Combs, Causality and Contractions in Atomic Markov Categories

Dario Stein Márk Széles
Radboud University Nijmegen, The Netherlands
{dario.stein,mark.szeles}@ru.nl

We present a counterexample showing that Markov categories with conditionals (such as BorelStoch) need not validate a natural scheme of axioms which we call contraction identities. These identities hold in every traced monoidal category, so in particular this shows that BorelStoch cannot be embedded in any traced monoidal category. We remedy this under the additional assumption of atomicity: Atomic Markov categories validate all contraction identities, and furthermore admit a notion of trace defined for non-signalling morphisms. We conclude that atomic Markov categories admit an intrinsic calculus of combs without having to assume an embedding into compact-closed categories.

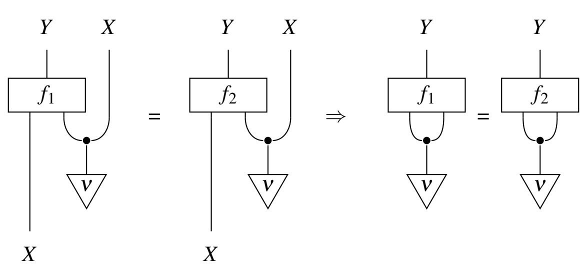
1 Introduction

Markov categories with conditionals have emerged as a general and powerful framework for studying stochastic processes and notions such as conditioning and independence in an abstract way that generalizes reasoning in graphical models (e.g. [4, 11, 5, 8, 9, 15, 21]). An important challenge is to understand the equational theory of such categories: assuming some equation or factorization holds, which further equations can be derived? This is crucial a reasoning tool, for example in causal inference: see [16] for a derivation of Pearl's front-door adjustment in such categorical terms.

Such problems are nontrivial because the existence of conditionals implies a range of non-obvious quasi-identities[1], such as the *positivity* and *causality axioms* [6] or that isomorphisms are deterministic:

$$f \circ g = \mathrm{id}_Y \wedge g \circ f = \mathrm{id}_X \quad \Rightarrow \quad \Delta_Y \circ f = (f \otimes f) \circ \Delta_X$$

In this work, we are interested in a particular schema of implications which we call *contraction identities*. A simple instance looks as follows:



(1)

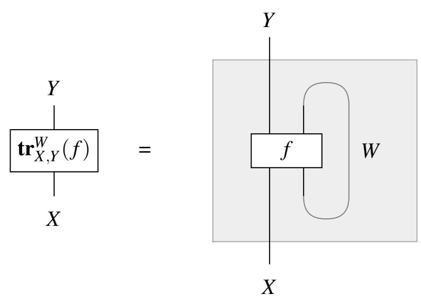
Every contraction identity is formed by connecting an outgoing wire (here X) to an incoming one and pulling the strings tight. This must be done in an acyclic way to make sure the resulting diagram is meaningful – we need to avoid genuine feedback loops. We will give the general, combinatorial definition of the contraction scheme in (25) using the language of free Markov categories [10] and hypergraphs.

[1] i.e. implications between equations. we will speak of *identities* for brevity

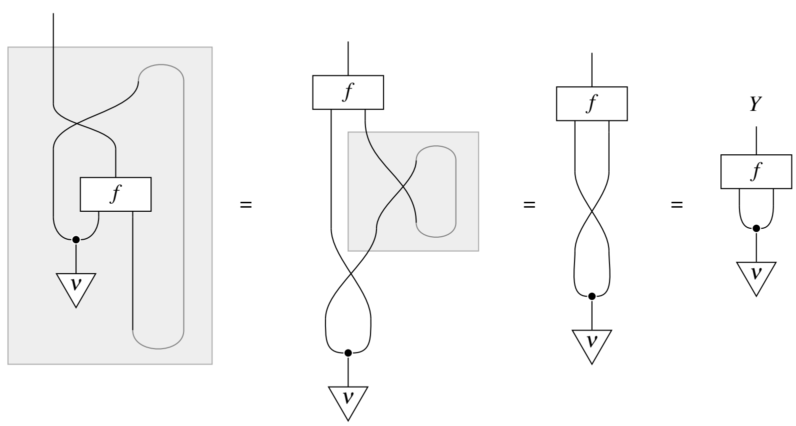
Submitted to:
ACT 2024

We show using a simple counterexample that (1) does not hold in the category **BorelStoch**, despite it having conditionals. We remedy this under a common additional assumption called *atomicity*. A Markov category is called atomic if $\Delta \circ p \ll p \otimes p$ holds for every morphism p. Common categories such as **FinStoch**, **Gauss** and **SetMulti** are atomic, while **Stoch**, **BorelStoch** and categories modelling fresh name generation are not. Our main theorem 27 states that atomic Markov categories with conditionals satisfy all contraction equations, which gives us additional power for equational reasoning in such categories. As we will discuss next, such categories admit a calculus of traces and combs.

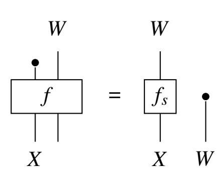
Causal Traces A *trace* on a monoidal category **C** is an operator which assigns to every morphism $f : X \otimes W \to Y \otimes W$ a morphism $\mathbf{tr}^W_{X,Y}(f) : X \to Y$ which represents evaluating a feedback loop of W into itself. We depict the trace operation as follows



The trace satisfies a number of axioms which encode the desired properties of such a feedback operation (see Appendix 7.2), and the graphical calculus reflects these [20]. We show that all contraction identities can be proved from the trace axioms (Proposition 26). For example in (1), we can obtain the the consequent equation via the trace



As a consequence of our counterexample, **BorelStoch** cannot be embedded in *any* traced monoidal category. We then establish the following converse result, in trying to obtain a canonical notion of trace on an atomic Markov category. It is hopeless to attempt this for an arbitrary morphism $f : X \otimes W \to Y \otimes W$. If however f satisfies the non-signalling condition, i.e. there exists a morphism f_s such that



this means the input W is not needed to compute the output W. So the feedback loop created by the trace remains causal. Our central definition 4 states that every atomic Markov category with conditionals admits a canonical notion of causal trace, that is a trace defined on non-signalling morphisms. We show that this construction satisfies a restricted version of the trace axioms, and compare with the notion of a partially traced category [12, 1]. The properties of the causal trace suffice to derive all contraction identities (Propositions 26 and 27).

Combs Combs are a widely-used tool for studying decompositions of string diagrams into more flexible shapes, with applications for example in quantum theory [2] and causal inference [16]. A comb C of type $(A,A') \to (B,B')$ is roughly a diagram of type $A \to A'$ which features a *hole* of type $B \to B'$. This hole can be filled by appropriate morphisms h, leading to a composite $C[h]$. This way, a comb describes a second-order process and lends itself to be rendered in an evocative shape (see Figure 1).

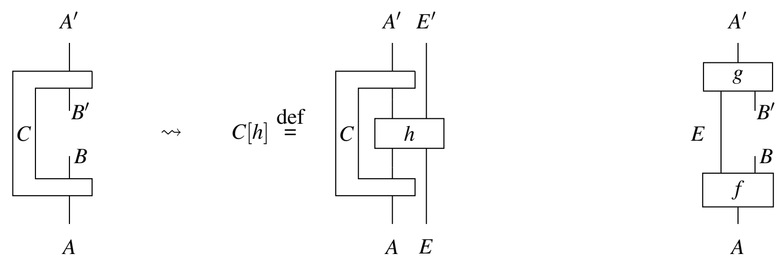
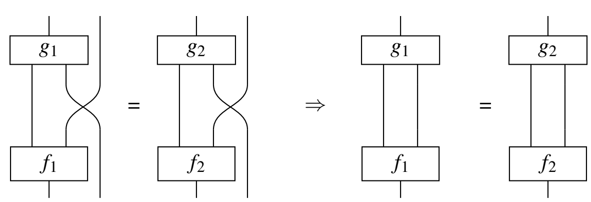


Figure 1: Left: Comb insertion (graphically), right: intensional presentation of a comb as a pair $\langle f|g\rangle$

There are various inequivalent ways of formally defining what a comb is. An extensional definition is to define a comb as a morphism $A \otimes B' \to A' \otimes B$ in $\mathbf{C}$ which is non-signalling from B' to B. An intensional definition is as a pair of morphisms $\langle f|g\rangle$ with $f : A \to E \otimes B$ and $g : E \otimes B' \to A'$, under some kind of equivalence relation. We review these notions in Section 5 following [13].

Relating the different definitions of combs to each other is tricky: To go from the extensional definition to comb insertion is an instance of a contraction identity:



(2)

For this reason, the theory of combs has commonly been developed in the setting of compact closed categories. There, second-order processes can be reduced to first-order processes by bending wires, and the various notions of comb are equivalent. The downside is that even when analyzing diagrams in a Markov category $\mathbf{C}$ using combs, we must assume that it comes with an embedding into a compact closed category (for example $\mathbf{FinStoch} \hookrightarrow \mathbf{Mat}(\mathbb{R}^+)$). Not only is this difficult for practical categories (such an embedding category requires developing a theory of exact conditioning, e.g. [17, 23]), but our counterexample shows that this is generally impossible in the absence of the axiomicity axiom: Identity (2) is invalid in $\mathbf{BorelStoch}$. We argue that causal traces are sufficient to develop an intrinsic theory of combs. We prove that in every atomic Markov category with conditionals, the extensional and intensional definitions of combs are equivalent.

Contributions This work is a shared refinement of three different developments [14, 13, 7]. The major starting point is [14], where the author develops a theory of causality and contraction. The statements of their Lemma 4.2.5 and Theorem 4.2.6 have the same content as our Lemma 10 and Corollary 27, but are proved under a different set of assumptions. Their work is built on the notion of *universal dilations* in semicartesian categories. We specialize this to Markov categories, which allows us to refine the monolithic notion of universal dilations into more modular pieces:

1. We recognize the role of the atomicity axiom, which has been introduced in [7]. Our theorems are proven under the assumption of conditionals and atomicity, which are weaker than universal dilations
2. our counterexample (5) gives a precise reason for the failure of universality in **BorelStoch**.
3. the combinatorics of the contraction identities are elegantly phrased using free Markov categories
4. Traces have been (purposefully) ignored in [14]. We make the notion of causal trace a central element of our theory, which offers new proofs, and enables connections to the theory of combs and compact closed categories

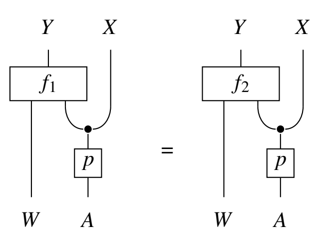
We then connect our theory to the different notions of combs compared in [13]. We generalize their results significantly from the compact-closed case to a large family of Markov categories, which removes the need to assume an embedding as done in [16]. We briefly return to the relationship between universal dilations and combs in Proposition 33.

Acknowledgements We'd like to thank many people for fruitful discussions, in particular Dylan Braithwaite, Tobias Fritz, Tomáš Gonda, Nicholas Gauguin Houghton-Larsen, and Nathaniel Virgo.

2 Atomic Markov Categories

We being by recalling the notion of atomicity in Markov categories, as well as the prerequisite notions of almost sure equality and absolute continuity found in [7]. We assume that the reader is familiar with monoidal Markov categories, but recall relevant definitions and notations in Appendix 7.1. Let **C** be a Markov category.

Definition 1 (Almost sure equality – [7])**.** *Given a morphism $p : A \to X$, we say $f_1, f_2 : W \otimes X \to Y$ are* p-almost surely equal *(written $f_1 =_p f_2$) if*



Definition 2 (Absolute continuity – [7])**.** *For two morphisms $p : A \to X$ and $q : B \to X$, we say that p is* absolutely continuous with respect to *q (written $p \ll q$), if $f_1 =_q f_2$ implies $f_1 =_p f_2$ for all $f_1, f_2 : W \otimes X \to Y$.*

We note the quantification over arbitrary W which is stronger than earlier definitions (e.g. [5]). These abstract notions capture the usual definitions of almost sure equality in our example categories (see [7, Section 2.2]):

1. in **FinStoch**, f_1, f_2 are p-almost surely equal if $f_1(y|x) = f_2(y|x)$ for all x with $p(x|a) > 0$ for some $a \in A$. We have $p \ll q$ if $\mathrm{supp}(p) \subseteq \mathrm{supp}(q)$ where
$$\mathrm{supp}(p) = \{x \in X : \exists a \in A, p(x|a) > 0\}$$
2. in **BorelStoch**, f_1, f_2 are p-almost surely equal if the set $D = \{x : f_1(x) \neq f_2(x)\}$ satisfies $p(D|a) = 0$ for all $a \in A$. We have $p \ll q$ if for every every measurable subset $S \subseteq X$, we have
$$(\forall b \in B, q(S|b) = 0) \Rightarrow \forall a \in A, p(S|a) = 0$$

Definition 3 (Atomicity – [7])**.** *We call a morphism $p : A \to X$* atomic *if $\Delta_X \circ p \ll p \otimes p$. We call the Markov category* **C** *atomic if every morphism in it is atomic.*

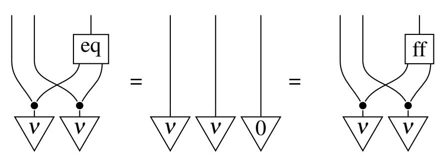
The following characterizations are known from [7], but we repeat them because they are instructive.

Example 4. *For every morphism $p : A \to X$ in* **FinStoch***, we have*
$$(p \otimes p)(x_1, x_2|a_1, a_2) = p(x_1|a_1)p(x_2|a_2), \qquad (\Delta_X \circ p)(x_1, x_2|a) = \begin{cases} p(x_1|a), & x_1 = x_2 \\ 0, & \textit{otherwise} \end{cases}$$
so $\mathrm{supp}(\Delta_X \circ p) \subseteq \mathrm{supp}(p \otimes p)$*. That is* **FinStoch** *is atomic.*

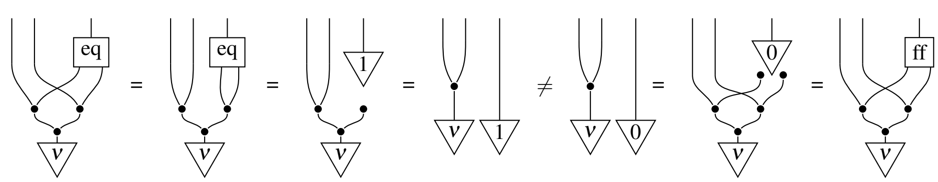
The name *atomicity* points to the fact that this property fails for distributions that are atomless, such as the Lebesgue measure. For a morphism $p : A \to X$ in **BorelStoch**, we define its set of atoms $\mathscr{A} \subseteq X$ as
$$\mathscr{A} = \{x \in X : \exists a \in A, p(\{x\}|a) > 0\}$$
We call p *completely atomic* if $p(\mathscr{A}|a) = 1$ for all $a \in A$, i.e. its probability mass is fully concentrated on its atoms. It is shown in [7, Theorem 3.2.7] that p is atomic in the sense of Definition 3 if and only if it is completely atomic. It is easy to see that the Lebesgue measure ν on the interval $[0, 1]$ is *not* atomic, as its set of atoms $\mathscr{A}$ is empty. We will now give a concrete counterexample, showing that $\Delta \circ \nu \not\ll \nu \otimes \nu$.

Example 5. **BorelStoch** *is not atomic.*

Proof. Let $X = [0, 1]$ and $\nu : I \to X$ be the Lebesgue measure, and consider the measurable functions $\mathrm{ff}, \mathrm{eq} : X \otimes X \to \{0, 1\}$ which are defined as follows
$$\mathrm{ff}(x, y) = 0, \qquad \mathrm{eq}(x, y) = \begin{cases} 1, & x = y \\ 0, & \text{otherwise} \end{cases}$$
Then we have $\mathrm{ff} =_{\nu \otimes \nu} \mathrm{eq}$, because intuitively, if $X, Y \sim \nu$ then $\Pr(\mathrm{eq}(X, Y) = 0) = 1$. Formally,



However $\mathrm{ff} \neq_{\Delta \circ \nu} \mathrm{eq}$, because $\mathrm{ff}(X, X)$ is constantly 0, while $\mathrm{eq}(X, X)$ is constantly 1.

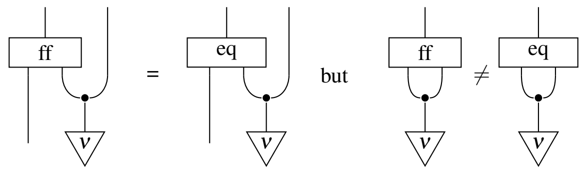


This shows $\Delta \circ \nu \not\ll \nu \otimes \nu$. □

As a simple corollary, we have

Example 6. **BorelStoch** *does not satisfy the contraction identity* (1) *from the introduction.*

Proof.



□

This type of counterexample can easily be adapted to other Markov categories which feature atomless distributions, such as quasi-Borel spaces or nominal sets ([22]). As we will see, such counterexamples are impossible when the category in question is atomic.

We remark that the notion of atomicity is dependent on the surrounding category. The Gaussian distribution $\mathcal{N}(0,1)$ is non-atomic in **BorelStoch**, but it is atomic in the subcategory **Gauss**. More generally, the category **TychStoch** of Tychonoff spaces and continuous Markov kernels is atomic [7], despite featuring atomless measures: Note that the map eq from the counterexample 6 fails to be continuous.

We return to the abstract properties of atomic morphisms

Proposition 7. *In a Markov category* **C***, the following morphisms p are always atomic*

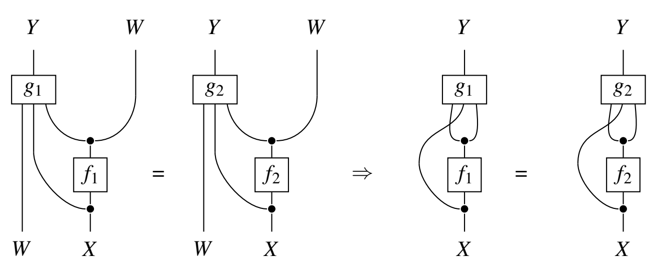
1. *morphisms of 'full support', in the sense that $f_1 =_p f_2$ implies $f_1 = f_2$.*
2. *deterministic morphisms*
3. *if* **C** *is satisfies the causality axiom, g is deterministic, and p is atomic, then the composite $g \circ p$ is atomic*

Note that atomic morphisms are generally not closed under composition, and atomicity cannot be checked on points (see Example 36). We now show that the atomicity axiom plus conditionals imply a basic contraction identity. These will be sufficient to prove all of them (3.1). We need the following technical notion to rule out pathological cases:

Definition 8. *An object W of a Markov category is called* cancellable *if* $\mathrm{del}_W \otimes f_1 = \mathrm{del}_W \otimes f_2$ *implies* $f_1 = f_2$*. We call the Markov category* cancellative *if every object W is cancellable.*

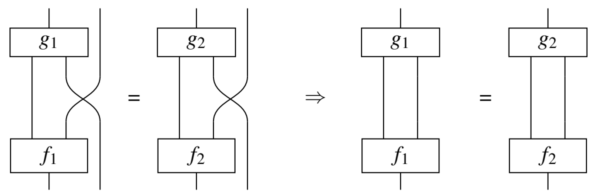
This condition is called *normality* in [14]. In practical examples, most objects are cancellable, for example any object that admits a state $I \to W$. In **FinStoch** and **BorelStoch**, every object except $W = \emptyset$ is cancellable. We may formally consider sub-Markov categories **FinStoch***, **BorelStoch*** on non-empty objects, or simply assume cancellability on the fly when needed. We will go with the latter approach.

Lemma 9. *In an atomic Markov category, we have for all $f_1, f_2 : X \to W$ and $g_1, g_2 : W \otimes X \otimes W \to Y$ that*



Proof. In the appendix. □

Lemma 10. *Let* **C** *be an atomic Markov category with conditionals. Then*

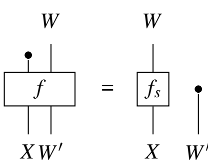


Proof. In the appendix. □

Note that this is precisely the statement of Lemma 4.2.5 of [14] in our setting.

3 Causal Traces

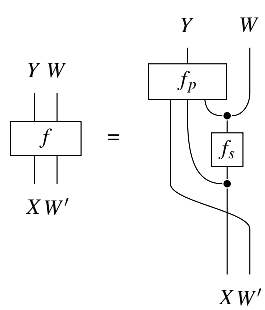
Definition 11. *We say a morphism $f : X \otimes W' \to Y \otimes W$ is* non-signalling *(from W' to W) if there exists a morphism $f_s : X \to W$ such that*



This captures the intuition that in order to determine W, we don't need access to W' (but we do to compute the joint output Y).

For non-signalling morphisms, we can hope to define a canonical trace as follows:

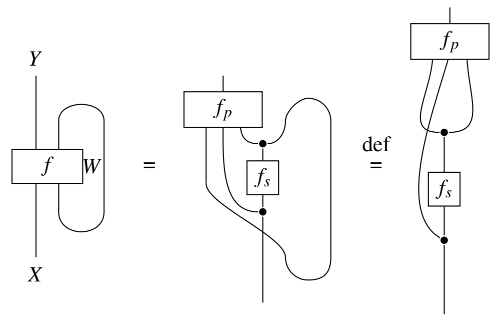
Proposition 12. *Let* **C** *be a Markov category with conditionals. Then $f : X \otimes W' \to Y \otimes W$ is non-signalling if and only if it can be written in the following form*



(3)

In this article, we will call the form (3) *a* disintegration *of f. Note that while f_s is unique (assuming cancellability), f_p is not.*

Definition 13. *Let* **C** *be an atomic Markov category with conditionals, and $f : X \otimes W \to Y \otimes W$ be* *non-signalling. Then we define the causal trace* $\mathbf{tr}^W(f): X \to Y$ *in terms of any disintegration as*



(4)

By Lemma 10, this definition does not depend on the choice of disintegration.

The definition of the causal trace is unique in that it is forced via the trace axiom. For a non-signalling morphism f, *any* trace must send f to (4). The special role of non-signalling is also reflected in the following proposition

Proposition 14. *Let* $\mathbf{C}$ *be a traced semicartesian monoidal category. If* $f: X \otimes W \to Y \otimes W$ *is a discardable morphism, then* $\mathbf{tr}^W(f): X \to Y$ *need not be discardable, unless* f *is non-signalling.*

Proof. In the appendix. □

We spell out concretely how to compute the trace in **FinStoch**:

Example 15. *Let* $f: X \otimes W \to Y \otimes W$ *be a non-signalling morphism in* **FinStoch**, *and assume* $W \neq \emptyset$. *As we can compute the causal trace using any embedding into a traced category, such as the usual embedding into* $\mathbf{Mat}(\mathbb{R})$. *There we have*

$$\mathbf{tr}^W(f)(y|x) = \sum_w f(y,w|x,w)$$

This will not define a normalized probability kernel for a general f. *However, if* f *is non-signalling then the sum* $\sum_y f(y,w|x,w')$ *does not depend on* w'. *Hence by fixing any* $w_0 \in W$, *we can prove normalization*

$$\sum_y \mathbf{tr}^W f(y|x) = \sum_y \sum_w f(y,w|x,w) = \sum_w \sum_y f(y,w|x,w) = \sum_w \sum_y f(y,w|x,w_0) = 1$$

from the normalization of f.

In the language of [24], we can see a non-signalling morphism as a degenerate kind of automaton whose future states don't depend on the current states. The causal trace takes a fixed point by re-using identical copies of states.

3.1 Laws of the Causal Trace

We now state and prove a list of properties which the causal trace enjoys. We will compare it with the notion of *partial trace* discussed by [12, 1]. This requires us to postulate a class of traceable morphism, which we choose as the non-signalling ones, $\mathbb{T}_W^{X,Y} = \{f: X \otimes W \to Y \otimes W \mid f \text{ is non-signalling}\}$. The axioms marked with a star* are weaker than the corresponding axioms for partial traces, while the ones without star are identical to the partial trace axioms. All proofs are given in the Appendix (Section 7.4).

Proposition 16 (Tightening – naturality in X,Y)**.** *If $f : X \otimes W \to Y \otimes W$ is traceable, then $(h \otimes \mathrm{id}_W) \circ f \circ (g \otimes \mathrm{id}_W) : X' \otimes W \to Y' \otimes W$ is traceable and*

$$\mathbf{tr}^W((h \otimes \mathrm{id}_W) \circ f \circ (g \otimes \mathrm{id}_W)) = h \circ \mathbf{tr}^W(f) \circ g \tag{5}$$

Proposition 17 (Sliding* – dinaturality in W)**.** *Let $f : X \otimes W' \to Y \otimes W$ and $g : W \to W'$. If f is traceable, then so are $(\mathrm{id}_Y \otimes g) \circ f$ and $f \circ (\mathrm{id}_X \otimes g)$, and*

$$\mathbf{tr}^U((\mathrm{id}_Y \otimes g) \circ f) = \mathbf{tr}^V(f \circ (\mathrm{id}_X \otimes g)) \tag{6}$$

Proposition 18 (Vanishing (Coherence with I))**.** *All $g : X \otimes I \to Y \otimes I$ are traceable, and $\mathbf{tr}^I(g) = g$.*

Proposition 19 (Associativity* (Coherence with $\otimes$))**.** *If $f : X \otimes (U \otimes V) \to Y \otimes (U \otimes V)$ is traceable, then so are $f : (X \otimes U) \otimes V \to (Y \otimes U) \otimes V$ and $\mathbf{tr}^V(f) : X \otimes U \to Y \otimes U$, and*

$$\mathbf{tr}^{U \otimes V}(f) = \mathbf{tr}^U(\mathbf{tr}^V(f))$$

Proposition 20 (Superposition (Strength))**.** *If $f : X \otimes W \to Y \otimes W$ is traceable, so is $g \otimes f : X' \otimes X \otimes W \to Y' \otimes Y \otimes W$, and*

$$\mathbf{tr}^W(g \otimes f) = g \otimes \mathbf{tr}^W(f) \tag{7}$$

Proposition 21 (Yanking)**.** *The symmetry is traceable, and*

$$\mathbf{tr}^W(\mathrm{swap}_{W,W}) = \mathrm{id}_W \tag{8}$$

4 Causal Traces in Free Markov Categories

In this section, we show that free Markov categories also have causal traces, and that interpretations of Markov string diagrams in cancellative, atomic Markov categories $\mathbf{C}$ with conditionals preserve causal traces. This is an analogue of Theorem 4.2.13 of [14], albeit in a different setting (Markov categories with extra structure instead of universal theories). In particular, we obtain that all contraction identities are satisfied. That is, in the language of [14], the causal trace of a non-signalling morphism in $\mathbf{C}$ can be computed via any stencil representation.

We use free Markov categories as constructed in [10]. Recall that every monoidal signature Σ gives rise to a finite hypergraph which we also denote by Σ. A labelling of wires and boxes in a hypergraph G is a hypergraph homomorphism to Σ. Finite hypergraphs and their homomorphisms form a category **FinHyp**. A *Markov string diagram* is a cospan $\underline{m} \xrightarrow{i} G \xleftarrow{j} \underline{n}$ with discrete hypergraphs $\underline{m}, \underline{n}$, satisfying the conditions of acyclicity, left monogamy, and having no eliminable boxes. The setting is recalled in more detail in Appendix 7.5.

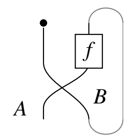
Proposition 22 ([10])**.** *The free Markov category $\mathbf{FreeMarkov}_\Sigma$ over a monoidal signature Σ can be constructed as follows:*

1. *Objects are hypergraph homomorphisms $\underline{m} \to \Sigma$, i.e. lists of types in Σ.*
2. *Morphisms are (isomorphism classes of) Markov string diagrams, which compose by pushout and subsequent normalisation (elimination of eliminable boxes).*
3. *The tensor is given by coproduct.*

To ease notation we will write $f : n \to m$ for a morphism $f : (\underline{m} \to \Sigma) \to (\underline{n} \to \Sigma)$, thus leaving the labelling implicit. It is straightforward to verify the following:

Proposition 23. *A morphism $f : n \otimes w \to m \otimes w$ in* $\mathbf{FreeMarkov}_\Sigma$ *is non-signalling if and only if there are no directed paths from input ports in w to output ports in w.*

We now show that $\mathbf{FreeMarkov}_\Sigma$ has causal traces, by describing the appropriate combinatorial contraction as a contraction of hypergraphs: Let $f : m \otimes w \to n \otimes w$ be non-signalling, represented by the cospan $\underline{m} + \underline{w} \xrightarrow{[i,i']} G \xleftarrow{[j,j']} \underline{n} + \underline{w}$. Graphically, forming the contraction amounts to gluing the wires connected to matching input and output ports in w, thus making them inner wires (not connected to any input or output port), and normalizing. See Figure 2 for an illustration. The need for normalisation is apparent from the following example:



More formally, we define the contracted string diagram $\mathbf{contr}^w(f) : m \to n$ as the normalization of the resulting cospan in the following diagram, where the central square is a pushout:z

$$\begin{tikzcd}
 & & \mathbf{contr}^w(G) & G \arrow[l, dashed] & \underline{n} \arrow[l, "j"'] \\
 & G \arrow[ur, dashed] & & \underline{w} \arrow[ul, dashed] & \\
\underline{m} \arrow[ur, "i"] & & \underline{w}+\underline{w} \arrow[ul, "{[i',j']}"'] \arrow[ur, "{[\mathrm{id},\mathrm{id}]}"] & &
\end{tikzcd}$$

We verify that the resulting string diagram is acyclic and left monogamous. Acyclicity follows from the non-signalling assumption. For left monogamy, observe that after gluing the only affected wires are the one connected to ports in w. Every wire connected to an output port is either also connected to an input port, or is an output of a box. In both cases, left monogamy is preserved, for if a wire was both connected to an input and an output port, the input port cannot be in w by the acyclicity assumption.

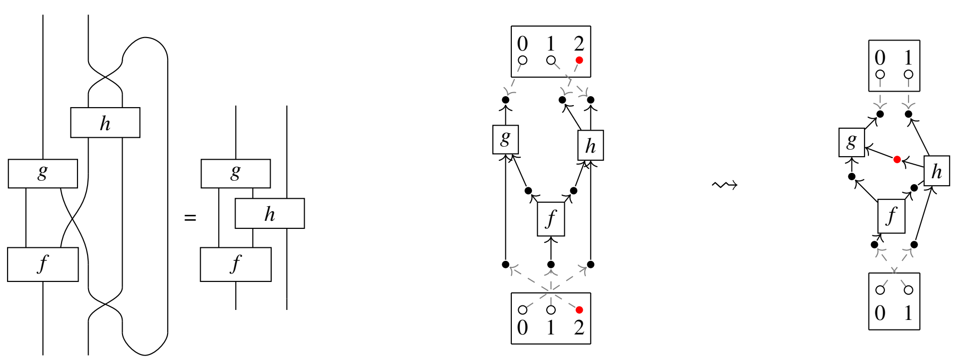


Figure 2: An example of computing the contraction in the hypergraph representation. (Left) The causal trace we want to compute (Right) Hypergraph representations. Black dots represent wires, white dots ports; the red ports are being contracted.

Proposition 24. *The contraction operation* $\mathbf{contr}$ *satisfies the causal trace axioms (as stated in Propositions 16 - 21) for* $\mathbf{FreeMarkov}_\Sigma$.

Proof. Vanishing, strength, and yanking are immediate. The other axioms require bit more caution around the normalization step that occurs in sequential composition and contraction. We can use the fact that normalization of string diagrams is an identity-on-objects gs-monoidal functor $\mathbf{FreeGS}_\Sigma \to \mathbf{FreeMarkov}_\Sigma$ from the free gs-monoidal category **FreeGS** ([10], Lemma 6.5) to solve these cases. $\square$

We finally are ready to formally define our notion of contraction identities.

Definition 25. *A Markov category* $\mathbf{C}$ *is said to satisfy all contraction identities if for all* $C_1, C_2 : m \otimes w \to n \otimes w$ *non-signalling in a free Markov category over any signature* Σ*, and all interpreting functors* $[\![-]\!] : \mathbf{FreeMarkov}_\Sigma \to \mathbf{C}$*, we have that if* $[\![C_1]\!] = [\![C_2]\!]$*, then* $[\![\mathbf{contr}^w(C_1)]\!] = [\![\mathbf{contr}^w(C_2)]\!]$*.*

This definition is analogous to the 'notions of contraction' in [14].We can show by induction that if $\mathbf{C}$ already has causal traces, then any Markov functor from $\mathbf{FreeMarkov}_\Sigma$ must preserve them:

Proposition 26. *Let* $\mathbf{C}$ *have causal traces, and let* $f : m \otimes w \to n \otimes w$ *in* $\mathbf{FreeMarkov}_\Sigma$*,* $[\![-]\!] : \mathbf{FreeMarkov}_\Sigma \to \mathbf{C}$ *an interpreting functor. Then if* f *is non-signalling, so is* $[\![f]\!]$*, and* $[\![\mathbf{contr}^w(f)]\!] = \mathbf{tr}^{[\![w]\!]}([\![f]\!])$*.*

Proof. The first point is clear. For the second point we proceed by strong induction on w. If $w = 0$, then f is of the form $g \otimes \mathrm{id}_I$. We are done by vanishing (18).

Now assume the statement for all $c < k+1$. Let $f : n+k+1 \to m+k+1$.

$$\mathbf{tr}^{[\![k+1]\!]}([\![f]\!]) = \mathbf{tr}^{[\![k]\!]}(\mathbf{tr}^{[\![1]\!]}([\![f]\!])) = \mathbf{tr}^{[\![k]\!]}([\![\mathbf{contr}^1(f)]\!] = [\![\mathbf{contr}^k(\mathbf{contr}^1(f))]\!] = [\![\mathbf{contr}^{k+1}(f)]\!]$$

We applied the induction hypothesis twice, and associativity (19) twice. $\square$

As a direct consequence of the previous proposition we obtain:

Corollary 27. *Every cancellable atomic Markov categories with conditionals satisfies all contraction identities.*

5 Calculus of Combs

We recall various definitions of combs and refer to [13, 18, 19] for reference.

Definition 28 (Comb)**.** *Let* $\mathbf{C}$ *be a symmetric monoidal category. A comb of type* $(A, A') \to (B, B')$ *is a triple* $C = \langle f | g \rangle_E$ *consisting of an object* E *and morphisms* $f : A \to E \otimes B$ *and* $g : E \otimes B' \to A'$*. We will omit the subscript* E *if it is clear from context. For a morphism* $h : B \otimes K \to B' \otimes K'$*, the comb insertion* $C[h] : A \otimes K \to A' \otimes K'$ *is defined as* $C[h] = (f \otimes \mathrm{id}_K);(\mathrm{id}_E \otimes h);(g \otimes \mathrm{id}_{K'})$*. The* extension *of the comb* C *is the joint morphism* $C[\mathrm{swap}_{B,B'}] : A \otimes B' \to A' \otimes B$*.*

Definition 29. *Two combs* $C_1, C_2 : (A, A') \to (B, B')$ *are called*

1. extensionally equivalent *if their extensions are equal*
2. contextually equivalent *if for all* $h : B \otimes K \to B' \otimes K'$*, we have* $C_1[h] = C_2[h]$
3. optically equivalent *if they are identified as elements of the coend*

$$\int^E \mathbf{C}(A, E \otimes B) \times \mathbf{C}(E \otimes B', A')$$

Concretely, this is the equivalence generated by 'sliding' for all $s : E \to E'$

$$\langle f;(s \otimes \mathrm{id}_B) | g \rangle_{E'} \sim \langle f | (s \otimes \mathrm{id}_{B'}); g \rangle_E$$

Proposition 30 ([13])**.** *Optic equivalence implies contextual equivalence. Contextual equivalence implies extensional equivalence.*

The converses are false in general. Comb insertion need not be well-defined under extensional equivalence; counterexample atomicity. It is shown in [13] that all three notions coincide in two special cases: compact closed categories and cartesian categories. We generalize this as follows:

Proposition 31. *In any category with causal traces, extensional and contextual equivalence coincide.*

Proof. We can compute the comb insertion from the extension using the causal trace from Figure 2. □

Theorem 32. *Let* **C** *be an atomic Markov category with conditionals. Then extensional and contextual equivalence coincide. Computing the extension defines a bijection between*

1. *morphisms* $f : A \otimes B' \to A' \otimes B$ *that are non-signalling*
2. *contextual equivalence classes of combs* $(A, A') \to (B, B')$

Proof. In the appendix. □

This theorem generalizes the examples covered by [13] to include many common Markov categories without assuming an embedding in a compact-closed category. The case of optic equivalence requires further structure.

Proposition 33. *If* **C** *has universal dilations in the sense of [14], then extensional and optic equivalence coincide.*

Proof. In the appendix. □

Because **FinStoch** has universal dilations ([14, Theorem 2.4.6]), this means that for **FinStoch**, all notions of combs coincide.

6 Conclusions and Future Work

We have shown that the contraction identities hold in every cancellative atomic Markov category with conditionals, and used this to develop a theory of causal traces and relate various notions of comb equivalence. Our work leaves an array of interesting open questions, including about converses of our results:

When does a Markov category embed into a traced category? It is known that every partially traced category embeds in a traced one [1], but our axioms for the causal trace are strictly weaker than the partial trace axioms. We conjecture that the difference is similar to what Houghton-Larsen achieves with his construction of *causal channels* in [14, Section 4.1].

To what extent is atomicity a necessary assumption? Does an embedding in a traced category imply atomicity? Atomicity is intimately related to supports [7], and in turn to universal dilations, though the precise relationship remains to be clarified.

Our developments in Section 4 suggest that the information-flow properties of free Markov categories are an interesting area of study beyond this work: We conjecture that these categories are atomic, and have supports but no conditionals. We also believe that free Markov categories embed in free hypergraph categories, even though this point was sidestepped in [10].

We would also like to explore if failure of atomicity poses formal challenges in causal inference, given that the approach to causal inference in [16] relied on combs, or if the approach can be refined to not rely on them.

References

[1] Marc Bagnol (2015): *Representation of Partial Traces*. Electronic Notes in Theoretical Computer Science 319, pp. 37–49.

[2] Giulio Chiribella, G Mauro D'Ariano & Paolo Perinotti (2008): *Quantum circuit architecture*. Physical review letters 101(6), p. 060401.

[3] Kenta Cho & B. Jacobs (2019): *Disintegration and Bayesian inversion via string diagrams*. Mathematical Structures in Computer Science 29, pp. 938 – 971.

[4] T. Fritz (2020): *A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics*. Adv. Math. 370.

[5] T. Fritz, Tomáš Gonda, P. Perrone & Eigil Fjeldgren Rischel (2020): *Representable Markov Categories and Comparison of Statistical Experiments in Categorical Probability*. Available at `https://arxiv.org/abs/2010.07416`.

[6] Tobias Fritz, Tomáš Gonda, Nicholas Gauguin Houghton-Larsen, Antonio Lorenzin, Paolo Perrone & Dario Stein (2023): *Dilations and information flow axioms in categorical probability*. Mathematical Structures in Computer Science 33(10), pp. 913–957.

[7] Tobias Fritz, Tomáš Gonda, Antonio Lorenzin, Paolo Perrone & Dario Stein (2023): *Absolute continuity, supports and idempotent splitting in categorical probability*. arXiv preprint arXiv:2308.00651.

[8] Tobias Fritz, Tomáš Gonda & Paolo Perrone (2021): *De Finetti's Theorem in Categorical Probability*. arXiv:2105.02639.

[9] Tobias Fritz & Andreas Klingler (2023): *The d-separation criterion in categorical probability*. Journal of Machine Learning Research 24(46), pp. 1–49.

[10] Tobias Fritz & Wendong Liang (2023): *Free gs-Monoidal Categories and Free Markov Categories*. Appl. Categorical Struct. 31(2), p. 21, doi:10.1007/S10485-023-09717-0. Available at `https://doi.org/10.1007/s10485-023-09717-0`.

[11] Tobias Fritz & Eigil Fjeldgren Rischel (2020): *Infinite products and zero-one laws in categorical probability*. Compositionality 2, doi:10.32408/compositionality-2-3. Available at `https://doi.org/10.32408/compositionality-2-3`.

[12] Esfandiar Haghverdi & Philip J Scott (2005): *Towards a typed geometry of interaction*. In: International Workshop on Computer Science Logic, Springer, pp. 216–231.

[13] James Hefford & Cole Comfort (2022): *Coend optics for quantum combs*. arXiv preprint arXiv:2205.09027.

[14] Nicholas Gauguin Houghton-Larsen (2021): *A Mathematical Framework for Causally Structured Dilations and its Relation to Quantum Self-Testing*. Ph.D. thesis, University of Copenhagen.

[15] B. Jacobs (2020): *A channel-based perspective on conjugate priors*. Mathematical Structures in Computer Science 30(1), p. 44–61, doi:10.1017/S0960129519000082.

[16] Bart Jacobs, Aleks Kissinger & Fabio Zanasi (2019): *Causal inference by string diagram surgery*. In: Foundations of Software Science and Computation Structures: 22nd International Conference, FOSSACS 2019, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2019, Prague, Czech Republic, April 6–11, 2019, Proceedings 22, Springer, pp. 313–329.

[17] Elena Di Lavore & Mario Román (2023): *Evidential Decision Theory via Partial Markov Categories*. arXiv:2301.12989.

[18] Mario Román (2020): *Comb diagrams for discrete-time feedback*. arXiv preprint arXiv:2003.06214.

[19] Mario Román (2020): *Open diagrams via coend calculus*. arXiv preprint arXiv:2004.04526.

[20] Peter Selinger (2011): *A survey of graphical languages for monoidal categories*. New structures for physics, pp. 289–355.

[21] Dario Stein (2021): *Structural Foundations for Probabilistic Programming Languages*. Ph.D. thesis, University of Oxford.

[22] Dario Stein (2021): *Structural Foundations for Probabilistic Programming Languages*. Ph.D. thesis, University of Oxford.

[23] Dario Stein & Sam Staton (2021): *Compositional Semantics for Probabilistic Programs with Exact Conditioning (long version)*. In: *Proceedings of Thirty-Sixth Annual ACM/IEEE Conference on Logic in Computer Science (LICS 2021)*.

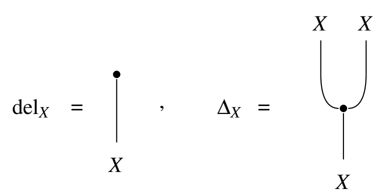
[24] Nathaniel Virgo (2023): *Unifilar Machines and the Adjoint Structure of Bayesian Models*. arXiv preprint arXiv:2305.02826.

7 Appendix

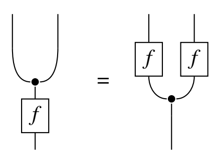
7.1 Recap on Markov Categories

This introductory material is taken from [4].

A *gs-monoidal category* or *CD-category* (due to [3]) is a symmetric monoidal category $(\mathbf{C}, \otimes, I)$ where every object is coherently equipped with the structure of a commutative comonoid $\Delta_X : X \to X \otimes X$, $\mathrm{del}_X : X \to I$. We render these graphically as



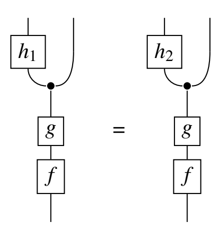
A morphism $f : X \to Y$ is called *discardable* if $\mathrm{del}_Y \circ f = \mathrm{del}_X$. A *Markov category* is a CD category in which every morphism is discardable. Equivalently, this means del is a natural transformation. Copying is not assumed to be natural – a morphism $f : X \to Y$ is called *deterministic* if it commutes with copying



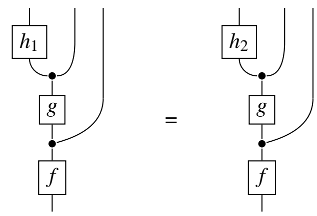
From now on, let $\mathbf{C}$ be a Markov category.

7.1.1 Information-flow Axioms

So called *information-flow axioms* (e.g. [4, 6]) are certain additional axioms which do not hold in every Markov category, but capture specific aspects of probabilistic reasoning. Important in this work is the so-called *causality* axiom: A Markov category is called *causal* if every equation of the form

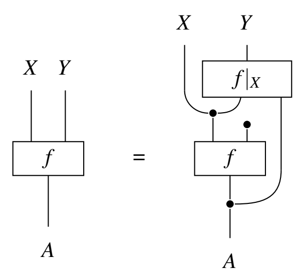


implies a stronger equation

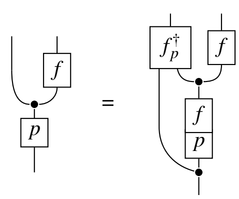


7.1.2 Conditionals

We say that $\mathbf{C}$ *has conditionals* if for every $f : A \to X \otimes Y$ there exists $f|_X : X \otimes A \to Y$ such that



For morphisms $p : A \to X$ and $f : X \to Y$, a *Bayesian inverse* is a morphism $f_p^\dagger : A \otimes Y \to X$ such that



Conditionals and Bayesian inverses are mutually interdefinable, and we will be using both in the present article. Conditionals are known to imply the information-flow axioms previously mentioned.

7.1.3 Example Categories

We briefly recall the Markov categories of interest, mainly to establish notation. For the full definitions, we ask the reader to consult the references.

Example 34. *The Markov category* **FinStoch** *consists of*

1. *objects are finite sets X*
2. *morphisms $p : X \to Y$ are stochastic matrices, with entries written $p(y|x) \in [0,1]$, subject to the axiom*
$$\forall x \in X, \sum_{y \in Y} p(y|x) = 1$$
3. *composition is matrix multiplication, a.k.a. the Kolmogorov-Chapman equation*
$$(gf)(z|x) = \sum_y g(z|y) f(y|x)$$

The hypergraph category $\mathbf{Mat}(\mathbb{R}^+)$ is defined like **FinStoch** allows arbitrary matrices with with nonnegative entries. **FinStoch** corresponds to the subcategory of $\mathbf{Mat}(\mathbb{R}^+)$ of morphisms which are discardable.

Example 35. *The Markov category* **BorelStoch** *consists of*

1. *morphisms are standard Borel spaces* (X, Σ_X)
2. *morphisms* $p : X \to Y$ *are Markov kernels* $p : X \times \Sigma_Y \to [0,1]$. *We write* $p(x)$ *for the measure* $p(x,-) : \Sigma_Y \to [0,1]$ *on* Y.
3. *composition is Lebesgue integration*

$$(gf)(x,E) = \int g(y,E) f(x, \mathrm{d}y)$$

For $x \in X$, we write δ_x for the Dirac distribution centered at x. We can consider every measurable function $f : (X, \Sigma_X) \to (Y, \Sigma_Y)$ as a **BorelStoch** morphism δ_f defined by $\delta_f(x) = \delta_{f(x)}$. As a slight abuse of notation, we will often write f for both the function and its induced Markov kernel.

The following Markov categories will not be relevant beyond mentioning in examples, but we will give references

1. **SetMulti**: sets and nonempty relations [4]
2. **TychStoch**: Tychonoff spaces and continuous Markov kernels [7]
3. **Gauss**: Markov kernels built from multivariate normal distributions and linear maps [4]

7.2 Traced Monoidal Categories

A traced monoidal category is a symmetric monoidal category $\mathbf{C}$ together with a family of operators

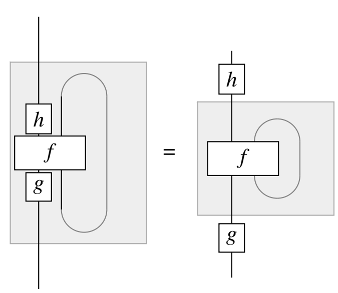
$$\mathbf{tr}^W_{X,Y} : \mathbf{C}(X \otimes W, Y \otimes W) \to \mathbf{C}(X,Y)$$

satisfying the following axioms:

1. Tightening (naturality in X,Y): For all $f : X \otimes W \to Y \otimes W$, $g : X' \to X$, $h : Y \to Y'$, we have

$$\mathbf{tr}^W((h \otimes \mathrm{id}_W) \circ f \circ (g \otimes W)) = h \circ \mathbf{tr}^W(f) \circ g$$

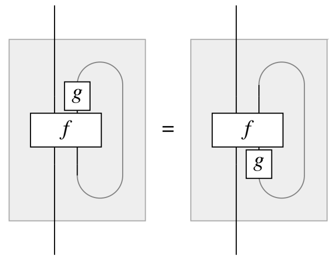
   or graphically



2. Sliding (dinaturality in W): For all $f : X \otimes W \to Y \otimes W'$ and $g : W' \to W$, we have

$$\mathbf{tr}^{W}((\mathrm{id}_Y \otimes g) \circ f) = \mathbf{tr}^{W'}(f \circ (\mathrm{id}_X \otimes g))$$

or graphically



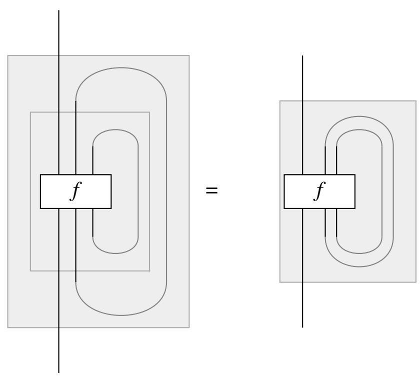
3. Vanishing (coherence with I). For all $f : X \otimes I \to Y \otimes I$, we have

$$\mathbf{tr}^{I}(f) = (X \xrightarrow{\rho_X^{-1}} X \otimes I \xrightarrow{f} Y \otimes I \xrightarrow{\rho_Y} Y)$$

4. Associativity (coherence with $\otimes$): For all $f : X \otimes U \otimes V \to Y \otimes U \otimes V$ we have

$$\mathbf{tr}^{U \otimes V}(f) = \mathbf{tr}^{U}(\mathbf{tr}^{V}(f))$$

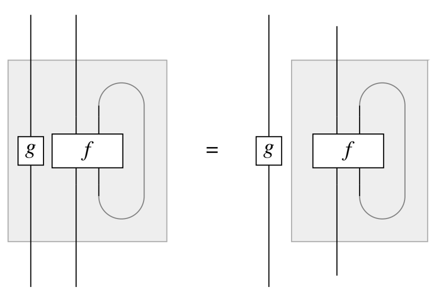
or graphically



5. Superposition (strength): For all $f : X \otimes W \to Y \otimes W$ and $g : X' \to Y'$, we have

$$\mathbf{tr}^{W}(g \otimes f) = g \otimes \mathbf{tr}^{W}(f)$$

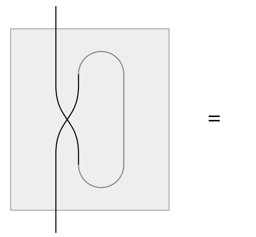
or graphically



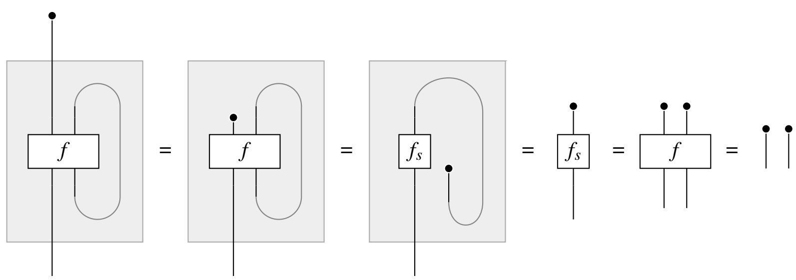
6. Yanking:

$$\mathbf{tr}^W(\mathrm{swap}_{W,W}) = \mathrm{id}_W$$

or graphically



Proof of Proposition 14. The trace of a discardable morphism f need not itself be discardable. For example, in the compact closed category $\mathbf{Mat}(\mathbb{R}^+)$, the trace of $\mathrm{id}_W : W \to W$ is the scalar $|W| : I \to I$ which is not normalized (i.e. equal to 1). On the other hand, if f is non-signalling, we obtain

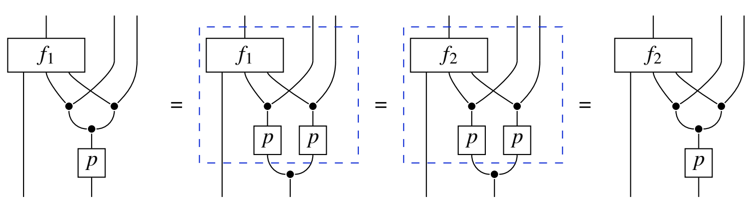


□

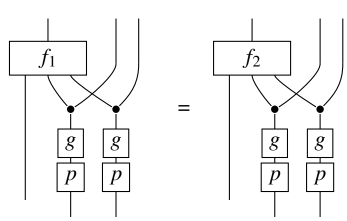
We will omit the grey shading of the trace boxes when it is clear from context how to interpret the trace.

7.3 Appendix to Section 2

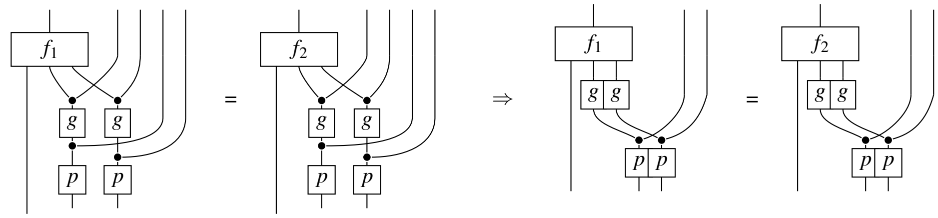
Proof of Proposition 7. The statement for morphisms of full support is immediate. For a deterministic morphism p, we reason straightforwardly



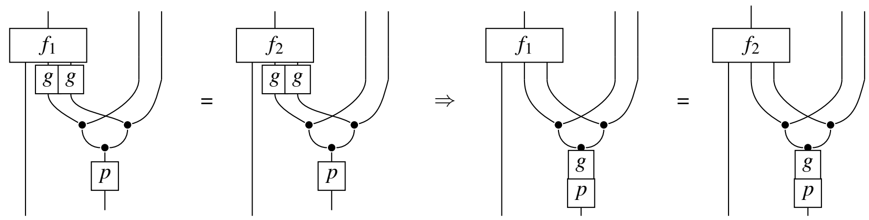
For the third point, let g be deterministic and p atomic, and assume



By causality [6, A.14], we may strengthen this equation and marginalize to obtain



Now, we use atomicity of p, postcompose the two rightmost wires with g and use determinism to simplify



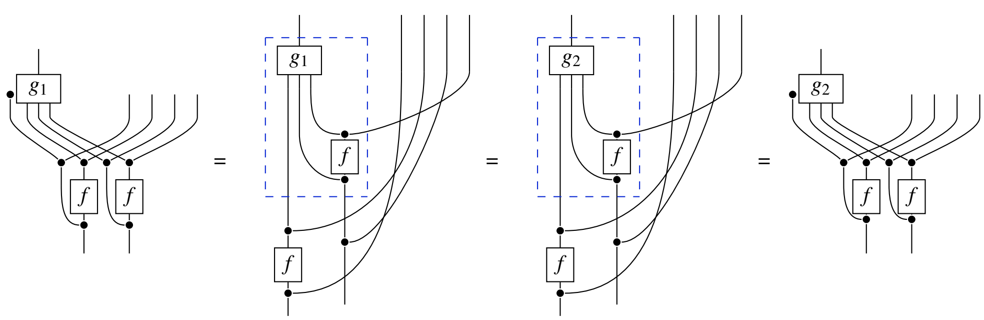
$\square$

Example 36. *Atomic morphisms need not be closed under composition: In* **BorelStoch***, let* $X = [0,1]$ *and define the morphism* $p : X \otimes 2 \to X$ *by*

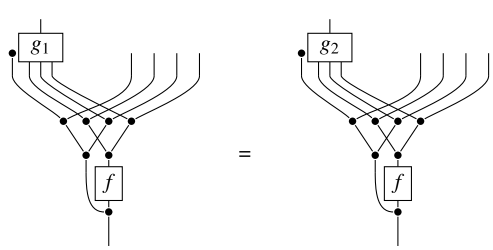
$$p(x,c) = \begin{cases} \delta_x, & c = 0 \\ \nu, & c = 1 \end{cases}$$

Then the atoms of p are all of X, i.e. p is atomic. However, for all deterministic states $\sigma : I \to X \otimes 2$ *of the form* $\sigma = \delta_{(x,1)}$*, we have that* $p \circ \sigma = \nu$ *is not atomic.*

Proof of Lemma 9. First, by marginalizing Y and cancellability of W, we conclude that $f_1 = f_2$ and write f indiscriminately. Now we show that the morphisms $c_i = \mathrm{del}_X \otimes g_i$ (for $i = 1,2$) and are $\phi \otimes \phi$-almost surely equal, where $\phi = (\mathrm{id}_X \otimes f) \circ \Delta_X$.

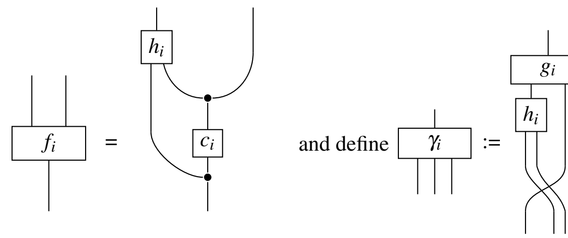


Applying atomicity to ϕ, we obtain that that c_1, c_2 are also $(\Delta \circ \phi)$-almost surely equal, i.e.

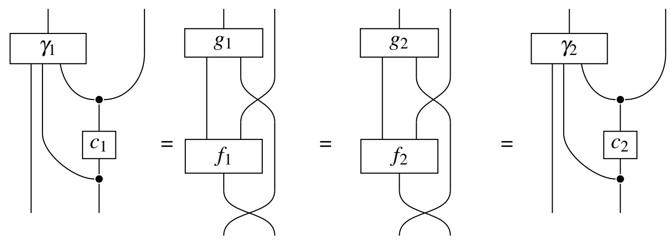


From this, we obtain the desired conclusion by marginalizing the four wires on the right. □

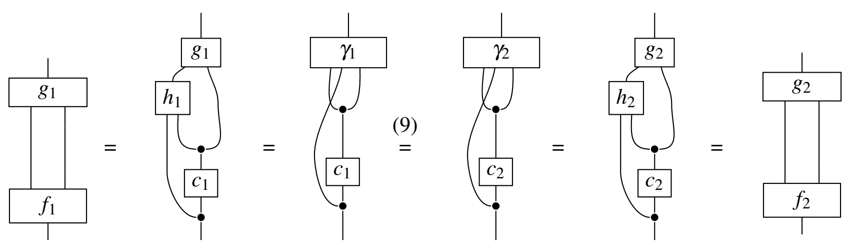
Proof of Lemma 10. For $i = 1, 2$, disintegrate f_i as



Then



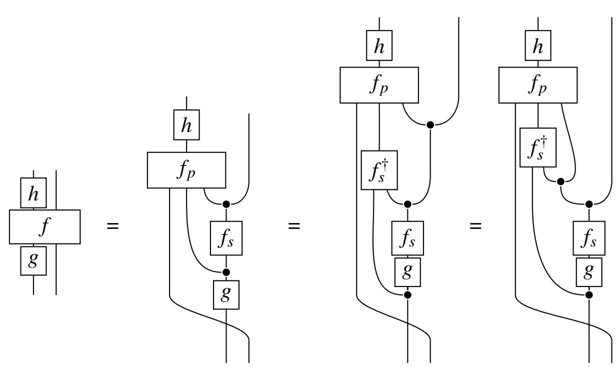
So by Lemma 9, we conclude



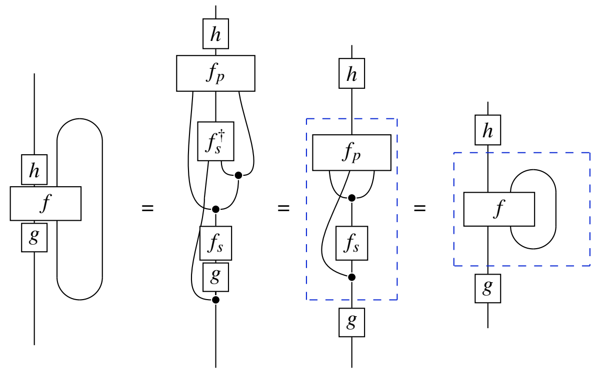
□

7.4 Appendix to Section 3.1

Proof of Tightening (Proposition 16). Write k for the composite $(h \otimes \mathrm{id}_W) \circ f \circ (g \otimes \mathrm{id}_W)$; it is easy to see that k is traceable. Using the Bayesian inverse $(f_s)^\dagger_g$, we obtain the following disintegration

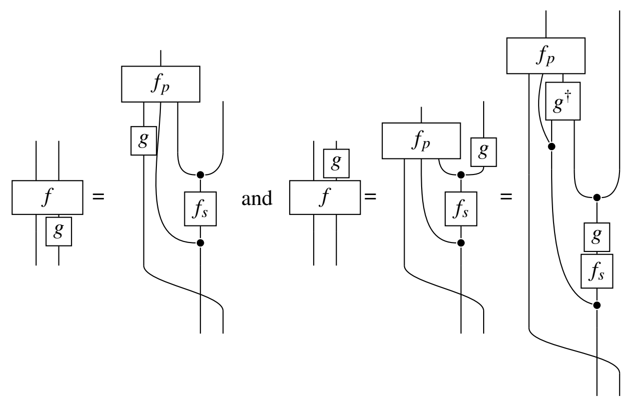


Using that disintegration, we compute as desired

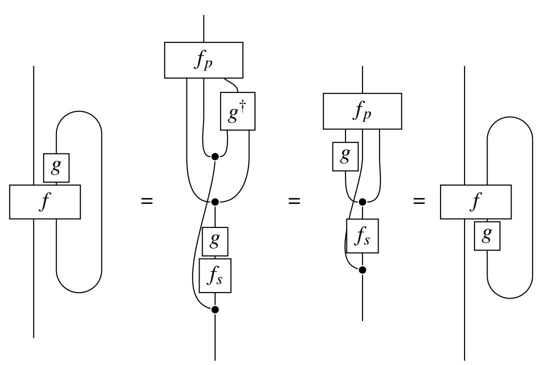


$\square$

Proof of Sliding (Proposition 17). Traceability is straightforward. Using the Bayesian inverse $g^\dagger_{f_s}$, we establish the following disintegrations:

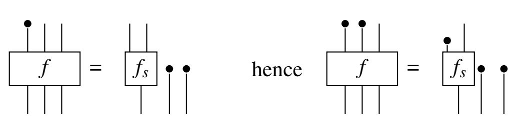


Using these disintegrations, we show

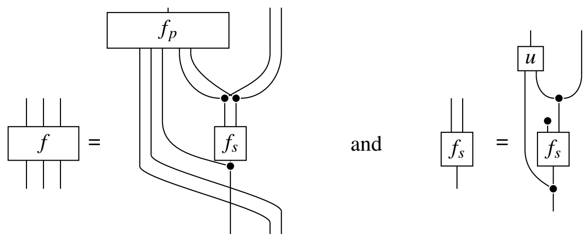


$\square$

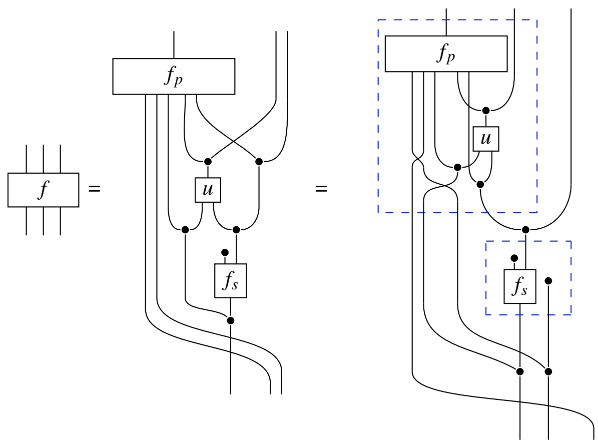
Proof of Associativity (Proposition 19). Assume that $f : X \otimes (U \otimes V) \to Y \otimes (U \otimes V)$ is non-signalling in $U \otimes V$; that is



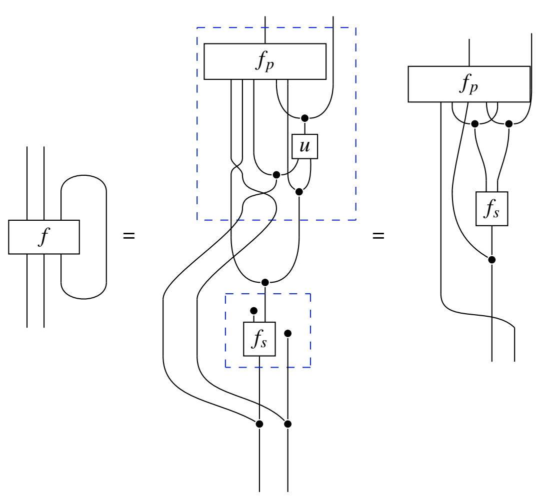
hence f is in particular non-signalling from V to V. Choose a disintegration f_p, f_s, and condition f_s further, to obtain



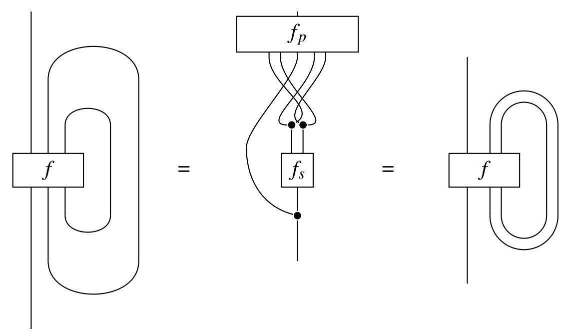
Then we can give the following disintegration with respect to V



Using this disintegration, we obtain $\mathbf{tr}^V(f)$ as

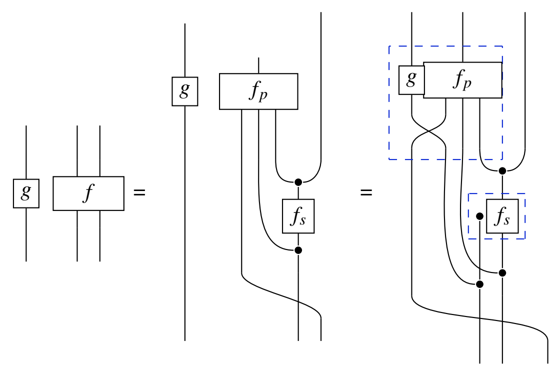


Of this, we can in turn compute the trace

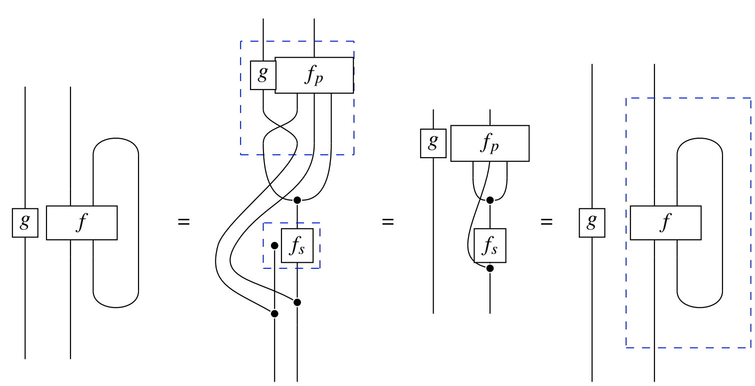


□

Proof of Superposition (Proposition 20). We make use of the following disintegration

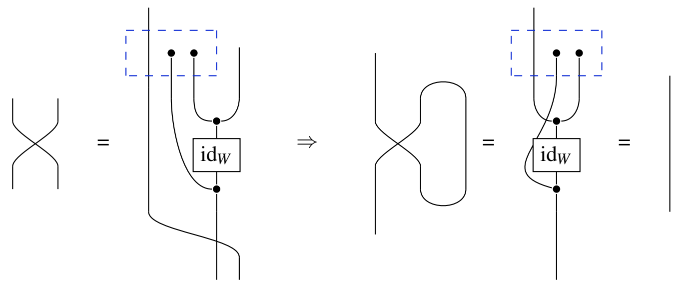


to obtain



□

Proof of Yanking (Proposition 21). The symmetry has the following disintegration, hence



□

7.5 Appendix to Section 4

We recall the construction of free Markov categories [10].

Definition 37 ([10])**.** *The category* $\mathbf{I}$ *is defined as follows.*

- *Objects are pairs of natural numbers* $(m,n) \in \mathbb{N} \times \mathbb{N}$*, and there is an extra object* $*$.
- *The only non-identity morphisms are* $\mathrm{in}_1, \cdots, \mathrm{in}_m, \mathrm{out}_1, \cdots, \mathrm{out}_n : (m,n) \to *$.

A hypergraph is a functor $\mathbf{I} \to \mathbf{Sets}$*. Hypergraphs form a functor category* **Hyp**.

We also use the following notation for a hypergraph $G : \mathbf{I} \to \mathbf{Sets}$.

- $W(G) = G(*)$ is the set of wires.
- $B_{m,n}(G)$ is the set of boxes with m inputs and n outputs. $B(G) = \coprod_{m,n \in \mathbb{N}} B_{m,n}(G)$ is the set of all boxes.
- We abbreviate $G(\mathrm{in}_i)$ to in_i and $G(\mathrm{out}_i)$ to out_i. These assign the ith input/output wire to each box.
- For $b \in B_{m,n}(G)$, $w \in W(G)$

$$\mathrm{in}(b,w) = \{\mathrm{in}_i(b) : i \in \{1, \cdots, , m\}\}$$
$$\mathrm{out}(b,w) = \{\mathrm{out}_i(b) : i \in \{1, \cdots, n\}\}$$

  These numbers tell how many times a wire is the input/output of a box.

Definition 38 ([10])**.** *A hypergraph is finite if the set of wires* $W(G)$ *and the set of boxes* $B(G)$ *is finite. We denote the subcategory of finite hypergraphs by* **FinHyp**.

Every monoidal signature Σ gives rise to a finite hypergraph which we also denote by Σ. A labelling of wires and boxes in a hypergraph G is a hypergraph homomorphism (natural transformation) to Σ.

Definition 39 ([10])**.** *Markov string diagrams over the monoidal signature* Σ *are (isomorphism classes of) cospans in the slice category* $\mathbf{FinHyp}/\Sigma$*, that is of the form*

$$\begin{tikzcd}
\underline{m} \arrow[r, "p"] \arrow[dr] & G \arrow[d] & \underline{n} \arrow[l, "q"'] \arrow[dl] \\
& \Sigma &
\end{tikzcd}$$

in **FinHyp***, such that*

1. $\underline{m}$ *and* $\underline{n}$ *are discrete hypergraphs* ($B(\underline{m}) = B(\underline{n}) = \emptyset$).

2. *G is acyclic, i.e. it contains no cycles. A path is a finite, alternating sequence of wires and boxes $(w_1, b_1, \cdots w_n, b_n, w_{n+1})$ such that $\mathrm{in}(b_i, w_i) > 0$, and $\mathrm{out}(b_i, w_{i+1} > 0)$ for all $1 \le i \le n$. A cycle is a path that additionally satisfies $w_1 = w_{n+1}$.*
3. *The cospan satisfies left monogamy: for all wires $w \in W(G)$*

$$|p^{-1}(w)| + \sum_{b \in B(G)} \mathrm{out}(b, w) = 1$$

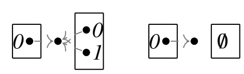
   That is, every wire is either connected to exactly one input port or is the output of a single box.
4. *There are no eliminable boxes. A box is eliminable if none of its output wires are connected to an output port or to the input of a box. That is, an eliminable box b satisfies for all $w \in W(G)$*

$$\mathrm{out}(b, w) > 0 \Longrightarrow q^{-1}(w) = 0 \wedge \forall b' \in B(G). \mathrm{in}(b', w) = 0$$

We leave the labelling implicit and write the cospan as $\underline{m} \xrightarrow{p} G \xleftarrow{q} \underline{n}$.

Proposition 40 ([10])**.** *The free Markov category* **FreeMarkov**$_\Sigma$ *over a monoidal signature Σ can be constructed as follows:*

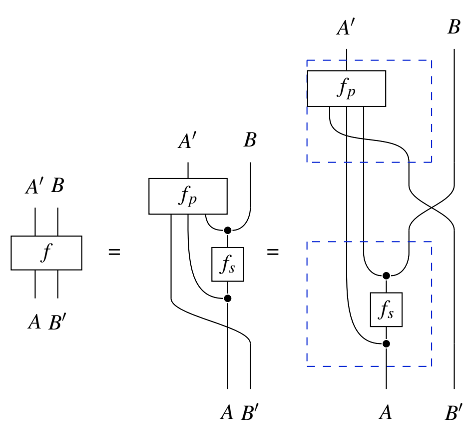
1. *Objects are homomorphisms $\underline{m} \to \Sigma$, that is an m-long list of types in Σ.*
2. *Morphisms are (isomorphism classes of) Markov string diagrams, which compose by pushout and subsequent normalisation (elimination of eliminable boxes).*
3. *The tensor is given by coproduct.*
4. *Copy and delete are represented by the cospans*



To ease notation we write $f : n \to m$ for a morphism $f : (\underline{m} \to \Sigma) \to (\underline{n} \to \Sigma)$, thus leaving the labelling implicit.

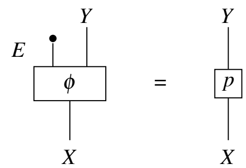
7.6 Appendix to Section 5

Proof of Theorem 32. For each comb $\langle f|g\rangle : (A, A') \to (B, B')$, its extension is a non-signalling morphism $A \otimes B' \to A' \otimes B$. Conversely, from a non-signalling morphism $f : A \otimes B' \to A' \otimes B$, we construct a disintegration

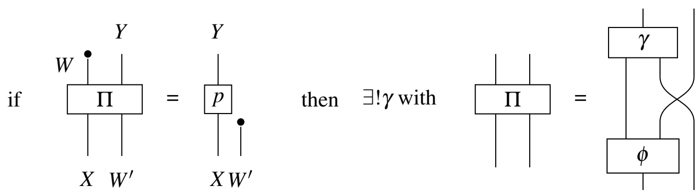


which defines a comb with environment $E = A \otimes B$ whose extension is f. By Proposition 31, the resulting comb is unique up to contextual equivalence. □

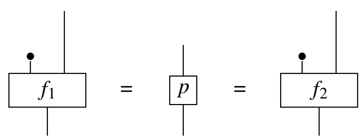
Universal Dilations In the terminology of [14, Definition 2.4.1], a *universal dilation* of a morphism $p : X \to Y$ is another morphism $\phi : X \to E \otimes Y$ satisfying



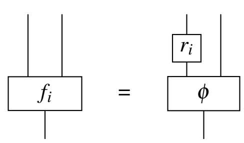
such that for all $\Pi : X \otimes W' \to W \otimes Y$



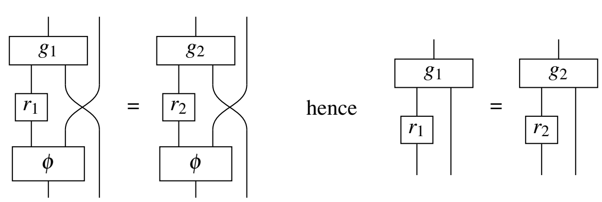
Proof of Proposition 33. Let $\langle f_1|g_1\rangle_{E_1}, \langle f_2|g_2\rangle_{E_2} : (A, A') \to (B, B')$ be two combs which are extensionally equivalent. Define using cancellability $p : A \to B$ as the common morphism



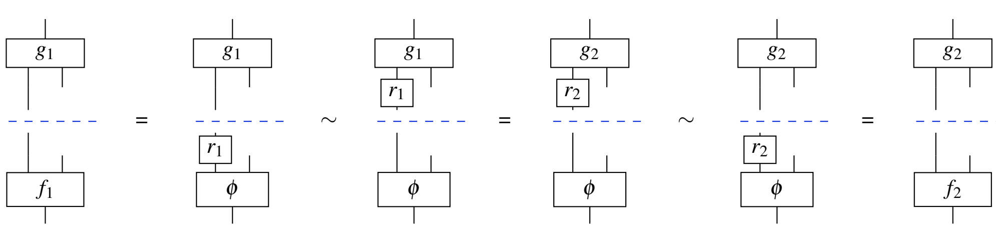
and choose a universal dilation $\phi : A \to E \otimes B$ of p. Using universality, the f_i must factor through ϕ as



By uniqueness of factorization, we have



Now we reason modulo optic equivalence that



□