Bimonoidal Structure of Probability Monads

Tobias Fritz[a,1] and Paolo Perrone[a,2]

[a] *Max Planck Institute for*
Mathematics in the Sciences
Leipzig, Germany

Abstract

We give a conceptual treatment of the notion of joints, marginals, and independence in the setting of categorical probability. This is achieved by endowing the usual probability monads (like the Giry monad) with a monoidal and an opmonoidal structure, mutually compatible (i.e. a bimonoidal structure). If the underlying monoidal category is cartesian monoidal, a bimonoidal structure is given uniquely by a commutative strength. However, if the underlying monoidal category is not cartesian monoidal, a strength is not enough to guarantee all the desired properties of joints and marginals. A bimonoidal structure is then the correct requirement for the more general case.
We explain the theory and the operational interpretation, with the help of the graphical calculus for monoidal categories. We give a definition of stochastic independence based on the bimonoidal structure, compatible with the intuition and with other approaches in the literature for cartesian monoidal categories. We then show as an example that the Kantorovich monad on the category of complete metric spaces is a bimonoidal monad for a non-cartesian monoidal structure.

Keywords: Monoidal monads, probabilistic powerdomain, strong monads, randomness, stochastic correlation, Kantorovich duality, Wasserstein distance.

1 Introduction

The standard way to treat randomness categorically is via a *probability monad*, of which classic examples are the Giry monad [7] and the probabilistic powerdomain [10]. The interpretation is the following: let C be a category whose objects we think of as spaces of possible values that a variable may assume. A probability monad P on C makes it possible to talk about random variables on objects $X \in \mathsf{C}$, or equivalently random elements of X: an element $p \in PX$ specifies the *law* of a random variable on X.

A central theme of probability theory is that random variables can form joints and marginals. For this to make sense in C, we need C to be a monoidal category, and we need P to interact well with the monoidal structure. We argue that this interaction is best modelled in terms of a bimonoidal structure.

A first structure which links a monad with the tensor product in a category is that of a *strength*. A strength for a probability monad is a natural map $X \otimes PY \to P(X \otimes Y)$, whose interpretation is the following: an element of X and a random element of Y determine uniquely a random element of $X \otimes Y$ which has the correct marginals, and whose randomness is all in the Y component. In the language of probability theory, $(x, q) \in X \otimes PY$ defines the product distribution of δ_x and q on $X \otimes Y$. In the literature, the operational meaning of a strength for a monad, which includes the usage in probability, is well explained in [16], and in [10] for the case of the probabilistic powerdomain. A compendium of probability monads appearing in the literature, with information about their strength, can be found in [9].

The monoidal structure can be thought of as a refinement of the idea of strength. The basic idea is that given two probability measures $p \in PX$ and $q \in PY$, one can canonically define a probability measure

---

This paper is electronically published in
Electronic Notes in Theoretical Computer Science
URL: www.elsevier.nl/locate/entcs

$p \otimes_\nabla q \in P(X \otimes Y)$, the "product distribution" [3]. This is not the only possible joint distribution that p and q have, but it can be obtained without additional knowledge (of their correlation). When a strength satisfies suitable symmetry conditions (commutative strength) it defines automatically a monoidal structure [11,8].

An opmonoidal structure formalizes the dual intuition, namely that given a joint probability distribution $r \in P(X \otimes Y)$ we canonically have the marginals on PX and PY as well. A bimonoidal structure is a compatible way of combining the two structures, in a way consistent with the usual properties of products and marginals in probability. When the underlying category is cartesian monoidal, then P is automatically opmonoidal. In this case, we show that if P carries a monoidal structure, then it is automatically bimonoidal. Therefore a commutative strong monad on a cartesian monoidal category is canonically bimonoidal. This is for example the case of the probabilistic powerdomain [10]. We argue that the bimonoidal structure is the structure of relevance for probability theory: if the underlying category is not cartesian monoidal, or the strength is not commutative, then one cannot talk about joints and marginals in the usual way just by having a strong monad.

However, not every probability monad in the literature is bimonoidal, not even strong; a famous counterexample is in [17]. While a non-bimonoidal probability monad could be of use in measure theory to talk about spaces of measures, it would be far from applications to probability, since it would not permit talking about concepts like stochastic independence and correlation, which in probability theory play a central role. We thus want to argue that in order for a monad to *really* count as a probability monad, it should be a bimonoidal monad.

In Section 2 we describe the setting of semicartesian monoidal categories and affine monads, which we argue is the one of relevance for classical probability theory. In such a setting, we will represent the concepts using a graphical calculus analogous to that of [14], presented in 2.1. In Section 3 we will sketch the basic theory and interpretation of a bimonoidal structure for probability monads, using the graphical calculus. The same definitions in terms of commutative diagrams can be found in Appendix A. In 3.1, we will show how this permits to talk about functions between products of random variables. In 3.2, we show how to define a category of probability spaces from a probability monad, in such a way that the monoidal structure is inherited. This permits to connect with other treatments of stochastic independence in the literature. In 3.3 we will see in more detail why this formalism generalizes the strength of probability monads on cartesian monoidal categories. In Section 4, we give a notion of stochastic independence based on the bimonoidal structure of the monad, and show that it satisfies some of the intuitively expected properties. In 4.1 we show that, if the base category is cartesian monoidal, our definition agrees with the one given by Franz [5], and it is compatible with the definition of independence structure given by Simpson [18]. Finally, in Section 5 we will give a nontrivial example of a bimonoidal monad, the Kantorovich monad on complete metric spaces [19,6]. The precise proofs and calculations of the statements of Section 5 can be found in Appendix B.

2 Semicartesian monoidal categories and affine monads

By definition, a *semicartesian monoidal category* is a monoidal category in which the monoidal unit 1 is a terminal object. For probability theory, this is a very appealing feature of a category, because such an object can be interpreted as a trivial space, having only one possible state. In other words, the object 1 would have the property that for every object X, $X \otimes 1 \cong X$ (monoidal unit), so that tensoring with 1 does not increase the number of possible states, and moreover there is a unique map $! : X \to 1$ (terminal object), which we can think of as "forgetting the state of X". Cartesian monoidal categories are in particular semicartesian. Not every monoidal category of interest in probability theory is cartesian, but most of them are semicartesian (in particular, all the ones listed in [9]).

Semicartesian monoidal categories have another appealing feature for probability: every tensor product space comes equipped with natural projections onto its factors:

$$X \otimes Y \xrightarrow{\text{id}\otimes!} X \otimes 1 \xrightarrow{\cong} X,$$

$$X \otimes Y \xrightarrow{!\otimes\text{id}} 1 \otimes Y \xrightarrow{\cong} Y,$$

which satisfy the universal property of the product projections if and only if the category is cartesian monoidal. These maps are important in probability theory, because they give the *marginals*. Since these projections are automatically natural in X and Y, a semicartesian monoidal category is always automatically a *tensor category with projections* in the sense of [5, Definition 3.3]; see [13] for more background.

---

[3] Our reason for denoting it by $p \otimes_\nabla q$ rather than by $p \otimes q$ is that we want to interpret $p : 1 \to PX$ and $q : 1 \to PY$ as morphisms, so that $p \otimes q : 1 \otimes 1 \to PX \otimes PY$ is not yet the product distribution. Rather, one needs to compose $p \otimes q$ with the monoidal structure $\nabla : PX \otimes PY \to P(X \otimes Y)$, which is the subject of the present paper, see Section 3.2.

Suppose now that P is a probability monad[4] on a semicartesian monoidal category C. Since we can interpret the unit 1 has having only one possible (deterministic) state, it is tempting to say that just as well there should be only one possible random state: if there is only one possible outcome, then there is no real randomness. In other words, it is appealing to require that $P(1) \cong 1$. A monad with this condition is called *affine*. Most monads of interest for probability are indeed affine (in particular, again, all the ones listed in [9]).

Unless otherwise stated, we will always work in a symmetric semicartesian monoidal category with an affine probability monad. These conditions simplify the treatment a lot, while keeping most other conceptual aspects interesting. By the remarks above, they seem to be the right framework for classical probability theory. The definition of monoidal, opmonoidal, and bimonoidal monads can however be given for general braided monoidal categories: the interested reader can find them in Appendix A.

2.1 Graphical calculus

Here we introduce a form of graphical calculus specializing that of Melliès [14] to our setting. Let C be a strict symmetric semicartesian monoidal category, and P an affine monad. We can represent objects X as vertical lines, and morphisms $f : X \to Y$ as boxes:

which we read from top to bottom.

Functor applications are represented by shadings. For example the image PX of X under a functor P and the functor image $Pf : PX \to PY$ of f are:

We can represent monoidal products by horizontal juxtaposition. For example, the map $f \otimes g : X \otimes A \to Y \otimes B$ can be represented as:

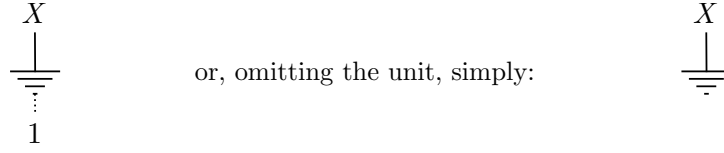The monoidal unit 1 is better represented by *nothing*, so that expressions like $X \otimes 1 \cong 1 \otimes X \cong X$ all have the same representation. However sometimes it is helpful to keep track of it, and in those cases we will draw it as a dotted line:
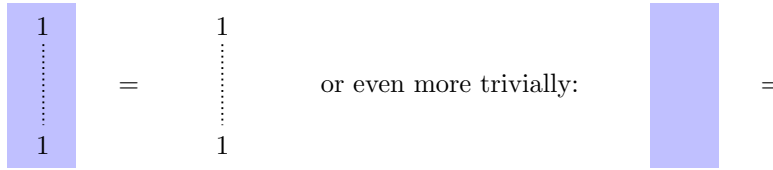

---

[4] In this work, "probability monad" is not a technical term: any monad could be in principle considered a probability monad. We merely use this term in order to indicate our intended interpretation in terms of randomness, as in the case of the Giry monad or the probabilistic powerdomain.

For every object X there is a unique map $! : X \to 1$, which we can interpret as "forgetting the state of X". We will represent such a map as a "ground wire", following the literature on quantum systems:

X or, omitting the unit, simply: X

1

The condition that P is affine, in picture, is

$$\begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} = \begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} \qquad \text{or even more trivially:} \qquad = $$

Since we are in a symmetric monoidal category, there is a canonical *braiding* isomorphism $X \otimes Y \to Y \otimes X$. We represent it as:

$$X \qquad Y$$

$$Y \qquad X$$
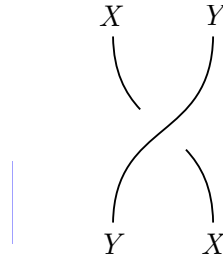
which one can think of as "swapping" X and Y. In a symmetric monoidal category, if we apply it twice, we obtain the identity.

We turn now to the monad structure of P. The monad unit $\delta : X \to PX$ is a natural transformation which "puts X into a shading", while the multiplication $E : PPX \to PX$ goes from a double shading to a single shading:

$$X \qquad\qquad \text{and} \qquad\qquad X$$

$$X \qquad\qquad\qquad\qquad\qquad\qquad X$$

We do not draw a box for these "structure maps": we consider them the canonical maps from their source to their target. The diagrams above will always denote δ and E, never other morphisms.

## 3   Monoidal structure of probability monads

Let P be an affine probability monad on a strict symmetric semicartesian monoidal category C. In this setting, a *monoidal structure* for the functor P amounts to a natural map $\nabla : PX \otimes PY \to P(X \otimes Y)$ with associativity and unitality conditions. In terms of graphical calculus, ∇ is a way to pass from $PX \otimes PY$, i.e.:

$$X \qquad\qquad Y$$

to $P(X \otimes Y)$, i.e.:


so we can represent it as:
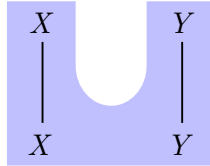
We again do not put any box, as we consider it the canonical map of the form given by the diagram above. The probabilistic interpretation is the following: given $p \in PX$ and $q \in PY$, there is a canonical (albeit not unique) way of obtaining a joint in $P(X \otimes Y)$, namely the product probability. Technically we also should need a map $1 \to P(1) \cong 1$, i.e.

or, omitting the unit, simply
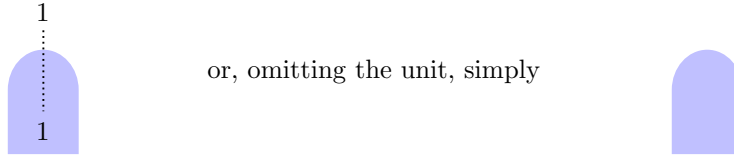
But due to our affineness assumption, such a map can only be the identity. The associativity condition now says that it does not matter in which way we multiply first:
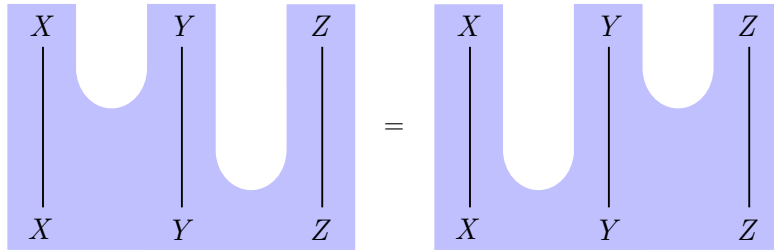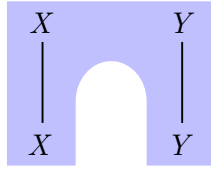

so that there is really just one way of forming a product of three probability distributions. The left and right unitality conditions say that:
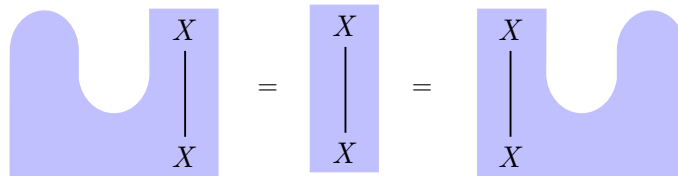
which means that the product distribution of some $p \in PX$ with the unique measure on 1 is the same as just p.

An *opmonoidal structure* for the functor P amounts to a natural map $\Delta : P(X \otimes Y) \to PX \otimes PY$, which
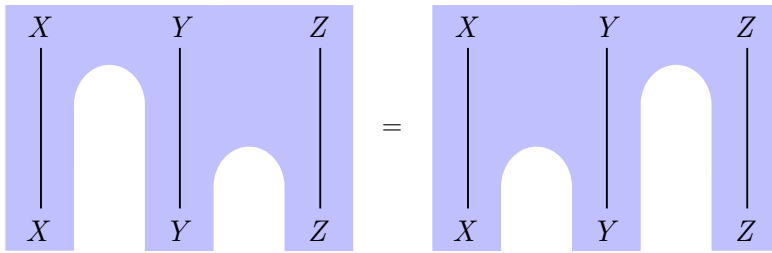
5

we represent as:

and again a map $P(1) \to 1$, i.e.

which in this setting can only be the identity. We have, dually, a coassociativity condition:

The probabilistic interpretation is that given a joint probability distribution $r \in P(X \otimes Y)$, we can canonically obtain marginal distributions on PX and PY, and again, if we have many factors, it does not matter in which order we take the marginals. Analogously, we have left and right counitality conditions:
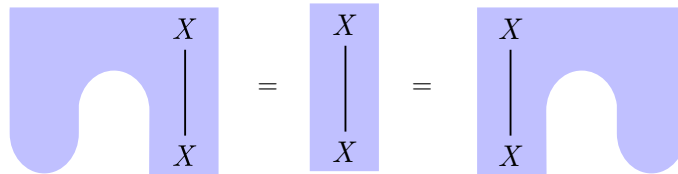
which say that the marginal distribution of some $p \in P(X \otimes 1)$ on the first factor (or of some $p \in P(1 \otimes X)$ on the second factor) is just p again.

The monoidal and opmonoidal structure should interact to form a *bimonoidal structure* [1] for the functor P. To have that, we have first of all some unit-counit conditions, which in our setting are trivially satisfied, since they only involve maps to 1. But more importantly, the following bimonoidality (or distributivity) condition needs to hold:

$$(1)$$
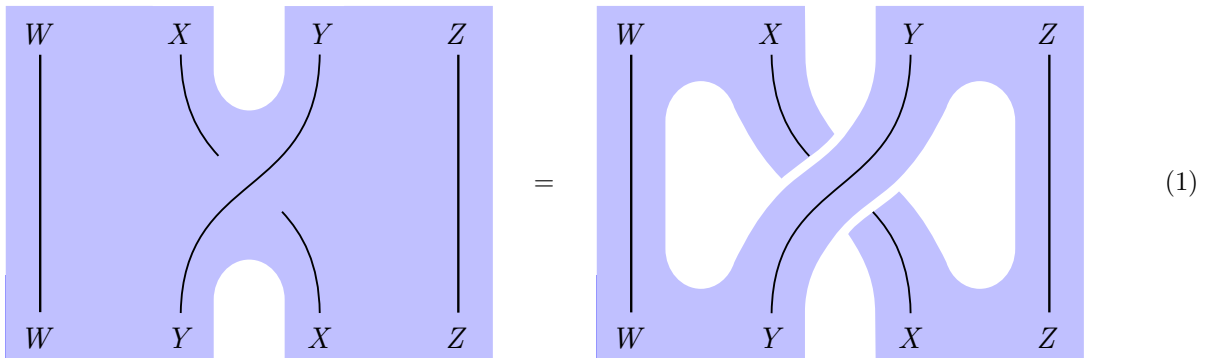
where the center of the diagram on the right is a swap of PX and PY. The probabilistic interpretation is a bit involved, and it has to do with stochastic independence. We will analyze it separately in Section 4.

We can say even more about the structure of joints and marginals: the whole monad structure should respect the bimonoidal structure of P, i.e. $\delta : X \to PX$ and $E : PPX \to PX$ commute with taking joint and marginals. In other words, we are saying that δ and E should be *bimonoidal natural transformations*. In terms of diagrams, we are saying that first of all, δ commutes with the monoidal multiplication and comultiplication:

Probabilistically, this means that the delta over the product is the product of the deltas, and that a delta over a product space has as marginals a pair of deltas over the projections.

The same can be said about the average map E. It commutes with the multiplication and with the comultiplication:

which means that the product of the average is the average of the product, and that the marginals of an average are the averages of the marginals. These last conditions may seem a bit obscure, but they come up naturally in probability: see as an example the case of the Kantorovich monad (Section 5 and its proofs in Appendix B).

We are in other words requiring that P is a bimonoidal monad.

Definition 3.1 A *bimonoidal monad* (P, δ, E) is a monad whose functor is a bimonoidal functor, and whose unit and multiplication are bimonoidal natural transformations.

The definition above works in general, however the particular conditions for the monoidal and opmonoidal structure which have been given here suffice only in the specific context of a semicartesian monoidal category with an affine monad. In Appendix A there is a more general definition, for generic symmetric monoidal categories. The definition given there specializes to the one given above in this context.

As far as we know, this kind of structure has not been considered before in this exact form. Monads in a general bicategory are a standard concept, however to the best of our knowledge the bicategory of monoidal categories, bimonoidal functors, and bimonoidal natural transformations has not been used explicitly. In particular, it has not been used in categorical probability. To avoid possible confusion, let us also point out that the notion of a bimonoidal monad is a distinct concept from that of a *bimonad* [21].

Most probability monads in the literature have an additional symmetry: the multiplication and comultiplication commute with the braiding, i.e. they are equivariant with respect to permutations of random variables.

This means in diagrams that
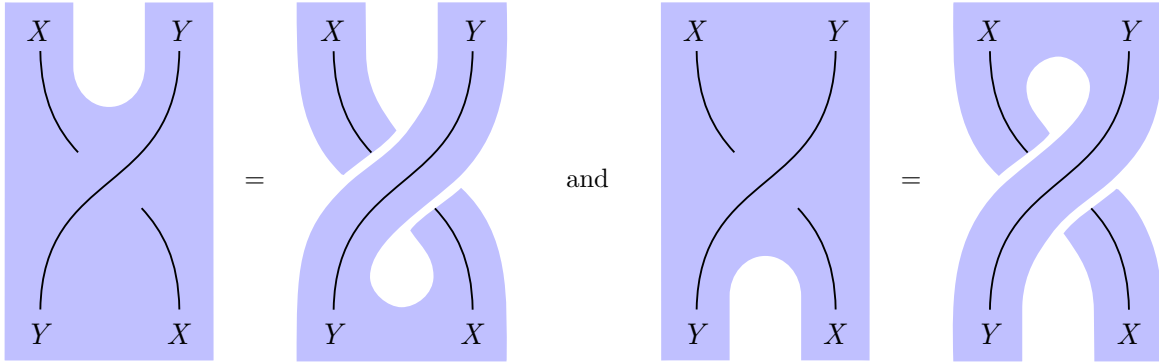
Such a functor (and such a monad) is called *braided* or *symmetric*. A definition in terms of traditional commutative diagrams can again be found in Appendix A.

3.1 Algebra and coalgebra of random variables

The so-called "law of the unconscious statistician" says that given a function $f : X \to Y$ and a random variable on X with law $p \in PX$, the law of the image random variable under f will be the push-forward of p along f. In categorical terms, this simply means that P is a functor, and that the image random variable has law $(Pf)(p)$, where $Pf : PX \to PY$ is given by the push-forward.

The bimonoidal structure of P comes into play whenever we have functions to and from product spaces. Consider a morphism $f : X \otimes Y \to Z$, which we represent as:
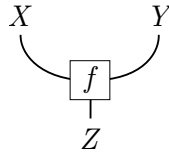
Given random variables X and Y, we can form an image random variable on Z in the following way: first we form the joint on $X \otimes Y$ using the monoidal structure, and then we form the image under f. In other words, in terms of laws we perform the following composition:
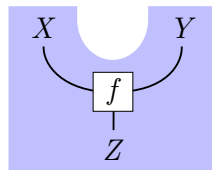
For maps in the form $g : X \to Y \otimes Z$ we can proceed analogously by forming the marginals, using the opmonoidal structure:
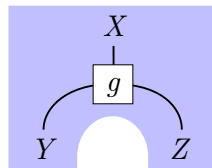
This way, together with associativity and coassociativity, one can form functions to and from arbitrary products of random variables.

Whenever we have an internal structure, like an internal monoid or group, this way we can extend the operations on the random elements, via convolution. For example, if X is a monoid, then also PX becomes a monoid, using $PX \otimes PX \to P(X \otimes X) \to PX$ for the multiplication. The analogous statements apply for coalgebraic structures. In other words, the bimonoidal structure allows to have an *algebra and coalgebra of random variables* whenever the deterministic variables form an internal algebraic structure.

3.2 The category of random elements

In the literature, many categorical treatments of statistical dependence work in categories whose objects are probability spaces, or fixed probability measures on a space, rather than categories with a probability monad [5,18]. One can form probability spaces from a probability monad in a canonical way:

Definition 3.2 Let C be a category with terminal object 1 and P a probability monad on C. Then the category $\mathsf{Prob}(\mathsf{C})$ is defined to be the comma category $1/P$. In other words:

- Objects of $\mathsf{Prob}(\mathsf{C})$ are objects X of C together with arrows $1 \to PX$ of C, which can be interpreted as probability measures on X;
- Morphisms of $\mathsf{Prob}(\mathsf{C})$ are maps $f : X \to Y$ of C which makes the diagram

$$
\begin{array}{ccc}
 & 1 & \\
 \swarrow & & \searrow \\
PX & \xrightarrow{\ Pf\ } & PY
\end{array}
$$

commute, which can be interpreted as measure-preserving maps.

In analogy with the category of elements, we can interpret $\mathsf{Prob}(\mathsf{C})$ as a *category of random elements*, or of probability spaces. This is because a map $1 \to PX$ determines in the usual settings a probability measure on X, or the law of a random element of X.

Under some mild assumptions, if C has a semicartesian monoidal structure we can transfer that structure to the category of random elements, with a construction analogous to that of Section 3.1.

Definition 3.3 Let C be a semicartesian monoidal category and P an affine probability monad on C with monoidal structure ∇. We define the following monoidal structure on $\mathsf{Prob}(\mathsf{C})$: given $p : 1 \to PX$ and $q : 1 \to PY$, we define $p \otimes_\nabla q : 1 \to P(X \otimes Y)$ to be the composition:

$$
1 \cong 1 \otimes 1 \xrightarrow{p \otimes q} PX \otimes PY \xrightarrow{\ \nabla\ } P(X \otimes Y).
$$

and for morphisms we proceed analogously.

This way $(\mathsf{Prob}(\mathsf{C}), \otimes_\nabla)$ is a semicartesian monoidal category, with the unit $1 \to 1$ isomorphic to the terminal object. In particular, it is always a tensor category with projections in the sense of [5], generalizing the construction given in Section 3.1 therein (in which the base category Meas is cartesian monoidal).

Some of the upcoming results will refer to $\mathsf{Prob}(\mathsf{C})$, whose objects we also call *laws*, as they generalize laws of random variables. In particular we will use the notation $p \otimes_\nabla q$ for the product probability.

3.3 Bimonoidal monads on a cartesian monoidal category

Suppose now that the monoidal structure of C is *cartesian* monoidal, i.e. that the monoidal product is given by the categorical product (so, in particular, C is semicartesian). The projection maps $\pi_1 : X \times Y \to X$ and $\pi_2 : X \times Y \to Y$ now satisfy a universal property. Let's now apply P, so that we get maps $P\pi_1 : P(X \times Y) \to PX$ and $P\pi_2 : P(X \times Y) \to PY$. By the universal property of the product, there is then a *unique* map $P(X \times Y) \to PX \times PY$ compatible with the projections, i.e. making the following diagram commute:

$$
\begin{array}{ccc}
 & P(X \times Y) & \\
{\scriptstyle P\pi_1}\swarrow & \downarrow & \searrow{\scriptstyle P\pi_2} \\
PX \xleftarrow{\ \pi_1\ } & PX \times PY & \xrightarrow{\ \pi_2\ } PY
\end{array}
$$

This gives a natural map $\Delta : P(X \times Y) \to PX \times PY$. Such a map exists and is unique for any (finite) number of factors, so it is automatically associative. Therefore P has a canonical opmonoidal structure. This is true

for all functors P between cartesian monoidal categories. Moreover, this opmonoidal structure is unique, due to naturality,

$$
\begin{array}{ccc}
P(X \times Y) & \xrightarrow{\Delta_{X,Y}} & PX \times PY \\
\downarrow & & \downarrow \\
P(X \times 1) & \xrightarrow{\Delta_{X,1} = 1_X} & PX
\end{array}
$$

Suppose now that P in addition has a (given) monoidal structure ∇. By the universal property of the product, it is straightforward to see that the bimonoid diagram (1) commutes automatically. Therefore, whenever C is cartesian monoidal, it suffices to have a monoidal structure to obtain a bimonoidal structure:

Proposition 3.4 *In a cartesian monoidal category, a bimonoidal monad is the same structure as a monoidal monad.*

In particular, since a monoidal structure is equivalent to a commutative strength (see [11] for the closed monoidal case, and [8, Appendix A4] for the general case), a commutative strong monad on a cartesian monoidal category is automatically bimonoidal in a unique way. This is what happens, for example, for the probabilistic powerdomain on the category of domains. However, not all bimonoidal probability monads arise in this way. In Section 5, we will give an example of a bimonoidal probability monad on a non-cartesian monoidal category, the Kantorovich monad on complete metric spaces.

4 Stochastic independence

Our framework allows to give a formal definition of stochastic dependence and independence in categorical terms, closely related to other notions appearing in the literature [5,18].

First of all, we look at an important consequence of the bimonoidality condition (1): stochastic *dependence* can only be forgotten, not created. Consider two spaces X and Y. Then given a joint distribution $r \in P(X \otimes Y)$, we can form the marginals $r_X \in PX$ and $r_Y \in PY$. If we try to form a joint again, via the product, the correlation is lost. Vice versa, instead, if we have two marginals, form their joint, and then divide them again into marginals, we expect to get our initial random variables back. Graphically:

$$(2)$$

This is indeed the case under the assumptions that we've made so far:

Proposition 4.1 *Let X, Y be objects of a symmetric semicartesian monoidal category C. Let $P : \mathsf{C} \to \mathsf{C}$ be a bimonoidal endofunctor, with $P(1) \cong 1$. Then $\Delta \circ \nabla = \mathrm{id}_{PX \otimes PY}$. In particular, $PX \otimes PY$ is a retract of $P(X \otimes Y)$.*

The proposition above is proved graphically in Appendix B.1. It is a special case of a standard result about the so-called *normal* bimonoidal functors, which can be found for example in [1, Section 3.5].

In general we do *not* get any condition $\nabla \circ \Delta = \mathrm{id}_{P(X \otimes Y)}$, i.e. in general

$$(3)$$

An example is given by $X = Y = \{0, 1\}$, with a perfectly correlated and uniform distribution. So correlation can be forgotten, but not created, by the bimonoidal structure maps.

Going further, we can use these structures in order to talk about probabilistic independence:

Definition 4.2 X and Y are *independent* for the law $r : 1 \to P(X \otimes Y)$ if and only if $\nabla \circ \Delta \circ r = r$.

That is, applying the left-hand side of (3) gives the same as applying the right-hand side if and only if we have independence.

We are now ready for the probabilistic interpretation of the bimonoidality condition (1), which gives its main motivation: Consider any joints WX and YZ, and form their product. In the resulting distribution, W will be independent of Y, and X will be independent of Z. More rigorously:

Proposition 4.3 *Let* W, X, Y, Z *be objects of a symmetric semicartesian monoidal category* C. *Let* $P : \mathsf{C} \to \mathsf{C}$ *be a bimonoidal functor, with* $P(1) \cong 1$. *Let* $r : 1 \to P(W \otimes X)$ *and* $s : 1 \to P(Y \otimes Z)$, *and consider the law* $r \otimes_\nabla s := \nabla \circ (r \otimes s)$ *on* $W \otimes X \otimes Y \otimes Z$. *Then after forgetting* X *and* Z, *for the resulting law* W *and* Y *are independent. Just as well, after forgetting* W *and* Y, *for the resulting law* X *and* Z *are independent.*

A graphical proof in terms of Definition 4.2 is given as well in Appendix B.1.

This result forms part of the semi-graphoid axioms [15] which axiomatize properties of conditional independence, namely in the case where the conditioning is trivial. Concretely, Proposition 4.3 corresponds to the axiom of decomposition, stating that if X is independent from (Y, Z), then X is also independent from Y. The semi-graphoid axiom of symmetry (if X is independent of Y, then Y is independent of X) is also satisfied whenever we have a symmetric bimonoidal monad.

4.1   Comparison with other notions of independence

Franz [5] defines stochastic independence in a semicartesian monoidal category in the following way: given objects A, B_1, B_2 (which one can think of as probability spaces), and arrows $f_1 : A \to B_1$ and $f_2 : A \to B_2$ (which one can think of measure-preserving maps), then f_1 and f_2 are independent if and only if there exists $h : A \to B_1 \otimes B_2$ making this diagram commute:

$$
\begin{array}{ccc}
& A & \\
{\scriptstyle f_1} \swarrow & \downarrow {\scriptstyle h} & \searrow {\scriptstyle f_2} \\
B_1 \xleftarrow{\ \pi_1\ } & B_1 \otimes B_2 & \xrightarrow{\ \pi_2\ } B_2
\end{array}
\tag{4}
$$

where π_1, π_2 are the projections of the tensor product. He then proves [5, Proposition 3.5] that in the category Prob of (traditional) probability spaces, this notion of independence is equivalent to the standard one of probability theory. We propose a generalization of that result, which holds for categories of random elements obtained by generic *cartesian* monoidal categories.

Proposition 4.4 *Let* C *be a cartesian monoidal category and* P *an affine bimonoidal probability monad. Consider a law* $s : 1 \to PA$ *and maps* $f_1 : A \to B_1$ *and* $f_2 : A \to B_2$. *Then* f_1 *and* f_2 *are independent in the sense of Franz [5] if and only if* B_1 *and* B_2 *are independent for the law* $P(f_1, f_2) \circ s$ *in the sense of Definition 4.2.*

So in the case of cartesian monoidal base categories, the two approaches agree. The proof can be found in Appendix B.2, and goes along the lines of the proof of [5, Proposition 3.5].

Simpson [18] defines an *independence structure* as a certain collection of multispans that contains the singleton families. Given again a *cartesian* monoidal category C and an affine monad P on C, and given a finite multispan $\{f_i : A \to B_i\}_{i \in I}$ in C, we can form a multispan in the category $\mathsf{Prob}(\mathsf{C})$ by precomposing with a law $r : 1 \to PA$. We can call such a resulting multispan *independent*, in analogy with Definition 4.2, iff

$$
\nabla_I \circ \Delta_I \circ P\left((f_i)_{i \in I}\right) \circ r = P\left((f_i)_{i \in I}\right) \circ r,
$$

where $(f_i)_{i \in I} : A \to \prod_{i \in I} B_i$ is the tupling of the f_i given by the cartesian monoidal structure, and ∇_I and Δ_I are the maps respectively $\prod_i (PB_i) \to P(\prod_i B_i)$ and $P(\prod_i B_i) \to \prod_i (PB_i)$ obtained by iterating respectively ∇ and Δ (by associativity and coassociativity, the resulting maps are unique). Independent multispans defined in this way form then an independence structure in the sense of [18, Definition 2.1], in a way analogous to Examples 2.1 and 2.2 therein: they are closed with respect to multispan composition, and to forming subfamilies. Therefore, again in the case of a cartesian monoidal base category, our definition is compatible with Simpson's approach.

5   Bimonoidal structure of the Kantorovich monad

The Kantorovich monad is a probability monad on complete metric spaces. It was first defined by van Breugel for compact and for complete 1-bounded metric spaces [19]. We will use here the definitions and results of [6],

which work for all complete metric spaces.

Consider the category CMet whose:

- Objects are complete metric spaces;
- Morphisms are short maps, i.e. functions $f : X \to Y$ such that

$$d\big(f(x), f(x')\big) \leq d(x, x')$$

for all $x, x' \in X$;

- As monoidal structure, we define $X \otimes Y$ to be the set $X \times Y$, with the metric:

$$d\big((x, y), (x', y')\big) := d(x, x') + d(y, y').$$

This category can be thought of as a category of enriched categories and functors [12, Section 2], and the monoidal structure is closed but not cartesian. Further motivation for the choice of this category is given in [6]. In particular, by choosing as morphisms the short maps, one can obtain PX as a colimit of spaces of empirical distributions of finite sequences [6, Section 3], which would not be possible if one allowed for more general morphisms (like continuous or Lipschitz functions).

We recall the basic definitions of [6].

Definition 5.1 Let X be a complete metric space.

- A Radon probability measure p on X is said to have *finite first moment* if for every short map $f : X \to \mathbb{R}$,

$$\int_X f \, dp < \infty.$$

Every such probability measure can be specified uniquely by its integration against short maps to $\mathbb{R}$: the set of such measures can be identified with the set of positive, Scott-continuous linear functionals on the space of Lipschitz functions on X. Hence, in the following, we explicitly construct such measures by specifying their action on short maps.

- The *Kantorovich-Wasserstein* space PX is the space of all Radon probability measures on X with finite first moment, equipped with the metric:

$$d(p, q) := \sup_{f : X \to \mathbb{R}} \left| \int_X f \, dp - \int_X f \, dq \right|,$$

where the supremum ranges over all short maps $X \to \mathbb{R}$. With this metric, PX is itself a complete metric space.

- Given $f : X \to Y$, we define $Pf : PX \to PY$ as the map assigning to $p \in PX$ its push-forward measure $(Pf)(p) := f_*p \in PY$. The latter is defined by saying that for all $g : Y \to \mathbb{R}$ short,

$$\int_Y g \, d(f_*p) := \int_X g \circ f \, dp.$$

f_*p also has finite first moment, and this assignment makes P into a functor.

A concise treatment of Wasserstein spaces can be found in [2] and a more comprehensive one in [20]. For the basic measure-theoretic setting, we refer the reader to [3,4].

The functor P admits a monad structure, with the unit $\delta : X \to PX$ given by the Dirac distributions

$$\int_X f(y) \, d(\delta(x))(y) := f(x),$$

and the multiplication $E : PPX \to PX$ given by forming the expected or average distribution,

$$\int_X f \, d(E\mu) := \int_{PX} \left(\int_X f(x) \, dp(x) \right) d\mu(p).$$

12

We can now define product joints and marginals, which will equip P with a bimonoidal structure.

Definition 5.2 Let $p \in PX, q \in PY$. We denote $p \otimes_\nabla q$ the joint probability measure on $X \otimes Y$ defined by:

$$\int_{X \otimes Y} f(x,y)\, d(p \otimes_\nabla q)(x,y) := \int_{X \otimes Y} f(x,y)\, dp(x)\, dq(y).$$

Let now $r \in P(X \otimes Y)$. We denote (r_X) the marginal probability on X defined by:

$$\int_X f(x)\, dr_X(x) := \int_{X \otimes Y} f(x)\, dr(x,y).$$

The marginal on Y is defined analogously.

It is straightforward to check that the functionals defined in Definition 5.2 are positive, linear, and Scott-continuous, therefore they specify uniquely Radon probability measures of finite first moment.

In the rest of this section we will show that the joints and marginals in Definition 5.2 equip the Kantorovich monad on CMet with a bimonoidal monad structure (Theorem 5.15). The proofs with the actual calculations are in Appendix B.

We will prove now that the product joint construction equips P with a monoidal structure.

Definition 5.3 Let $X, Y \in$ CMet. We define the map $\nabla : PX \otimes PY \to P(X \otimes Y)$ as mapping $(p,q) \in PX \otimes PY$ to the joint $p \otimes_\nabla q \in P(X \otimes Y)$.

Proposition 5.4 $\nabla : PX \otimes PY \to P(X \otimes Y)$ is short.

Therefore, ∇ is a morphism of CMet. This would not be the case if we took as monoidal structure for CMet the cartesian product: for the product metric, ∇ is Lipschitz, but in general not short. The fact that ∇ equips P with a monoidal structure now follows directly from the naturality and associativity of the product probability construction (as sketched in Section 3). In other words, the proofs of the next three statements (see Appendix B.3) can be adapted to most other categorical contexts in which the map ∇ is of a similar form.

Proposition 5.5 $\nabla : PX \otimes PY \to P(X \otimes Y)$ is natural in X and Y.

Proposition 5.6 $(P, \mathrm{id}_1, \nabla)$ is a symmetric lax monoidal functor CMet $\to$ CMet.

Proposition 5.7 (P, δ, E) is a symmetric monoidal monad.

We know that a monoidal monad is the same as a commutative monad, and therefore obtain:

Corollary 5.8 P is a commutative strong monad, with strength $X \otimes PY \to P(X \otimes Y)$ given by:

$$(x, q) \mapsto \delta_x \otimes_\nabla q \in P(X \otimes Y).$$

We now turn to the analogous statements for the marginals, and show that they equip P with an opmonoidal structure.

Definition 5.9 Let $X, Y \in$ CMet. We define the map $\Delta : P(X \otimes Y) \to PX \otimes PY$ as mapping $r \in P(X \otimes Y)$ to the pair of marginals $(r_X, r_Y) \in PX \otimes PY$.

Proposition 5.10 $\Delta : P(X \otimes Y) \to PX \otimes PY$ is short.

Therefore Δ is a morphism of CMet. Again, the following statements follow just from the properties of marginals, and their proofs (see Appendix B.4) can be adapted to most other categorical contexts provided that Δ is of a similar form.

Proposition 5.11 $\Delta : P(X \otimes Y) \to PX \otimes PY$ is natural in X, Y.

Proposition 5.12 The marginal map together with the trivial counitor defines a symmetric oplax monoidal functor $(P, \mathrm{id}_1, \Delta)$.

Proposition 5.13 (P, δ, E) is a symmetric opmonoidal monad.

The lax and oplax monoidal structure interact to give a bimonoidal structure. The following statements also follow just from the properties of joints and marginals.

Proposition 5.14 *P is a symmetric bilax monoidal functor.*

The main result then just follows as a corollary:

Theorem 5.15 *The Kantorovich monad is a symmetric bimonoidal monad, with monoidal structure given by the product joint, and opmonoidal structure given by the marginals.*

By Proposition 4.1, we therefore have:

Corollary 5.16 $\Delta_{X,Y} \circ \nabla_{X,Y} = \mathrm{id}_{PX \otimes PY}$. *Therefore, the inclusion ∇ of product measures into general joints, is an isometric embedding for the Kantorovich metric, and its image is a retract of the space of all joints.*

A Monoidal, opmonoidal and bimonoidal monads

We recall the definition of the different monoidal structures for a functor, for the case of braided (including symmetric) monoidal categories. For more results and more general definitions, we refer to [1].

Let $(\mathsf{C}, \otimes)$ and $(\mathsf{D}, \otimes)$ be braided monoidal categories.

Definition A.1 A *lax monoidal functor* $(\mathsf{C}, \otimes) \to (\mathsf{D}, \otimes)$ is a triple (F, η, ∇), such that:

(i) $F : C \to D$ is a functor;

(ii) The "unit" $\eta : 1_\mathsf{D} \to F(1_\mathsf{C})$ is a morphism of D;

(iii) The "multiplication" $\nabla : F(-) \otimes F(-) \Rightarrow F(- \otimes -)$ is a natural transformation of functors $\mathsf{C} \times \mathsf{C} \to \mathsf{D}$;

(iv) The following "associativity" diagram commutes for every X, Y, Z in C:

$$
\begin{array}{ccc}
(FX \otimes FY) \otimes FZ & \xrightarrow{\cong} & FX \otimes (FY \otimes FZ) \\
\downarrow{\scriptstyle \nabla_{X,Y} \otimes \mathrm{id}} & & \downarrow{\scriptstyle \mathrm{id} \otimes \nabla_{Y,Z}} \\
F(X \otimes Y) \otimes FZ & & FX \otimes F(Y \otimes Z) \\
\downarrow{\scriptstyle \nabla_{X \otimes Y, Z}} & & \downarrow{\scriptstyle \nabla_{X, Y \otimes Z}} \\
F((X \otimes Y) \otimes Z) & \xrightarrow{\cong} & F(X \otimes (Y \otimes Z))
\end{array}
$$

(v) The following "unitality" diagrams commute for every X in C:

$$
\begin{array}{ccc}
1_\mathsf{D} \otimes FX & \xrightarrow{\eta \otimes \mathrm{id}} & F(1_\mathsf{C}) \otimes FX \\
\downarrow{\scriptstyle \cong} & & \downarrow{\scriptstyle \nabla_{1_\mathsf{C}, X}} \\
FX & \xleftarrow{\cong} & F(1_\mathsf{C} \otimes X)
\end{array}
\qquad
\begin{array}{ccc}
FX \otimes 1_\mathsf{D} & \xrightarrow{\mathrm{id} \otimes \eta} & FX \otimes F(1_\mathsf{C}) \\
\downarrow{\scriptstyle \cong} & & \downarrow{\scriptstyle \nabla_{X, 1_\mathsf{C}}} \\
FX & \xleftarrow{\cong} & F(X \otimes 1_\mathsf{C})
\end{array}
$$

We say that (F, η, ∇) is also *braided*, or *symmetric* if C is symmetric, if in addition the multiplication commutes with the braiding:

$$
\begin{array}{ccc}
FX \otimes FY & \xrightarrow{\cong} & FY \otimes FX \\
\downarrow{\scriptstyle \nabla} & & \downarrow{\scriptstyle \nabla} \\
F(X \otimes Y) & \xrightarrow{\cong} & F(Y \otimes X)
\end{array}
$$

Definition A.2 Let (F, η_F, ∇_F) and (G, η_G, ∇_G) be lax monoidal functors $(\mathsf{C}, \otimes) \to (\mathsf{D}, \otimes)$. A *lax monoidal natural transformation*, or just *monoidal natural transformation* when it's clear from the context, is a natural transformation $\alpha : F \Rightarrow G$ which is compatible with the unit and multiplication map. In particular, the following diagrams must commute (for all $X, Y \in \mathsf{C}$):

$$
\begin{array}{ccc}
1_\mathsf{D} & \xrightarrow{\eta_F} & F(1_\mathsf{C}) \\
& {\scriptstyle \eta_G} \searrow & \downarrow{\scriptstyle \alpha_{1_\mathsf{C}}} \\
& & G(1_\mathsf{C})
\end{array}
\qquad
\begin{array}{ccc}
FX \otimes FY & \xrightarrow{\nabla_F} & F(X \otimes Y) \\
\downarrow{\scriptstyle \alpha_X \otimes \alpha_Y} & & \downarrow{\scriptstyle \alpha_{X \otimes Y}} \\
GX \otimes GY & \xrightarrow{\nabla_G} & G(X \otimes Y)
\end{array}
$$

14

Definition A.3 An *oplax monoidal functor* $(\mathsf{C}, \otimes) \to (\mathsf{D}, \otimes)$ is a triple (F, ϵ, Δ), such that:

(i) $F : C \to D$ is a functor;

(ii) The "counit" $\epsilon : F(1_\mathsf{C}) \to 1_\mathsf{D}$ is a morphism of D;

(iii) The "comultiplication" $\Delta : F(- \otimes -) \Rightarrow F(-) \otimes F(-)$ is a natural transformation of functors $\mathsf{C} \times \mathsf{C} \to \mathsf{D}$;

(iv) The following "coassociativity" diagram commutes for every X, Y, Z in C:

$$
\begin{array}{ccc}
F((X \otimes Y) \otimes Z) & \xrightarrow{\cong} & F(X \otimes (Y \otimes Z)) \\
\downarrow{\scriptstyle \Delta_{X \otimes Y, Z}} & & \downarrow{\scriptstyle \Delta_{X, Y \otimes Z}} \\
F(X \otimes Y) \otimes FZ & & FX \otimes F(Y \otimes Z) \\
\downarrow{\scriptstyle \Delta_{X,Y} \otimes \mathrm{id}} & & \downarrow{\scriptstyle \mathrm{id} \otimes \Delta_{Y,Z}} \\
(FX \otimes FY) \otimes FZ & \xrightarrow{\cong} & FX \otimes (FY \otimes FZ)
\end{array}
$$

(v) The following "counitality" diagrams commute for every X in C:

$$
\begin{array}{ccc}
F(1_\mathsf{C} \otimes X) & \xrightarrow{\Delta_{1_\mathsf{C}, X}} & F(1_\mathsf{C}) \otimes FX \\
\downarrow{\scriptstyle \cong} & & \downarrow{\scriptstyle \epsilon \otimes \mathrm{id}} \\
FX & \xleftarrow{\cong} & 1_\mathsf{D} \otimes FX
\end{array}
\qquad
\begin{array}{ccc}
F(X \otimes 1_\mathsf{C}) & \xrightarrow{\Delta_{X, 1_\mathsf{C}}} & FX \otimes F(1_\mathsf{C}) \\
\downarrow{\scriptstyle \cong} & & \downarrow{\scriptstyle \mathrm{id} \otimes \epsilon} \\
FX & \xleftarrow{\cong} & FX \otimes 1_\mathsf{D}
\end{array}
$$

We say that (F, ϵ, Δ) is also *braided*, or *symmetric* if C is symmetric, if in addition the comultiplication commutes with the braiding:

$$
\begin{array}{ccc}
F(X \otimes Y) & \xrightarrow{\cong} & F(Y \otimes X) \\
\downarrow{\scriptstyle \Delta} & & \downarrow{\scriptstyle \Delta} \\
FX \otimes FY & \xrightarrow{\cong} & FY \otimes FX
\end{array}
$$

Definition A.4 Let $(F, \epsilon_F, \Delta_F)$ and $(G, \epsilon_G, \Delta_G)$ be oplax monoidal functors $(\mathsf{C}, \otimes) \to (\mathsf{D}, \otimes)$. An *oplax monoidal natural transformation*, or just *monoidal natural transformation* when it's clear from the context, is a natural transformation $\alpha : F \Rightarrow G$ which is compatible with the counit and comultiplication map. In particular, the following diagrams must commute (for all $X, Y \in \mathsf{C}$):

$$
\begin{array}{ccc}
1_\mathsf{D} & \xleftarrow{\epsilon_F} & F(1_\mathsf{C}) \\
& {\scriptstyle \epsilon_G} \nwarrow & \downarrow{\scriptstyle \alpha_{1_\mathsf{C}}} \\
& & G(1_\mathsf{C})
\end{array}
\qquad
\begin{array}{ccc}
FX \otimes FY & \xleftarrow{\Delta_F} & F(X \otimes Y) \\
\downarrow{\scriptstyle \alpha_X \otimes \alpha_Y} & & \downarrow{\scriptstyle \alpha_{X \otimes Y}} \\
GX \otimes GY & \xleftarrow{\Delta_G} & G(X \otimes Y)
\end{array}
$$

Definition A.5 A *bilax monoidal functor* $(\mathsf{C}, \otimes) \to (\mathsf{D}, \otimes)$ is a "quintuplet" $(F, \eta, \nabla, \epsilon, \Delta)$ such that:

(i) $(F, \eta, \nabla) : (\mathsf{C}, \otimes) \to (\mathsf{D}, \otimes)$ is a lax monoidal functor;

(ii) $(F, \epsilon, \Delta) : (\mathsf{C}, \otimes) \to (\mathsf{D}, \otimes)$ is an oplax monoidal functor;

(iii) The following "bimonoidality" diagram commutes:

$$
\begin{array}{c}
F(W \otimes X) \otimes F(Y \otimes Z) \\
\end{array}
$$

with morphisms $\nabla_{W \otimes X, Y \otimes Z}$ and $\Delta_{W,X} \otimes \Delta_{Y,Z}$ to

$$F(W \otimes X \otimes Y \otimes Z) \qquad F(W) \otimes F(X) \otimes F(Y) \otimes F(Z)$$

$$\cong \downarrow \qquad\qquad \downarrow \cong \qquad\qquad (A.1)$$

$$F(W \otimes Y \otimes X \otimes Z) \qquad F(W) \otimes F(Y) \otimes F(X) \otimes F(Z)$$

$$\Delta_{W \otimes Y, X \otimes Z} \searrow \qquad \swarrow \nabla_{W,Y} \otimes \nabla_{X,Z}$$

$$F(W \otimes Y) \otimes F(X \otimes Z)$$

(iv) The following three "unit/counit" diagrams commute:

$$
\begin{array}{ccc}
1 \xrightarrow{\eta} F(1) & 1 \xrightarrow{\eta} F(1) \xrightarrow{\cong} F(1 \otimes 1) & 1 \xleftarrow{\epsilon} F(1) \xleftarrow{\cong} F(1 \otimes 1) \\
\quad \searrow \downarrow \epsilon & \cong \downarrow \qquad\qquad \downarrow \Delta_{1,1} & \cong \uparrow \qquad\qquad \uparrow \nabla_{1,1} \\
\qquad 1 & 1 \otimes 1 \xrightarrow{\eta \otimes \eta} F(1) \otimes F(1) & 1 \otimes 1 \xleftarrow{\epsilon \otimes \epsilon} F(1) \otimes F(1)
\end{array}
$$

Definition A.6 Let $(F, \epsilon_F, \Delta_F)$ and $(G, \epsilon_G, \Delta_G)$ be bilax monoidal functors $(\mathsf{C}, \otimes) \to (\mathsf{D}, \otimes)$. A *bilax monoidal natural transformation*, or just *monoidal natural transformation* when it's clear from the context, is a natural transformation $\alpha : F \Rightarrow G$ which is a lax and oplax natural transformation.

Definition A.7 Now, we define:

- A *monoidal monad* is a monad in the bicategory of monoidal categories, lax monoidal functors, and monoidal natural transformations;

- An *opmonoidal monad* is a monad in the bicategory of monoidal categories, oplax monoidal functors, and monoidal natural transformations;

- A *bimonoidal monad* is a monad in the bicategory of braided monoidal categories, bilax monoidal functors, and monoidal natural transformations.

In the third definition, we need the symmetry (or at least a braiding) in order to express the bimonoid equation that is part of the definition of bilax monoidal functor [1], even if the functor itself if not braided. If the functor is braided, we can define in addition:

- A *braided* (resp. *symmetric*) *monoidal monad* is a monad in the bicategory of braided (resp. symmetric) monoidal categories, braided lax monoidal functors, and monoidal natural transformations;

- An *braided* (resp. *symmetric*) *opmonoidal monad* is a monad in the bicategory of braided (resp. symmetric) monoidal categories, braided oplax monoidal functors, and monoidal natural transformations;

- A *braided* (resp. *symmetric*) *bimonoidal monad* is a monad in the bicategory of braided (resp. symmetric) monoidal categories, braided bilax monoidal functors, and monoidal natural transformations.

B   Proofs

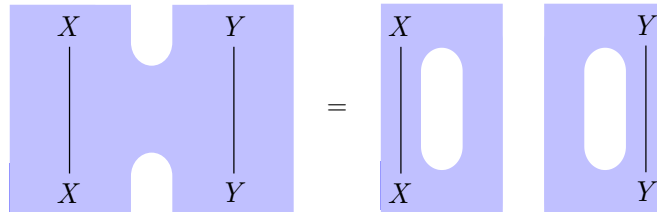Here are the detailed proofs of the statements in the main text.

B.1 Graphical proofs

Proof of Proposition 4.1 Let $X = Y = 1_{\mathsf{C}}$ in the bimonoidality diagram (1), and rename W to X and Z to Y for convenience. Then we get:
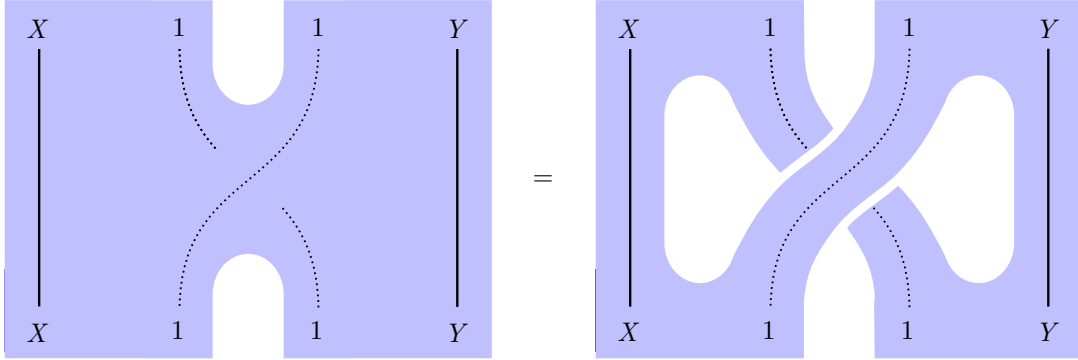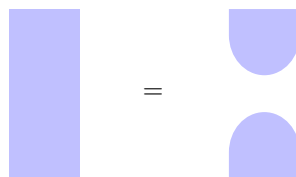

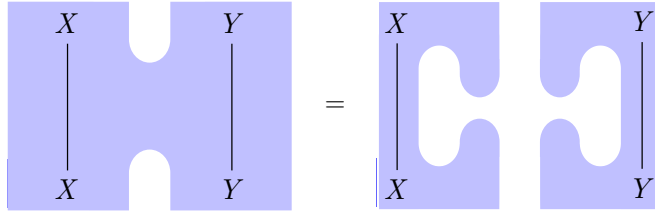Now since the braiding at $1 \otimes 1$ is just the identity, we can even simplify the condition to:
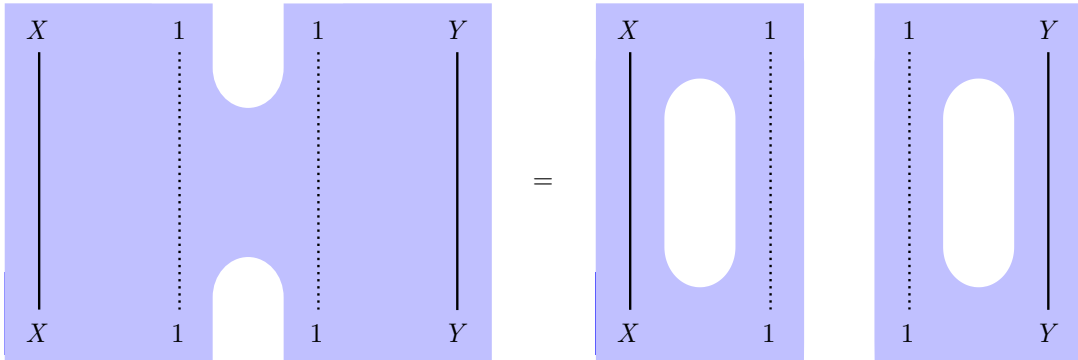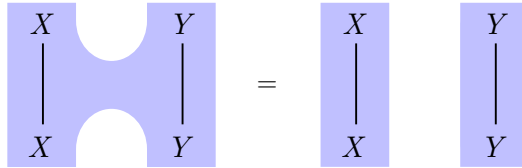

or more concisely:

We now notice that:

since both maps are just the identities at 1. This is the crucial step. We are left with:

$$
\vcenter{\hbox{(string diagram: X, Y top; X, Y bottom; cup/cap)}}
\;=\;
\vcenter{\hbox{(string diagram: X, Y top; X, Y bottom)}}
$$

which because of all the unit and counit conditions is equivalent to

$$
\vcenter{\hbox{(string diagram: X, Y top; X, Y bottom)}}
\;=\;
\vcenter{\hbox{(string diagram: X, Y top; X, Y bottom)}}
$$

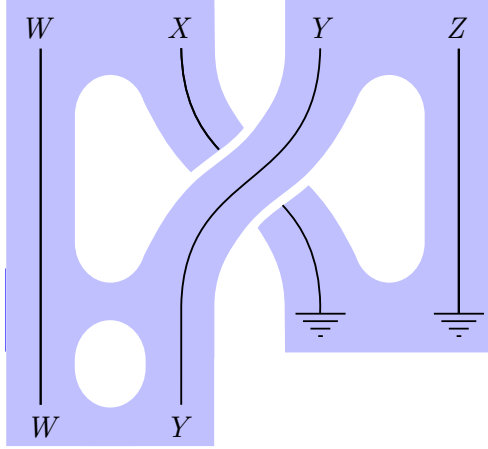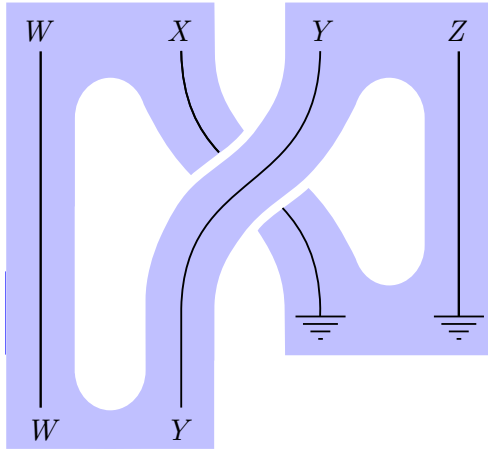i.e. equation (2).   □

Proof of Proposition 4.3 Consider the left side of (1) and forget X and Z using the unique maps to 1, and compose at the remaining $W \otimes Y$ with the left-hand side of (3). We get:
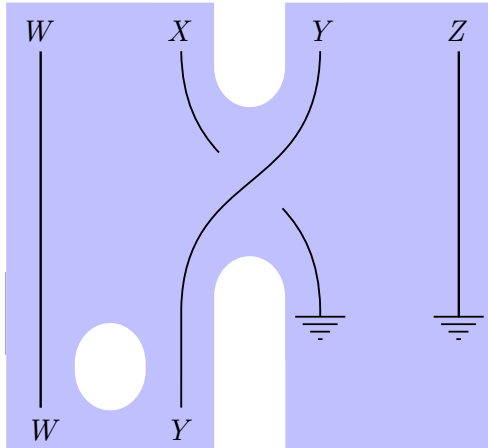
Applying (1) on the left and affinity of P on the right we get:

and applying (2) on the left we now get:

which, before applying the ground wire maps, is the right-hand side of (1), which is therefore equal to its left-hand side. Hence by Definition 4.2, W is independent of Y for any law in the form given in the hypothesis. For X and Z we can proceed analogously. □

B.2   Proof of equivalence of the notions of independence

Proof of Proposition 4.4 In $\mathsf{Prob}(\mathsf{C})$, f_1 and f_2 are independent in the sense of Franz with respect to the law $s : 1 \to PA$ if and only if there exists $h : A \to B_1 \times B_2$ such that the following diagram commutes:

$$\text{(B.1)}$$

where π_1 and π_2 are the projections of C, where the dotted arrows from 1, with a slight abuse of notation, denote Kleisli morphisms $(s : 1 \to PA,$ etcetera), and where r_1 and r_2 denote the resulting laws on B_1 and B_2.

Now suppose that such an h exists. By the universal property of the product, it must necessarily be equal to (f_1, f_2). Therefore $P(f_1, f_2) \circ s = r_1 \otimes_\nabla r_2 = \nabla \circ (r_1 \otimes r_2)$. Now using Proposition 4.1,

$$\nabla \circ \Delta \circ P(f_1, f_2) \circ s = \nabla \circ \Delta \circ \nabla \circ (r_1 \otimes r_2) = \nabla \circ (r_1 \otimes r_2) = P(f_1, f_2) \circ s,$$

so B_1 and B_2 are independent in the sense of Definition 4.2.

Conversely, suppose that $\nabla \circ \Delta \circ P(f_1, f_2) \circ s = P(f_1, f_2) \circ s$. Then we have

$$P(f_1, f_2) \circ s = \nabla \circ \Delta \circ P(f_1, f_2) \circ s = \nabla \circ (r_1, r_2) = r_1 \otimes_\nabla r_2,$$

as was to be shown. $\qquad\square$

B.3 Monoidal structure of the Kantorovich monad

In order to prove Proposition 5.4, first a useful result:

Proposition B.1 *Let $f : X \otimes Y \to \mathbb{R}$ be short. Let $p \in PX$. Then the function*

$$\left(\int_X f(x, -) \, dp(x) \right) : Y \to \mathbb{R}$$

is short as well.

Proof of Proposition B.1 First of all, $f : X \otimes Y \to \mathbb{R}$ being short means that for every $x, x' \in X, y, y' \in Y$:

$$|f(x, y) - f(x', y')| \le d(x, x') + d(y, y').$$

Now:

$$
\begin{aligned}
\left| \int_X f(x, y) \, dp(x) - \int_X f(x, y') \, dp(x) \right| \\
= \left| \int_X \big(f(x, y) - f(x, y') \big) \, dp(x) \right| \\
\le \int_X |f(x, y) - f(x, y')| \, dp(x) \\
\le \int_X \big(d(x, x) + d(y, y') \big) \, dp(x) \\
= \int_X d(y, y') \, dp(x) \\
= d(y, y').
\end{aligned}
$$

$\qquad\square$

Proof of Proposition 5.4 To prove that ∇ it is short, let $p, p' \in PX, q, q' \in PY$. Then

$$d\big(\nabla(p,q), \nabla(p',q')\big)$$

$$= d\big(p \otimes_\nabla q, p' \otimes_\nabla q'\big)$$

$$= \sup_{f:X\otimes Y\to\mathbb{R}} \int_{X\otimes Y} f(x,y)\, d(p\otimes_\nabla q - p'\otimes_\nabla q')(x,y)$$

$$= \sup_{f:X\otimes Y\to\mathbb{R}} \int_{X\otimes Y} f(x,y)\, d\big(p\otimes_\nabla q - p'\otimes_\nabla q + p'\otimes_\nabla q - p'\otimes_\nabla q'\big)(x,y)$$

$$= \sup_{f:X\otimes Y\to\mathbb{R}} \int_{X\otimes Y} f(x,y)\, d\big((p-p')\otimes q + p'\otimes_\nabla (q-q')\big)(x,y)$$

$$= \sup_{f:X\otimes Y\to\mathbb{R}} \int_X \left\{ \int_Y f(x,y)\, dq(y) \right\} d(p-p')(x)$$

$$\qquad + \int_Y \left\{ \int_X f(x,y)\, dp'(x) \right\} d(q-q')(y)$$

$$\leq \sup_{g:X\to\mathbb{R}} \int_X g(x)d(p-p')(x) + \sup_{h:Y\to\mathbb{R}} \int_Y h(y)d(q-q')(y)$$

$$= d(p,p') + d(q,q')$$

$$= d\big((p,q),(p',q')\big),$$

where by replacing the partial integral of f by g we have used Proposition B.1. □

Proof of Proposition 5.5 By symmetry, it suffices to show naturality in X. Let $f : X \to Z$. We need to show that this diagram commutes:

$$
\begin{array}{ccc}
PX \otimes PY & \xrightarrow{\nabla_{X,Y}} & P(X\otimes Y) \\
\downarrow{\scriptstyle f_*\otimes\mathrm{id}} & & \downarrow{\scriptstyle (f\otimes\mathrm{id})_*} \\
PZ \otimes PY & \xrightarrow{\nabla_{Z,Y}} & P(Z\otimes Y)
\end{array}
$$

Now let $p \in PX, q \in PY$, and $g : Z \otimes Y \to \mathbb{R}$. Then

$$\int_{Z\otimes Y} f(z,y)\, d\big((f\otimes\mathrm{id})_*\nabla_{X,Y}(p,q)\big)(z,y) = \int_{X\otimes Y} g(f(x),y)\, d(\nabla_{X,Y}(p,q))(x,y)$$

$$= \int_{X\otimes Y} g(f(x),y)\, dp(x)\, dq(y)$$

$$= \int_{Z\otimes Y} g(z,y)\, d(f_*p)(z)\, dq(y)$$

$$= \int_{Z\otimes Y} g(z,y)\, d\big((f_*p)\otimes q\big)(z,y)$$

$$= \int_{Z\otimes Y} g(z,y)\, d\big(\nabla_{Z,Y}\circ(f_*\otimes\mathrm{id})(p,q)\big)(z,y).$$

□

Proof of Proposition 5.6 Since both maps are natural, we only need to check the coherence diagrams. Since the unitor is just the identity at the terminal object, the unit diagrams commute. The associativity diagram

21

at each X, Y, Z

$$PX \otimes PY \otimes PZ \xrightarrow{\mathrm{id} \otimes \nabla_{Y,Z}} PX \otimes P(Y \otimes Z)$$

$$\downarrow \nabla_{X,Y} \otimes \mathrm{id} \qquad\qquad \downarrow \nabla_{X,Y \otimes Z}$$

$$P(X \otimes Y) \otimes PZ \xrightarrow{\nabla_{X \otimes Y, Z}} P(X \otimes Y \otimes Z)$$

gives for $(p, q, r) \in PX \otimes PY \otimes PZ$ on one path

$$(p, q, r) \mapsto (p \otimes_\nabla q, r) \mapsto (p \otimes_\nabla q) \otimes_\nabla r,$$

and on the other path

$$(p, q, r) \mapsto (p, q \otimes_\nabla r) \mapsto p \otimes_\nabla (q \otimes_\nabla r).$$

The product of probability distributions is now associative, as a simple calculation can show.

The symmetry condition is straightforward. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proof of Proposition 5.7 We know that $(P, \mathrm{id}_1, \nabla)$ is a lax monoidal functor. We need to check now that δ and E are monoidal natural transformations. Again we only need to show the commutativity with the multiplication, since the unitor is trivial. For $\delta : \mathrm{id}_{\mathsf{CMet}} \Rightarrow P$ we need to check that this diagram commute for each X, Y:

$$X \otimes Y \xrightarrow{\delta \otimes \delta} PX \otimes PY$$

$$\searrow^{\delta} \qquad \downarrow^{\nabla_{X,Y}}$$

$$P(X \otimes Y)$$

which means that for each $x \in X, y \in Y$ $\delta_x \otimes_\nabla \delta_y = \delta_{(x,y)}$, which is easy to check (the delta over the product is the product of the deltas). For $E : PP \Rightarrow P$ we first need to find the multiplication map $\nabla^2_{X,Y} : PPX \otimes PPY \to PP(X \otimes Y)$ (the unit is just twice the deltas, and the unit diagram again trivially commutes). This map is given by

$$P(PX) \otimes P(PY) \xrightarrow{\nabla_{PX,PY}} P(PX \otimes PY) \xrightarrow{(\nabla_{X,Y})_*} P(P(X \otimes Y))$$

and more explicitly, if $\mu \in PPX, \nu \in PPY$, and $f : P(X \times Y) \to \mathbb{R}$,

$$\int_{P(X \otimes Y)} f(r) \, d\big(\nabla^2_{X,Y}(\mu, \nu)\big)(r) = \int_{P(X \otimes Y)} f(r) \, d\big((\nabla_{X,Y})_* \circ \nabla_{PX,PY}(\mu, \nu)\big)(r)$$

$$= \int_{P(X \otimes Y)} f(r) \, d\big((\nabla_{X,Y})_*(\mu \otimes_\nabla \nu)\big)(r)$$

$$= \int_{PX \otimes PY} f(\nabla_{X,Y}(p, q)) \, d(\mu \otimes_\nabla \nu)(p, q)$$

$$= \int_{PX \otimes PY} f(p \otimes_\nabla q) \, d\mu(p) \, d\nu(q).$$

Now we have to check that this map makes this multiplication diagram commute:

$$PPX \otimes PPY \xrightarrow{E_X \otimes E_Y} PX \otimes PY$$

$$\downarrow^{\nabla^2_{X,Y}} \qquad\qquad \downarrow^{\nabla_{X,Y}}$$

$$PP(X \otimes Y) \xrightarrow{E_{X \otimes Y}} P(X \otimes Y)$$

Now let $\mu \in PPX, \nu \in PPY$, and $g : X \times Y \to \mathbb{R}$. We have, using the formula for ∇^2 found above,

$$\int_{X \otimes Y} g(x,y) \, d\big(\nabla_{X,Y} \circ (E_X, E_Y)(\mu, \nu)\big)(x,y) =$$

$$= \int_{X \otimes Y} g(x,y) \, d\big(\nabla_{X,Y}(E\mu, E\nu)\big)(x,y)$$

$$= \int_{X \otimes Y} g(x,y) \, d\big(E\mu \otimes_\nabla E\nu\big)(x,y)$$

$$= \int_{PX \otimes PY} \left\{ \int_{X \otimes Y} g(x,y) \, dp(x) \, dq(y) \right\} d\mu(p) \, d\nu(q)$$

$$= \int_{PX \otimes PY} \left\{ \int_{X \otimes Y} g(x,y) \, d(p \otimes_\nabla q)(x,y) \right\} d\mu(p) \, d\nu(q)$$

$$= \int_{P(X \times Y)} \left\{ \int_{X \otimes Y} g(x,y) \, dr(x,y) \right\} d(\nabla^2_{X,Y}(\mu, \nu))(r)$$

$$= \int_{X \otimes Y} g(x,y) \, d\big(E_{X \otimes Y} \circ \nabla^2_{X,Y}(\mu, \nu)\big)(x,y).$$

Therefore the diagram commutes, and (P, δ, E) is a monoidal monad. □

B.4 Opmonoidal structure of the Kantorovich monad

Just as in the case of joints, to prove the Proposition 5.10 we first prove the following useful result.

Proposition B.2 Let $f : X \to \mathbb{R}$ and $g : Y \to \mathbb{R}$ be short. Then $(f + g) : X \otimes Y \to \mathbb{R}$ given by $(x,y) \mapsto f(x) + g(y)$ is short.

Proof of Proposition B.2 Let $x, x' \in X$ and $y, y \in Y$. Then

$$|f(x) + g(y) - f(x') - f(y')| \le |f(x) - f(x')| + |g(y) - g(y')|$$

$$\le d(x, x') + d(y, y') = d\big((x,y), (x', y')\big).$$

□

Proof of Proposition 5.10 To prove that Δ is short, let $p, q \in P(X \otimes Y)$, and denote p_X, p_Y, q_X, q_Y their marginals. Then:

$$d\big(\Delta(p), \Delta(q)\big) = d\big((p_X, p_Y), (q_X, q_Y)\big) = d(p_X, q_X) + d(p_Y, q_Y)$$

$$= \sup_{f : X \to \mathbb{R}} \int_X f(x) \, d(p_X - q_X)(x) + \sup_{g : Y \to \mathbb{R}} \int_Y g(y) \, d(p_Y - q_Y)(y)$$

$$= \sup_{f : X \to \mathbb{R}} \int_{X \otimes Y} f(x) \, d(p - q)(x,y) + \sup_{g : Y \to \mathbb{R}} \int_{X \otimes Y} g(y) \, d(p - q)(x,y)$$

$$= \sup_{f : X \to \mathbb{R}} \sup_{g : Y \to \mathbb{R}} \int_{X \otimes Y} \Big(f(x) + g(y)\Big) \, d(p - q)(x,y)$$

$$\le \sup_{h : X \otimes Y \to \mathbb{R}} h(x,y) \, d(p - q)(x,y)$$

$$= d_{P(X \otimes Y)}(p, q),$$

where by replacing $f + g$ with h we have used Proposition B.2. □

Proof of Proposition 5.11 By symmetry, it suffices to show naturality in X. Let $f : X \to Z$. We need to show that this diagram commutes:

$$\begin{array}{ccc} P(X \otimes Y) & \xrightarrow{\Delta_{X,Y}} & PX \otimes PY \\ \downarrow{\scriptstyle (f \otimes \mathrm{id})_*} & & \downarrow{\scriptstyle f_* \otimes \mathrm{id}} \\ P(Z \otimes Y) & \xrightarrow{\Delta_{Z,Y}} & PZ \otimes PY \end{array}$$

Let now $p \in P(X \otimes Y)$. We have to prove that:

$$\Delta_{Z,Y} \circ (f \otimes \mathrm{id})_* p = (f_* \otimes \mathrm{id}) \circ \Delta_{X,Y}(p).$$

On one hand:

$$(f_* \otimes \mathrm{id}) \circ \Delta_{X,Y}(p) = (f_* \otimes \mathrm{id})(p_X, p_Y)$$
$$= (f_* p_X, p_Y).$$

On the other hand, let $h : Z \to \mathbb{R}$ and $g : Y \to \mathbb{R}$ be short. Then:

$$\int_Z h(z)\, d(((f \otimes \mathrm{id})_* p)_Z)(z) = \int_{Z \otimes Y} h(z)\, d((f \otimes \mathrm{id})_* p)(z, y)$$
$$= \int_{X \otimes Y} h(f(x))\, dp(x, y)$$
$$= \int_X h(f(x))\, dp_X(x)$$
$$= \int_Z h(z)\, d(f_* p_X)(x),$$

and:

$$\int_Y g(y)\, d(((f \otimes id)_* p)_Y)(y) = \int_{Z \otimes Y} g(y)\, d((f \otimes id)_* p)(z, y)$$
$$= \int_{X \otimes Y} g(y)\, dp(x, y)$$
$$= \int_Y g(y)\, dp_Y(y),$$

so the two components are again $(f_* p_X, p_Y)$. $\qquad\square$

Proof of Proposition 5.12 We already have naturality of the maps, and the counitor is trivial, we just have to check coassociativity. Namely, that the following diagrams commutes for each X, Y, Z:

$$\begin{array}{ccc} P(X \otimes Y \otimes Z) & \xrightarrow{\Delta_{X \otimes Y, Z}} & P(X \otimes Y) \otimes P(Z) \\ \downarrow{\scriptstyle \Delta_{X, Y \otimes Z}} & & \downarrow{\scriptstyle \Delta_{X \otimes Y} \otimes \mathrm{id}} \\ P(X) \otimes P(Y \otimes Z) & \xrightarrow[\mathrm{id} \otimes \Delta_{Y \otimes Z}]{} & P(X) \otimes P(Y) \otimes P(Z) \end{array}$$

Now given $p \in P(X \otimes Y \otimes Z)$, we get:

$$(\Delta_{X \otimes Y} \otimes \mathrm{id}) \circ \Delta_{X \otimes Y, Z}(p) = (\Delta_{X \otimes Y} \otimes \mathrm{id})(p_{XY}, p_Z) = (p_X, p_Y, p_Z),$$

and:

$$(\mathrm{id} \otimes \Delta_{Y \otimes Z}) \circ \Delta_{X, Y \otimes Z}(p) = (\mathrm{id} \otimes \Delta_{Y \otimes Z})(p_X, p_{YZ}) = (p_X, p_Y, p_Z),$$

since there is only one way of forming marginals.

The symmetry condition is again straightforward. □

Proof of Proposition 5.13 We know that $(P, \mathrm{id}_1, \Delta)$ is an oplax monoidal functor. We need to check now that δ and E are comonoidal natural transformations. Again we only need to show the commutativity with the comultiplication, since the counitor is trivial. For $\delta : \mathrm{id}_{\mathsf{CMet}} \Rightarrow P$ we need to check that this diagram commute for each X, Y:

$$
\begin{array}{ccc}
X \otimes Y & \xrightarrow{\ \delta\ } & P(X \otimes Y) \\
& \searrow{\scriptstyle \delta \otimes \delta} & \downarrow{\scriptstyle \Delta_{X,Y}} \\
& & PX \otimes PY
\end{array}
$$

which means that for each $x \in X, y \in Y$, $(\delta_{(x,y)})_X = \delta_x$ and $(\delta_{(x,y)})_Y = \delta_y$, which is again easy to check (the marginals of a delta are the deltas at the projections). For $E : PP \Rightarrow P$ we first need to find the comultiplication map $\Delta^2_{X,Y} : PP(X \otimes Y) \to PPX \otimes PPY$ (the unit is just twice the deltas, and the unit diagram again trivially commutes). This map is given by:

$$
P(P(X \otimes Y)) \xrightarrow{(\Delta_{XY})_*} P(PX \otimes PY) \xrightarrow{\Delta_{PX,PY}} P(PX) \otimes P(PY)
$$

and more explicitly, if $\mu \in P(P(X \otimes Y))$, and $f : PX \to \mathbb{R}$ and $g : PY \to \mathbb{R}$ are short:

$$
\int_{PX} f(p) \, d\big(((\Delta_{XY})_* \mu)_{PX}\big)(p) = \int_{PX \otimes PY} f(p) \, d\big(((\Delta_{XY})_* \mu)_{PX}\big)(p, q)
$$

$$
= \int_{P(X \otimes Y)} f(r_X) \, d\mu(r)
$$

since g only depends on PX, and analogously:

$$
\int_{PY} g(q) \, d\big(((\Delta_{XY})_* \mu)_{PY}\big)(q) = \int_{P(X \otimes Y)} f(r_Y) \, d\mu(r).
$$

We have to check that this map makes this multiplication diagram commute:

$$
\begin{array}{ccc}
PP(X \otimes Y) & \xrightarrow{\ E_{X \otimes Y}\ } & P(X \otimes Y) \\
\downarrow{\scriptstyle \Delta^2_{X,Y}} & & \downarrow{\scriptstyle \Delta_{X,Y}} \\
PPX \otimes PPY & \xrightarrow{\ E_X \otimes E_Y\ } & PX \otimes PY
\end{array}
$$

25

Now let $\mu \in P(P(X \otimes Y))$, and $f : X \to \mathbb{R}$ and $g : Y \to \mathbb{R}$ short. We have, using the formula for Δ^2 found above:

$$\int_X f(x) \, d\left((E_{X \otimes Y}\mu)_X\right)(x) = \int_{X \otimes Y} f(x) \, d\left(E_{X \otimes Y}\mu\right)(x, y)$$

$$= \int_{P(X \otimes Y)} \left\{ \int_{X \otimes Y} f(x) \, dr(x, y) \right\} d\mu(r)$$

$$= \int_{P(X \otimes Y)} \left\{ \int_X f(x) \, d(r_X)(x) \right\} d\mu(r)$$

$$= \int_{PX \otimes PY} \left\{ \int_X f(x) \, dp(x) \right\} d\left((\Delta_{XY})_*\mu\right)(p, q)$$

$$= \int_{PX} \left\{ \int_X f(x) \, dp(x) \right\} d\left(((\Delta_{XY})_*\mu)_{PX}\right)(p)$$

$$= \int_X f(x) \, d\left(E_X((\Delta_{XY})_*\mu)_{PX}\right)(x),$$

and analogously:

$$\int_Y g(y) \, d\left((E_{X \otimes Y}\mu)_Y\right)(y) = \int_Y f(y) \, d\left(E_Y((\Delta_{XY})_*\mu)_{PY}\right)(y),$$

which means:

$$\Delta_{X,Y} \circ E_{X \otimes Y}\mu = (E_X \otimes E_Y) \circ \Delta_{PX,PY}(\Delta_{XY})_*\mu$$

$$= (E_X \otimes E_Y) \circ (\Delta_{PX,PY} \circ (\Delta_{XY})_*)\mu$$

$$= (E_X \otimes E_Y) \circ \Delta_{X,Y}^2 \mu.$$

Therefore the diagram commutes, and (P, δ, E) is an opmonoidal monad. $\square$

B.5 *Bimonoidal structure of the Kantorovich monad*

Proof of Proposition 5.14 We already know that P is lax and oplax. We only need to check the compatibility diagrams between the two structures. The unit diagrams are trivial, because the unitors are trivial. The bimonoidality diagram:

$$
\begin{array}{ccc}
 & P(W \otimes X) \otimes P(Y \otimes Z) & \\
{}^{\nabla_{W \otimes X, Y \otimes Z}}\swarrow & & \searrow^{\Delta_{W,X} \otimes \Delta_{Y,Z}} \\
P(W \otimes X \otimes Y \otimes Z) & & P(W) \otimes P(X) \otimes P(Y) \otimes P(Z) \\
\cong \downarrow & & \downarrow \cong \\
P(W \otimes Y \otimes X \otimes Z) & & P(W) \otimes P(Y) \otimes P(X) \otimes P(Z) \\
 \searrow^{\Delta_{W \otimes Y, X \otimes Z}} & & \swarrow_{\nabla_{W,Y} \otimes \nabla_{X,Z}} \\
 & P(W \otimes Y) \otimes P(X \otimes Z) &
\end{array}
$$

says that given $p \in P(W \otimes X), q \in P(Y \otimes Z)$:

$$\Delta_{W \otimes Y, X \otimes Z} \circ \nabla_{W \otimes X, Y \otimes Z}(p, q) = (\nabla_{W,Y} \otimes \nabla_{X,Z}) \circ (\Delta_{W,X} \otimes \Delta_{Y,Z})(p, q)$$

Now on one hand:

$$(\nabla_{W,Y} \otimes \nabla_{X,Z}) \circ (\Delta_{W,X} \otimes \Delta_{Y,Z})(p,q) = (\nabla_{W,Y} \otimes \nabla_{X,Z})(p_W, p_X, q_Y, q_Z)$$

$$= (p_W \otimes_\nabla q_Y, p_X \otimes_\nabla q_Z).$$

On the other hand:

$$\Delta_{W \otimes Y, X \otimes Z} \circ \nabla_{W \otimes X, Y \otimes Z}(p,q) = \Delta_{W \otimes Y, X \otimes Z}(p \otimes_\nabla q).$$

The marginal of $p \otimes q$ on $W \otimes Y$ is, by Fubini's theorem, let $f : W \otimes Y \to \mathbb{R}$:

$$\int_{W \otimes Y} f(w,y) \, d((p \otimes_\nabla q)_{WY})(w,y) = \int_{W \otimes X \otimes Y \otimes Z} f(w,y) \, d(p \otimes_\nabla q)(w,x,y,z)$$

$$= \int_{W \otimes X \otimes Y \otimes Z} f(w,y) \, dp(w,x) \, dq(y,z)$$

$$= \int_{W \otimes Y} f(w,y) \, dp_W(w) \, dq_Y(y)$$

$$= \int_{W \otimes Y} f(w,y) \, d(p_W \otimes_\nabla q_Y)(w,y),$$

and similarly the marginal on $X \otimes Z$ is given by $p_X \otimes_\nabla q_Z$. In other words, if the pairs are independent, the components from different pairs are also independent. It follows that P is bilax monoidal. $\qquad \square$

Acknowledgement

The authors would like to thank the anonymous reviewers for the very helpful comments.

References

[1] Aguiar, M. and S. Mahajan, "Monoidal Functors, Species and Hopf Algebras," AMS, 2010.

[2] Basso, G., *A Hitchhiker's guide to Wasserstein distances* (2015), Available at: http://n.ethz.ch.

[3] Bogachev, V. I., "Measure Theory," Springer, 2000.

[4] Edgar, G. A., "Integral, Probability, and Fractal Measures," Springer, 1998.

[5] Franz, U., *What is stochastic independence?*, Non-commutativity, infinite-dimensionality and probability at the crossroads. Proceedings of the RIMS workshop on infinite-dimensional analysis and quantum probability. (2001).

[6] Fritz, T. and P. Perrone, *A Probability Monad as the Colimit of Finite Powers* (2017), submitted. arXiv:1712.05363.

[7] Giry, M., *A Categorical Approach to Probability Theory*, in: *Categorical aspects of topology and analysis*, Lecture Notes in Mathematics **915**, 1982 .

[8] Golbeaut-Larrecq, J., S. Lasota and D. Nowak, *Logical relations for monadic types*, Mathematical Structures in Computer Science **18** (2008), pp. 1169–1217.

[9] Jacobs, B., *From Probability Monads to Commutative Effectuses*, Journ. of Logical and Algebraic Methods in Programming (2017), in press. Available at http://www.cs.ru.nl/B.Jacobs/PAPERS/probability-monads.pdf.

[10] Jones, C. and J. D. Plotkin, *A Probabilistic Powerdomain of Evaluations*, Proceedings of the Fourth Annual Symposium of Logics in Computer Science (1989).

[11] Kock, A., *Strong functors and monoidal monads*, Arch. Math. (Basel) **23** (1972), pp. 113–120.

[12] Lawvere, W., *Metric spaces, generalized logic and closed categories*, Rendiconti del seminario matematico e fisico di Milano **43** (1973).
URL http://www.tac.mta.ca/tac/reprints/articles/1/tr1abs.html

[13] Leinster, T., *Monoidal categories with projections* (2016), *n*-Category Café blog post, https://golem.ph.utexas.edu/category/2011/05/an_operadic_introduction_to_en.html.

[14] Melliès, P.-A., *Functorial Boxes in String Diagrams*, International Workshop on Computer Science Logic (2006).

[15] Pearl, J. and A. Paz, *Graphoids: a graph-based logic for reasoning about relevance relations*, UCLA Computer Science Dept., Technical Report 850038 (1985).

[16] Plotkin, G. D. and J. Power, *Notion of Computation Determine Monads* (2002).

[17] Sato, T., *The Giry Monad is not strong for the canonical symmetric monoidal closed structure on Meas* (2016), arXiv:1612.05939.
URL http://arxiv.org/abs/1612.05939

[18] Simpson, A., *Category-theoretic structure for independence and conditional independence*, Electronic Notes in Theoretical Computer Science **336** (2018), pp. 281–297.

[19] van Breugel, F., *The Metric Monad for Probabilistic Nondeterminism* (2005), Available at http://www.cse.yorku.ca.

[20] Villani, C., "Optimal transport: old and new," Grundlehren der mathematischen Wissenschaften **338**, Springer, 2009.

[21] Willerton, S., *A diagrammatic approach to Hopf monads*, Arabian Journal of Science and Engineering C **33** (2008).