Bart Jacobs for sharing their LaTeX code for the string diagrams used in this paper. Most importantly, I thank Benjamin Russo for many years of fruitful discussions. This project began while I was an Assistant Research Professor at the University of Connecticut.

References

[1] Stéphane Attal, *Lecture 6: Quantum channels*, 2014. ↑29, 31

[2] Howard Barnum and Emanuel Knill, *Reversing quantum dynamics with near-optimal quantum and classical fidelity*, J. Math. Phys. **43** (2002), no. 5, 2097–2106. ↑3

[3] Kenta Cho and Bart Jacobs, *Disintegration and Bayesian inversion via string diagrams*, Mathematical Structures in Computer Science (2019), 1–34, available at 1709.00322. ↑3, 4, 6, 9, 13, 15

[4] Florence Clerc, Vincent Danos, Fredrik Dahlqvist, and Ilias Garnier, *Pointless learning*, Foundations of software science and computation structures, 2017, pp. 355–369. ↑4

[5] Bob Coecke and Aleks Kissinger, *Picturing quantum processes: A first course in quantum theory and diagrammatic reasoning*, Cambridge University Press, 2017. ↑3

[6] Jared Culbertson and Kirk Sturtz, *A categorical foundation for Bayesian probability*, Appl. Categ. Structures **22** (2014), no. 4, 647–662. ↑2, 4

[7] Chandler Davis, *A Schwarz inequality for convex operator functions*, Proc. Amer. Math. Soc. **8** (1957), 42–44. ↑10

[8] Douglas R. Farenick, *Algebras of linear transformations*, Universitext, Springer-Verlag, New York, 2001. ↑7

[9] Brendan Fong, *Causal theories: A categorical perspective on Bayesian networks*, Master's Thesis, 2012. Available at 1301.6201 [math.PR]. ↑4

[10] Tobias Fritz, *A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics*, 2019. arXiv preprint: 1908.07021 [math.ST]. ↑2, 3, 4, 6, 7, 9, 10, 11, 12, 13, 16, 24

[11] Michèle Giry, *A categorical approach to probability theory*, Categorical aspects of topology and analysis (Ottawa, Ont., 1980), 1982, pp. 68–85. ↑2

[12] Mikhail Gromov, *Probability, symmetry, linearity*, Institut des Hautes Études Scientifiques, 2014. Lecture recording available at https://www.youtube.com/watch?v=aJAQVletzdY. ↑2, 26

[13] Chris Heunen and Jamie Vicary, *Categories for quantum theory: an introduction*, Oxford University Press, USA, 2019. ↑3

[14] Bart Jacobs, *From probability monads to commutative effectuses*, J. Log. Algebr. Methods Program. **94** (2018), 200–237. ↑2

[15] ———, *The mathematics of changing one's mind, via Jeffrey's or via Pearl's update rule*, J. Artificial Intelligence Res. **65** (2019), 783–806, available at 1807.05609. ↑29

[16] Richard V. Kadison, *A generalized Schwarz inequality and algebraic invariants for operator algebras*, Ann. of Math. (2) **56** (1952), 494–503. ↑10

[17] F. W. Lawvere, *The category of probabilistic mappings*, 1962. Preprint. ↑2

[18] Matthew S. Leifer, *Quantum dynamics as an analog of conditional probability*, Phys. Rev. A **74** (2006), 042310, available at 0606022. ↑20

[19] ———, *Conditional density operators and the subjectivity of quantum operations*, Foundations of probability and physics - 4, aip conference proceedings, 2007, pp. 172–186. ↑3

[20] Matthew S. Leifer and Robert W. Spekkens, *Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference*, Phys. Rev. A **88** (2013), 052130, available at 1107.5849. ↑3

[21] Hans Maassen, *Quantum probability and quantum information theory*, Quantum information, computation and cryptography, 2010, pp. 65–108. ↑8, 10, 11, 32, 33

[22] Saunders Mac Lane, *Natural associativity and commutativity*, Rice Univ. Studies **49** (1963), no. 4, 28–46. ↑7

[23] Ashwin Nayak and Pranab Sen, *Invertible quantum operators and perfect encryption of quantum states*, Quantum Inf. Comput. **7** (2007), no. 1-2, 103–110, available at 0605041. ↑3

[24] Arthur J. Parzygnat, *Discrete probabilistic and algebraic dynamics: a stochastic Gelfand-Naimark theorem*, 2017. arXiv preprint: 1708.00091 [math.FA]. ↑4, 10

[25] ______, *From observables and states to Hilbert space and back: A 2-categorical adjunction*, Appl. Categ. Structures **26** (2018), 1123–1157, available at 1609.08975. ↑30

[26] ______, *Jeffrey conditioning and Bayesian inference in quantum mechanics*, 2020. In preparation. ↑3, 20, 29

[27] Arthur J. Parzygnat and Benjamin P. Russo, *Non-commutative disintegrations: existence and uniqueness in finite dimensions*, 2019. arXiv preprint: 1907.09689 [quant-ph]. ↑2, 3, 4, 7, 15, 19, 21, 23, 26, 29, 34

[28] ______, *A non-commutative Bayes' theorem*, 2020. In preparation. ↑3, 4, 22, 24

[29] Dénes Petz, *Sufficiency of channels over von Neumann algebras*, Q. J. Math. **39** (1988), no. 1, 97–108. ↑3

[30] Shôichirô Sakai, C^**-algebras and* W^**-algebras*, Springer-Verlag, New York-Heidelberg, 1971. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 60. ↑15

[31] John H. Selby, Carlo Maria Scandolo, and Bob Coecke, *Reconstructing quantum theory from diagrammatic postulates*, 2018. arXiv preprint: 1802.00367 [quant-ph]. ↑9

[32] W. Forrest Stinespring, *Positive functions on* C^**-algebras*, Proc. Amer. Math. Soc. **6** (1955), no. 2, 211–216. ↑12

[33] Abraham Westerbaan, *Quantum programs as Kleisli maps*, Proceedings 13th International Conference on Quantum Physics and Logic, 2017, pp. 215–228. ↑2

A. Parzygnat, Institut des Hautes Études Scientifiques, 35 Routes de Chartres 91440, Bures-sur-Yvette, France

E-mail address, A. Parzygnat: parzygnat@ihes.fr