Weakly affine monads

Tobias Fritz
Department of Mathematics, University of Innsbruck, AT

Paolo Perrone
Department of Computer Science, University of Oxford, UK

Fabio Gadducci
Department of Computer Science, University of Pisa, Pisa, IT

Davide Trotta
Department of Computer Science, University of Pisa, Pisa, IT

Abstract

Introduced in the 1990s in the context of the algebraic approach to graph rewriting, gs-monoidal categories are symmetric monoidal categories where each object is equipped with the structure of a commutative comonoid. They arise for example as Kleisli categories of commutative monads on cartesian categories, and as such they provide a general framework for effectful computation. Recently proposed in the context of categorical probability, Markov categories are gs-monoidal categories where the monoidal unit is also terminal, and they arise for example as Kleisli categories of commutative *affine* monads, where affine means that the monad preserves the monoidal unit.

The aim of this paper is to study a new condition on the gs-monoidal structure, resulting in the concept of *weakly Markov categories*, which is intermediate between gs-monoidal categories and Markov ones. In a weakly Markov category, the morphisms to the monoidal unit are not necessarily unique, but form a group. As we show, these categories exhibit a rich theory of conditional independence for morphisms, generalising the known theory for Markov categories. We also introduce the corresponding notion for commutative monads, which we call weakly affine, and for which we give two equivalent characterisations.

The paper argues that these monads are relevant to the study of categorical probability. A case at hand is the monad of finite non-zero measures, which is weakly affine but not affine. Such structures allow to investigate probability without normalisation within an elegant categorical framework.

2012 ACM Subject Classification Theory of Computation –> Models of computations

Keywords and phrases String diagrams, gs-monoidal and Markov categories, categorical probability, affine monads.

Digital Object Identifier 10.4230/LIPIcs...

Funding *Tobias Fritz*: FWF P 35992-N
Fabio Gadducci: MIUR PRIN 2017FTXR “IT-MaTTerS”.
Davide Trotta: MIUR PRIN 2017FTXR “IT-MaTTerS”.

1 Introduction

The idea of *gs-monoidal categories*, symmetric monoidal categories equipped with copy and discard morphisms making every object a comonoid, was first introduced in the context of algebraic approaches to term graph rewriting [3], and then developed in a series of papers [4, 5, 6]. Two decades later, similar structures have been rediscovered independently in the context of categorical probability theory, in particular in [1] and [9], under the names of *copy-discard (CD) categories* and *Markov categories*. While “CD-categories” and “gs-monoidal categories” are synonyms, Markov categories have the additional condition that the monoidal unit is the terminal object (i.e. every morphism commutes with the discard maps), a condition corresponding to normalisation of probability. See [13, Remark 2.2] for a more detailed history of these ideas.

A canonical way of obtaining a gs-monoidal category is as the Kleisli category of a commutative monad on a cartesian monoidal category. As argued in [17], commutative monads can be seen as generalising theories of distributions of some kind, and the fact that their Kleisli categories are gs-monoidal can be seen as the correspondence between distributions and (possibly unnormalised) probability theory. In particular, when the monad is affine (i.e. it preserves the monoidal unit [16, 14]), the Kleisli category is Markov – this can be seen as the correspondence between normalised distributions and probability theory.

In this work we introduce and study an intermediate notion between gs-monoidal and Markov categories, which we call *weakly Markov categories*. These are defined as gs-monoidal categories where for every object its morphisms to the monoidal unit form a group (Definition 3.2). Weakly Markov categories can be interpreted intuitively as gs-monoidal categories where each morphism is discardable up to an invertible normalisation (see Proposition 3.4 for the precise mathematical statement). The choice of the name is due to the fact that every Markov category is (trivially) weakly Markov.

In parallel to weakly Markov categories we also introduce *weakly affine monads*, which are commutative monads on cartesian monoidal categories preserving the (internal) group structure of the terminal object (Definition 3.5). As a particular concrete example of relevance to probability and measure theory, we consider the monad of finite non-zero measures on **Set** (Example 3.7), and we use it as a running example in the rest of the work. As we show (see Proposition 3.6), a commutative monad on a cartesian monoidal category is weakly affine if and only if its Kleisli category is weakly Markov, analogously to what happens with affine monads and Markov categories.

Markov categories come equipped with a notion of *conditional independence*, which has been one of the main motivations for their use in categorical probability and statistics [1, 9, 12]. It is noteworthy that a notion of conditional independence can also be given for any gs-monoidal category. As we show, for weakly Markov categories it has convenient properties which can be considered "up-to-normalisation" versions of their corresponding Markov-categorical counterpart. These concepts allow us to provide an equivalent condition for weak affinity of a monad, namely a pullback condition on the associativity diagram of the structural morphisms $c_{X,Y} : TX \times TY \to T(X \times Y)$ (Theorem 4.7), widely generalising the elementary statement that a monoid is a group if and only if its associativity diagram is a pullback (Proposition 2.1). As such, we believe that weak affine monads are relevant to the study of categorical probability, as they allow to investigate probability without normalisation within an elegant categorical framework.

Outline.

In Section 2 we review the main structures used in this work, in particular group and monoid objects, gs-monoidal and Markov categories, and their interaction with commutative monads.

In Section 3 we define the main original concepts, namely weakly Markov categories and weakly affine monads. We study their relationship and we prove that a commutative monad on a cartesian monoidal category is weakly affine if and only if its Kleisli category is weakly Markov (Proposition 3.6). We then turn to concrete examples using finite measures and group actions (Section 3.3).

In Section 4 we extend the concept of conditional independence from Markov categories to general gs-monoidal categories. We specialise to the weakly Markov case and show that the situation is then similar to what happens in Markov categories, but in a certain precise sense only up to normalisation. We use this formalism to equivalently reformulate weak affinity in terms of a pullback condition (Theorem 4.7). Together with the newly introduced concepts, this result can be considered the main outcome of our work.

Finally, in the concluding Section 5, we pose further questions, such as when we can iterate the construction of weakly Markov categories by means of weakly affine monads, and the relation to strongly affine monads in the sense of Jacobs [15].

2 Background

In this section, we develop some relevant background material for later reference. To begin, the following categorical characterisation of groups will be useful to keep in mind.

▶ **Proposition 2.1.** *A monoid (M, m, e) in* **Set** *is a group if and only if the associativity square*

$$\begin{array}{ccc} M \times M \times M & \xrightarrow{m \times \mathrm{id}} & M \times M \\ \downarrow{\scriptstyle \mathrm{id} \times m} & & \downarrow{\scriptstyle m} \\ M \times M & \xrightarrow{\quad m \quad} & M \end{array} \tag{1}$$

is a pullback.

Proof. The square (1) is a pullback of sets if and only if given $a, g, h, c \in M$ such that $ag = hc$, there exists a unique $b \in M$ such that $g = bc$ and $h = ab$. First, suppose that G is a group. Then the only possible choice of b is

$$b = a^{-1}h = gc^{-1}$$

which is unique by uniqueness of inverses.

Conversely, suppose that (1) is a pullback. We can set $g, h = e$ and $c = a$ so that $ae = ea = a$. Instantiating the pullback property on these elements gives b such that $ab = e$ and $ba = e$, that is, $b = a^{-1}$. ◀

Proposition 2.1 holds generally for a monoid object in a cartesian monoidal category, where the element-wise proof still applies thanks to the following standard observation.

▶ **Remark 2.2.** Given an object M in a cartesian monoidal category $\mathcal{D}$, there is a bijection between internal monoid structures on M and monoid structures on every hom-set $\mathcal{D}(X, M)$ such that pre-composition with any $f : X \to Y$ defines a monoid homomorphism

$$\mathcal{D}(Y, M) \longrightarrow \mathcal{D}(X, M).$$

The proof is straightforward by the Yoneda lemma. It follows that Proposition 2.1 holds for internal monoids in cartesian monoidal categories in general.

For the consideration of categorical probability, we now recall the simplest version of a commutative monad of measures. It works with measures taking values in any semiring instead of $[0, \infty)$ (see e.g. [7, Section 5.1]), but we restrict to the case of $[0, \infty)$ for simplicity.

▶ **Definition 2.3.** Let X be a set. Denote by MX the set of *finitely supported measures on* X, i.e. the functions $m : X \to [0, \infty)$ that are zero for all but a finite number of $x \in X$. Given a function $f : X \to Y$, denote by $Mf : MX \to MY$ the function sending $m \in MX$ to the assignment

$$(Mf)(m) \; : \; y \longmapsto \sum_{x \in f^{-1}(y)} p(x).$$

This makes M into a functor, and even a monad with the unit and multiplication maps

$$\begin{aligned} X &\xrightarrow{\delta} MX \\ x &\longmapsto \delta_x, \end{aligned} \qquad\qquad \begin{aligned} MMX &\xrightarrow{E} MX \\ \xi &\longmapsto E\xi, \end{aligned}$$

where

$$\delta_x(x') = \begin{cases} 1 & x = x', \\ 0 & x \neq x', \end{cases} \qquad\qquad (E\xi)(x) = \sum_{m \in MX} \xi(m)\, m(x).$$

Call M the *measure monad* on **Set**.

Denote also by $DX \subseteq MX$ the subset of *probability measures*, i.e. those finitely supported $p : X \to [0, \infty)$ such that

$$\sum_{x \in X} p(x) = 1.$$

D forms a sub-monad of M called the *distribution monad*.

It is known that M is a commutative monad [7]. The corresponding lax monoidal structure

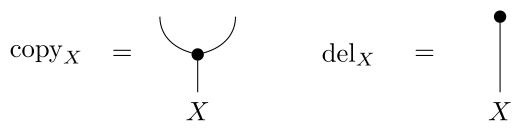
$$MX \times MY \xrightarrow{c} M(X \times Y)$$

is exactly the formation of product measures given by $c(m, m')(x, y) = m(x)m'(y)$. Also D is a commutative monad with the induced lax monoidal structure, since the product of probability measures is again a probability measure.

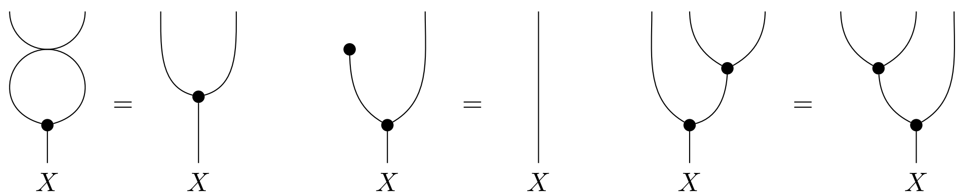
2.1 GS-monoidal and Markov categories

We recall here the basic definitions adopting the graphical formalism of string diagrams, referring to [18] for some background on various notions of monoidal categories and their associated diagrammatic calculus.

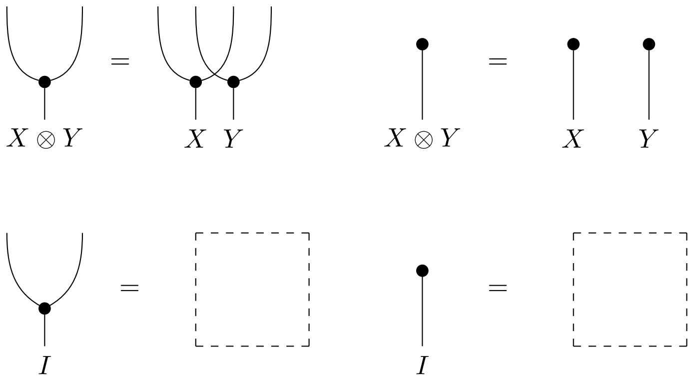
▶ **Definition 2.4.** A **gs-monoidal category** is a symmetric monoidal category $(\mathcal{C}, \otimes, I)$ with a commutative comonoid structure on each object X consisting of a comultiplication and a counit



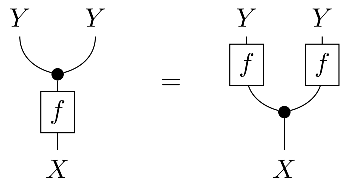
which satisfy the commutative comonoid equations:



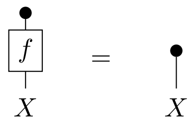
These comonoid structures must be multiplicative with respect to the monoidal structure



▶ **Definition 2.5.** A morphism $f : X \to Y$ in a gs-monoidal category is called **copyable** or **functional** if



It is called **discardable** or **full** if



▶ **Example 2.6.** The category **Rel** of sets and relations with the monoidal operation $\otimes : \mathbf{Rel} \times \mathbf{Rel} \to \mathbf{Rel}$ given by the direct product of sets is a gs-monoidal category [6]. In this gs-monoidal category, the copyable arrows are precisely the partial functions, and the discardable arrows are the total relations.

▶ **Remark 2.7.** It is well-known that if every morphism is copyable and discardable, or equivalently if the copy and discard maps are natural, then the monoidal product is the categorical product, and thus the category is cartesian monoidal [8]. In other words, the following conditions are equivalent for a gs-monoidal category $\mathcal{C}$

- $\mathcal{C}$ is cartesian monoidal;
- every morphism is copyable and discardable;
- the copy and discard maps are natural.

In recent work [10] it has been shown that gs-monoidal categories naturally arise in several ways, such as Kleisli categories of commutative monads or span categories. In the following proposition, we recall the result regarding Kleisli categories.

▶ **Proposition 2.8.** *Let T be a commutative monad on a cartesian monoidal category $\mathcal{D}$. Then its Kleisli category Kl_T is canonically a gs-monoidal category with the copy and discard structure induced by that of $\mathcal{D}$.*

▶ **Example 2.9.** The Kleisli categories of the monads M and D of Definition 2.3 are gs-monoidal. We can write their Kleisli categories concretely as follows

- a morphism $k : X \to Y$ of Kl_M is a *matrix* with rows indexed by Y and columns indexed by X, and non-negative entries $k(y|x)$ such that for each $x \in X$, the number $k(y|x)$ is non-zero only for finitely many x;

- a morphism $k : X \to Y$ of Kl_D is a morphism of Kl_M such that moreover, for all $x \in X$, the sum of each column satisfies

$$\sum_{y \in Y} k(y|x) = 1$$

  If X and Y are finite, such a matrix is called a *stochastic matrix*.

In both categories, identities are identity matrices, composition is matrix composition, monoidal structure is the cartesian product on objects and the Kronecker product on matrices, and the copy and discard maps are the images of the standard copy and discard maps on **Set** under the Kleisli inclusion functor.

Nowadays, *Markov categories* [9] represent one of the more interesting specialisations of the notion of gs-monoidal category. Based on the interpretation of their arrows as generalised Markov kernels, Markov categories are considered the foundation for a categorical approach to probability theory.

▶ **Definition 2.10.** A gs-monoidal category is said to be a **Markov category** if any (hence all) of the following equivalent conditions are satisfied

- the monoidal unit is terminal;
- the discard maps are natural;
- every morphism is discardable.

We recall from [16, 14] the notion of *affine monad*.

▶ **Definition 2.11.** A monad T on a cartesian monoidal category is called **affine** if $T1 \cong 1$.

It was observed in [9, Corollary 3.2] that if the monad preserves the terminal object, then every arrow of the Kleisli category is discardable, and this makes the Kleisli category into a Markov category. Since the converse is easy to see, we have the following addendum to Proposition 2.8.

▶ **Proposition 2.12.** *Let T be a commutative monad on a cartesian monoidal category $\mathcal{D}$. Then Kl_T is Markov if and only if T is affine.*

▶ **Example 2.13.** The distribution monad D of Definition 2.3 is affine, and so its Kleisli category (Example 2.9) is a Markov category. It is one of the simplest examples of categories of relevance for categorical probability.

The measure monad M is not affine, as it is easy to see that $M1 \cong [0, \infty)$, and so its Kleisli category is not Markov.

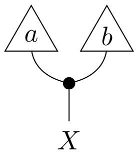
3 Weakly Markov categories and weakly affine monads

In this section, we introduce an intermediate level between gs-monoidal and Markov called *weakly Markov*, and its corresponding notion for monads, which we call *weakly affine*.

3.1 The monoid of effects

In a gs-monoidal category $\mathcal{C}$ we call a *state* a morphism from the monoidal unit $p : I \to X$, and *effect* a morphism to the monoidal unit $a : X \to I$. As is standard convention, we represent such morphisms as triangles as follows

Effects, i.e. elements of the set $\mathcal{C}(X, I)$, form canonically a commutative monoid as follows: the monoidal unit is the discard map $X \to I$, and given $a, b : X \to I$, their product ab is given by copying[1]



If a morphism $f : X \to Y$ is copyable and discardable, the pre-composition with f induces a morphism of monoids $\mathcal{C}(Y, I) \to \mathcal{C}(X, I)$.

▶ **Remark 3.1.** The monoidal unit I of a monoidal category is canonically a monoid object via the coherence isomorphisms $I \otimes I \cong I$ and $I \cong I$. However, in a general (i.e. not necessarily cartesian) gs-monoidal category $\mathcal{C}$, the monoid structure on $\mathcal{C}(X, I)$ is not, as in Remark 2.2, coming from considering the presheaf represented by I. Indeed, in order for Remark 2.2 to hold, we would need that every pre-composition is a morphism of monoids. As remarked above, this fails in general unless all morphisms are copyable and discardable (i.e. if $\mathcal{C}$ is not cartesian monoidal).

Let us now consider the case where the gs-monoidal structure comes from a commutative monad on a cartesian monoidal category $\mathcal{D}$. In this case, the monoid structure on Kleisli morphisms $X \to 1$ does come from the canonical internal monoid structure on $T1$ (and from the one on 1) in $\mathcal{D}$. Indeed, $T1$ is a monoid object with the following unit and multiplication [17, Section 10]

$$1 \xrightarrow{\eta} T1, \qquad T1 \times T1 \xrightarrow{c_{1,1}} T(1 \times 1) \xrightarrow{\cong} T1.$$

For example, for the monad of measures M, we obtain $M1 = [0, \infty)$ with its usual multiplication. The resulting monoid structure on Kleisli morphisms $X \to 1$ is now given as follows. The unit is given by

$$X \xrightarrow{\mathrm{del}_X} 1 \xrightarrow{\eta} T1,$$

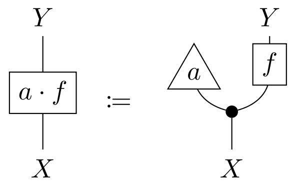
and the multiplication of Kleisli morphisms $f, g : X \to 1$ represented by $f^\sharp, g^\sharp : X \to T1$ is the Kleisli morphism represented by

$$X \xrightarrow{\mathrm{copy}_X} X \times X \xrightarrow{f^\sharp \times g^\sharp} T1 \times T1 \xrightarrow{c_{1,1}} T(1 \times 1) \xrightarrow{\cong} T1.$$

For the monad of measures M, Kleisli morphisms $X \to 1$ are represented by functions $X \to [0, \infty)$, and this description shows that their product is the point-wise product.

For a general $\mathcal{C}$, the commutative monoid $\mathcal{C}(X, I)$ acts on the set $\mathcal{C}(X, Y)$: given $a : X \to I$ and $f : X \to Y$, the resulting $a \cdot f$ is given as follows

[1] See also e.g. the $\odot$ product in [2, Proposition 3.10].



It is straightforward to see that this indeed amounts to an action of the monoid $\mathcal{C}(X, I)$ on the set $\mathcal{C}(X, Y)$. For the monad of measures M, this action is given by point-wise rescaling.

Moreover, for a general $\mathcal{C}$ the operation

$$\begin{aligned} \mathcal{C}(X, Y) \times \mathcal{C}(X, Z) &\longrightarrow \mathcal{C}(X, Y \otimes Z) \\ (f, g) &\longmapsto f \cdot g \coloneqq (f \otimes g) \circ \mathrm{copy}_X \end{aligned}$$

commutes with this action in each variable (separately).

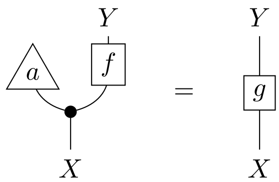
3.2 Main definitions

▶ **Definition 3.2.** A gs-monoidal category $\mathcal{C}$ is called **weakly Markov** if for every object X, the monoid $\mathcal{C}(X, I)$ is a group.

Every Markov category is weakly Markov: for every object X, the monoid $\mathcal{C}(X, I)$ is the trivial group.

▶ **Definition 3.3.** Given two parallel morphisms $f, g : X \to Y$ in a weakly Markov category $\mathcal{C}$, we say that f and g are called **equivalent**, denoted $f \sim g$, if they lie in the same orbit for the action of $\mathcal{C}(X, I)$, i.e. if there is $a \in \mathcal{C}(X, I)$ such that $a \cdot f = g$.

Note that if $a \cdot f = g$ for some a, then a is unique. This can be seen by discarding Y in the following diagram



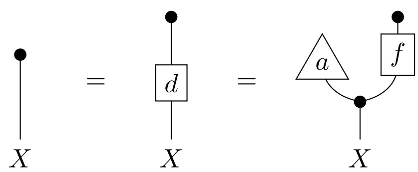
In other words, the action of $\mathcal{C}(X, I)$ on $\mathcal{C}(X, Y)$ is free, i.e. it has trivial stabilisers.

For the next statement, let us first call the *mass* of a morphism $f : X \to Y$ in a gs-monoidal category $\mathcal{C}$ the morphism $m_f \coloneqq \mathrm{del}_Y \circ f : X \to I$. Note that f is discardable if and only if $m_f = \mathrm{del}_X$, i.e. if its mass is the unit of the monoid $\mathcal{C}(X, I)$.

▶ **Proposition 3.4.** *Every morphism $f : X \to Y$ in a weakly Markov category is equivalent to a unique discardable morphism.*

We call the discardable morphism the *normalisation* of f and denote it by $n_f : X \to Y$.

Proof. Consider the mass m_f, and denote its group inverse by m_f^{-1}. The morphism $n_f \coloneqq m_f^{-1} \cdot f$ is discardable and equivalent to f. Suppose now that $d : X \to Y$ is discardable and equivalent to f, i.e. there exists $a : X \to I$ such that $d = a \cdot f$. Since d is discardable



which means that $a = m_f^{-1}$, i.e. $d = n_f$. ◀

In other words, every morphism f can be written as its mass times its normalisation.

Let us now look at the Kleisli case.

▶ **Definition 3.5.** A commutative monad T on a cartesian monoidal category is called **weakly affine** if $T1$ with its canonical internal commutative monoid structure is a group.

This choice of terminology is motivated by the following proposition, which can be seen as a "weakly" version of Proposition 2.12.

▶ **Proposition 3.6.** *Let $\mathcal{D}$ be a cartesian monoidal category and T a commutative monad on $\mathcal{D}$. Then the Kleisli category of T is weakly Markov if and only if T is weakly affine.*

Proof. First, suppose that $T1$ is an internal group, and denote by $\iota : T1 \to T1$ its inversion map. The inverse of a Kleisli morphism $a : X \to 1$ in $\mathrm{Kl}_T(X, 1)$ represented by $a^\sharp : X \to T1$ is represented by $\iota \circ a^\sharp$: indeed, the following diagram in $\mathcal{D}$ commutes

$$\begin{tikzcd}
X \arrow[r, "\mathrm{copy}_X"] \arrow[d, "a^\sharp"'] \arrow[dd, bend right=60, "\mathrm{del}_X"'] & X \times X \arrow[d, "a^\sharp \times a^\sharp"'] \arrow[dr, "a^\sharp \times (\iota \circ a^\sharp)"] & & \\
T1 \arrow[r, "\mathrm{copy}_{T1}"'] \arrow[d, "\mathrm{del}_{T1}"'] & T1 \times T1 \arrow[r, "\mathrm{id} \times \iota"'] & T1 \times T1 \arrow[r, "c_{1,1}"] & T(1 \times 1) \arrow[d, "\cong"] \\
1 \arrow[rrr, "\eta"] & & & T1
\end{tikzcd}$$

where the bottom rectangle commutes since ι is the inversion map for $T1$. The analogous diagram with $\iota \times \mathrm{id}$ in place of $\mathrm{id} \times \iota$ similarly commutes.

Conversely, suppose that for every X, the monoid structure on $\mathrm{Kl}_T(X, 1)$ has inverses. Then in particular we can take $X = T1$, and the inverse of the Kleisli morphism $\mathrm{id} : T1 \to T1$ is an inversion map for $T1$. ◀

This result can also be thought of in terms of the Yoneda embedding, via Remark 2.2: since the Yoneda embedding preserves and reflects pullbacks (and all limits), the associativity square for $T1$ is a pullback in $\mathcal{D}$ if and only if the associativity squares of all the monoids $\mathcal{D}(X, T1)$ are pullbacks. Note that Remark 2.2 applies since we are assuming that $\mathcal{D}$ is *cartesian* monoidal. In the proof of Proposition 3.6, this is reflected by the fact in the main diagram, the morphism $a^\sharp$ commutes with the copy maps.

3.3 Examples of weakly affine monads

Every affine monad is a weakly affine monad. Below you find a few less trivial examples.

▶ **Example 3.7.** Let $M^* : \mathbf{Set} \to \mathbf{Set}$ be the monad assigning to every set the set of finitely supported discrete *non-zero* measures on M^*, or equivalently let $M^*(X)$ for any set X be the set of non-zero finitely supported functions $X \to [0, \infty)$. It is a sub-monad $M^* \subseteq M$, meaning that the monad structure is defined in terms of the same formulas as for the monad of measures M (Definition 2.3). Similarly, the lax structure components

$$c_{X,Y} \; : \; M^*X \times M^*Y \longrightarrow M^*(X \times Y)$$

are also given by the formation of product measures, or equivalently point-wise products of functions $X \to [0, \infty)$.

Since $M^*1 \cong (0,\infty) \ncong 1$, this monad is not affine. However the monoid structure of $(0,\infty)$ induced by M^* is the usual multiplication of positive real numbers, which form a group. Therefore M^* is weakly affine, and its Kleisli category is weakly Markov.

On the other hand, if the zero measure is included, we have $M1 \cong [0,\infty)$ which is not a group under multiplication, so M is not weakly affine.

▶ **Example 3.8.** Let A be a commutative monoid. Then the functor $T_A \coloneqq A \times -$ on **Set** has a canonical structure of commutative monad, where the lax structure components $c_{X,Y}$ are given by multiplying elements in A while carrying the elements of X and Y along.

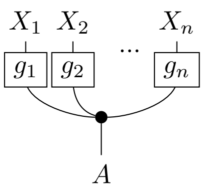
Since $T_A 1 \cong A$, the monad T_A is weakly affine if and only if A is a group, and affine if and only if $A \cong 1$.

▶ **Example 3.9.** As for negative examples, consider the free abelian group monad F on **Set**. Its functor takes a set X and forms the set FX of finite multisets (with repetition, where order does not matter) of elements of X and their formal inverses. We have that $F1 \cong \mathbb{Z}$, which is an abelian group under addition. However, the monoid structure on $F1$ induced by the monoidal structure of the monad corresponds to the *multiplication* on $\mathbb{Z}$, which does not have inverses. Therefore F is not weakly affine.

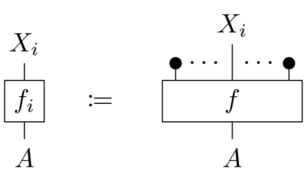
4 Conditional independence in weakly Markov categories

Markov categories have a rich theory of conditional independence in the sense of probability theory [12]. It is noteworthy that some of those ideas can be translated and generalised to the setting of weakly Markov categories.

▶ **Definition 4.1.** A morphism $f : A \to X_1 \otimes \cdots \otimes X_n$ in a gs-monoidal category $\mathcal{C}$ is said to exhibit **conditional independence of the X_i given A** if and only if it can be expressed as a product of the following form



Note that this formulation is a bit different from the earlier definitions given in [1, Definition 6.6] and [9, Definition 12.12], which were formulated for morphisms in Markov categories and state that f exhibits conditional independence if the above holds with the g_i being the *marginals* of f, which are



Indeed, in a Markov category, conditional independence in our sense holds if and only if it holds with $g_i = f_i$ [9, Lemma 12.11]. We also say that f is the *product of its marginals.*

▶ **Example 4.2.** In the Kleisli category of the distribution monad D, which is Markov, a morphism $f : A \to X \otimes Y$ exhibits conditional independence if and only if its value at every $a \in A$ is the product of its marginals [9, Section 12].

Here is what conditional independence looks like in the Kleisli case.

▶ **Proposition 4.3.** *Let $\mathcal{D}$ be a cartesian monoidal category and T a commutative monad on $\mathcal{D}$. Then a Kleisli morphism represented by $f^\sharp : A \to T(X_1 \times \cdots \times X_n)$ exhibits conditional independence of the X_i given A if and only if it factors as*

$$\begin{tikzcd} A \arrow[d, "(g_1^\sharp,\dots,g_n^\sharp)"'] \arrow[dr, "f^\sharp"] & \\ TX_1 \times \cdots \times TX_n \arrow[r, "c"'] & T(X_1 \times \cdots \times X_n) \end{tikzcd}$$

for some Kleisli maps $g_i^\sharp : A \to TX_i$, where the map c above is the one obtained by iterating the lax monoidal structure (which is unique by associativity).

Proof. In terms of the base category $\mathcal{D}$, a Kleisli morphism in the form of Definition 4.1 reads as follows

$$A \xrightarrow{\text{copy}} A \times \cdots \times A \xrightarrow{g_1^\sharp \times \cdots \times g_n^\sharp} TX_1 \times \cdots \times TX_n \xrightarrow{\quad c \quad} T(X_1 \times \cdots \times X_n).$$

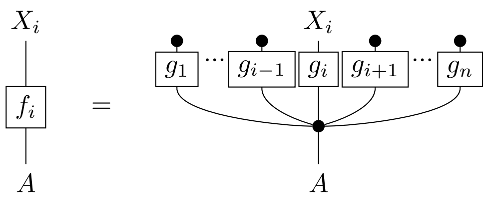
Therefore $f^\sharp : A \to T(X_1 \times \cdots \times X_n)$ exhibits the conditional independence if and only if it is of the form above. ◀

▶ **Example 4.4.** In the Kleisli category of the measure monad M, and for any objects, the morphism $A \to X_1 \otimes \cdots \otimes X_n$ given by the zero measure on every $a \in A$ exhibits conditional independence of its outputs given its input. For example, for $A = 1$, the zero measure on $X \times Y$ is the product of the zero measure on X and the zero (or any other) measure on Y. Notice that both marginals of the zero measure are zero measures – therefore, the factors appearing in the product are not necessarily related to the marginals.

In a weakly Markov category, the situation is similar to the Markov case discussed above, but up to equivalence: an arrow exhibits conditional independence if and only if it is *equivalent to* the product of its marginals.

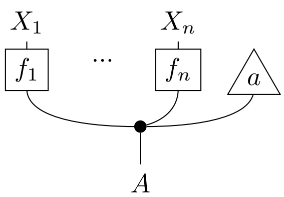
▶ **Proposition 4.5.** *Let $f : A \to X_1 \otimes \cdots \otimes X_n$ be a morphism in a weakly Markov category $\mathcal{C}$. Then f exhibits conditional independence of the X_i given A if and only if it is equivalent to the product of all its marginals.*

Proof. Denote the marginals of f by $f_1, \dots, f_n$. Suppose that f is a product as in Definition 4.1. By marginalising, for each $i = 1, \dots, n$ we get



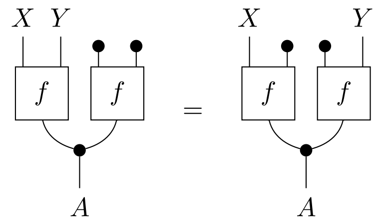
Therefore for each i we have that $f_i \sim g_i$.

Conversely, suppose that f is equivalent to the product of its marginals, i.e. that there exists $a : X \to I$ such that f is equal to the following



One can then choose $g_i = f_i$ for all $i < n$, and $g_n = a \cdot f_n$, so that f is in the form of Definition 4.1. ◀

▶ **Remark 4.6.** For $n = 2$, a morphism $f : A \to X \otimes Y$ in a weakly Markov category $\mathcal{C}$ exhibits conditional independence of X and Y given A if and only if the following equation holds



4.1 Main result

The concept of conditional independence for general weakly Markov categories allows us to give an equivalent characterisation of weakly affine monads. The condition is a pullback condition on the associativity diagram, and it recovers Proposition 2.1 when applied to the monads of the form $A \times -$ for A a commutative monoid.

▶ **Theorem 4.7.** *Let $\mathcal{D}$ be a cartesian monoidal category and T a commutative monad on $\mathcal{D}$. Then the following conditions are equivalent*

1. *T is weakly affine;*
2. *the Kleisli category Kl_T is weakly Markov;*
3. *for all objects X, Y, and Z, the following associativity diagram is a pullback*

$$
\begin{tikzcd}
T(X) \times T(Y) \times T(Z) \arrow[r, "\mathrm{id} \times c_{Y,Z}"] \arrow[d, "c_{X,Y} \times \mathrm{id}"'] & T(X) \times T(Y \times Z) \arrow[d, "c_{X, Y\times Z}"] \\
T(X \times Y) \times T(Z) \arrow[r, "c_{X\times Y, Z}"] & T(X \times Y \times Z)
\end{tikzcd}
\tag{2}
$$

We prove the theorem by means of the following property of weakly Markov categories.

▶ **Lemma 4.8** (localised independence property)**.** *Let $\mathcal{C}$ be a weakly Markov category. Whenever a morphism $f : A \to X \otimes Y \otimes Z$ exhibits conditional independence of $X \otimes Y$ (jointly) and Z given A, as well as conditional independence of X and $Y \otimes Z$ given A, then it exhibits conditional independence of X, Y, and Z given A.*

Proof of Lemma 4.8. Suppose $f : A \to X \otimes Y \otimes Z$ exhibits conditional independence of $X \otimes Y$ (jointly) and Z given A, as well as conditional independence of X and $Y \otimes Z$ given A. By marginalising out X, we have that f_{YZ} exhibits conditional independence of Y and Z given A. Since by hypothesis f exhibits conditional independence of X and $Y \otimes Z$ given A, by Proposition 4.5 we have that f is equivalent to the product of f_X and f_{YZ}. But, again by Proposition 4.5, f_{YZ} is equivalent to the product of f_Y and f_Z, so we have that f is equivalent to the product of all its marginals. Using Proposition 4.5 in the other direction, this means that f exhibits conditional independence of X, Y and Z given A. ◀

We are now ready to prove the theorem.

Proof of Theorem 4.7. $1 \Leftrightarrow 2$: see Proposition 3.6.

$1 \Rightarrow 3$: By the universal property of products, a cone over the cospan in (2) consists of maps $g_1^\sharp : A \to TX$, $g_{23}^\sharp : A \to T(Y \times Z)$, $g_{12}^\sharp : A \to T(X \times Y)$ and $g_3^\sharp : A \to TZ$ such that the following diagram commutes

$$\begin{tikzcd}
A \arrow[drr, bend left, "(g_1^\sharp, g_{23}^\sharp)"] \arrow[ddr, bend right, "(g_{12}^\sharp, g_3^\sharp)"'] & & \\
& T(X) \times T(Y) \times T(Z) \arrow[r, "\mathrm{id} \times c_{Y,Z}"'] \arrow[d, "c_{X,Y} \times \mathrm{id}"] & T(X) \times T(Y \times Z) \arrow[d, "c_{X,Y \times Z}"] \\
& T(X \times Y) \times T(Z) \arrow[r, "c_{X \times Y,Z}"'] & T(X \times Y \times Z)
\end{tikzcd}$$

By Proposition 4.3, this amounts to a Kleisli morphism $f^\sharp : A \to T(X \times Y \times Z)$ exhibiting conditional independence of X and $Y \otimes Z$ given A, as well as of $X \otimes Y$, and Z given A. By the localised independence property (Lemma 4.8), we then have that f exhibits conditional independence of all X, Y and Z given A, and so, again by Proposition 4.3, $f^\sharp$ factors through the product $TX \times TY \times TZ$. More specifically, by marginalising over Z, we have that $g_{12}^\sharp$ factors through $TX \times TY$, i.e. the following diagram on the left commutes for some $h_1^\sharp : A \to TX$ and $h_2^\sharp : A \to TY$, and similarly, by marginalising over X, the diagram on the right commutes for some $\ell_2^\sharp : A \to TY$ and $\ell_3^\sharp : A \to TZ$

$$\begin{tikzcd}
A \arrow[d, "(h_1^\sharp, h_2^\sharp)"'] \arrow[dr, "g_{12}^\sharp"] & \\
TX \times TY \arrow[r, "c_{X,Y}"'] & T(X \times Y)
\end{tikzcd}
\qquad
\begin{tikzcd}
A \arrow[d, "(\ell_2^\sharp, \ell_3^\sharp)"'] \arrow[dr, "g_{23}^\sharp"] & \\
TY \times TZ \arrow[r, "c_{Y,Z}"'] & T(Y \times Z)
\end{tikzcd}$$

In other words, the upper and the left curved triangles in the following diagram commute

$$\begin{tikzcd}
A \arrow[drr, bend left, "(g_1^\sharp, g_{23}^\sharp)"] \arrow[dr, shift left, "(g_1^\sharp, \ell_2^\sharp, \ell_3^\sharp)"] \arrow[dr, shift right, "(h_1^\sharp, h_2^\sharp, g_3^\sharp)"'] \arrow[ddr, bend right, "(g_{12}^\sharp, g_3^\sharp)"'] & & \\
& T(X) \times T(Y) \times T(Z) \arrow[r, "\mathrm{id} \times c_{Y,Z}"'] \arrow[d, "c_{X,Y} \times \mathrm{id}"] & T(X) \times T(Y \times Z) \arrow[d, "c_{X,Y \times Z}"] \\
& T(X \times Y) \times T(Z) \arrow[r, "c_{X \times Y,Z}"'] & T(X \times Y \times Z)
\end{tikzcd}$$

By marginalising over Y and Z, and by weak affinity of T, there exists a unique $a^\sharp : A \to T1$ such that $h_1 = a \cdot g_1$. Therefore

$$g_{12} = h_1 \cdot h_2 = (a \cdot g_1) \cdot h_2 = g_1 \cdot (a \cdot h_2),$$

and so in the diagram above we can equivalently replace h_1 and h_2 with g_1 and $a \cdot h_2$. Similarly, by marginalising over X and Y, there exists a unique $c^\sharp : A \to T1$ such that $\ell_3 = c \cdot g_3$, so that

$$g_{23} = \ell_2 \cdot \ell_3 = \ell_2 \cdot (c \cdot g_3) = (c \cdot \ell_2) \cdot g_3$$

and in the diagram above we can replace ℓ_2 and ℓ_3 with $c \cdot \ell_2$ and g_3, as follows

$$\begin{tikzcd}
A \arrow[drr, shift left, "(g_1^\sharp, (c\cdot\ell_2)^\sharp, g_3^\sharp)"] \arrow[drr, shift right, "(g_1^\sharp, (a\cdot h_2)^\sharp, g_3^\sharp)"'] \arrow[drrr, bend left, "(g_1^\sharp, g_{23}^\sharp)"] \arrow[ddrr, bend right, "(g_{12}^\sharp, g_3^\sharp)"'] & & & \\
& & T(X) \times T(Y) \times T(Z) \arrow[r, "\mathrm{id} \times c_{Y,Z}"'] \arrow[d, "c_{X,Y} \times \mathrm{id}"] & T(X) \times T(Y \times Z) \arrow[d, "c_{X, Y\times Z}"] \\
& & T(X \times Y) \times T(Z) \arrow[r, "c_{X\times Y, Z}"'] & T(X \times Y \times Z)
\end{tikzcd}$$

Now, marginalising over X and Z, we see that necessarily $a \cdot h_2 = c \cdot \ell_2$. Therefore there is a unique map $A \to TX \times TY \times TZ$ making the whole diagram commute, which means that (2) is a pullback.

$3 \Rightarrow 1$: If T is weakly affine, then taking $X = Y = Z = 1$ in (2) shows that this monoid must be an abelian group: we obtain a unique arrow $\iota \colon T1 \to T1$ making the following diagram commute

$$\begin{tikzcd}
T1 \arrow[dr, "(\mathrm{id}, \iota, \mathrm{id})"] \arrow[drr, bend left, "(\mathrm{id}, \eta_1 \mathrm{del}_{T1})"] \arrow[ddr, bend right, "(\eta_1 \mathrm{del}_{T1}, \mathrm{id})"'] & & & \\
& T1 \times T1 \times T1 \arrow[r, "\mathrm{id} \times c_{1,1}"] \arrow[d, "c_{1,1} \times \mathrm{id}"'] & T1 \times T(1 \times 1) \arrow[r, "\cong"] \arrow[d, "c_{1, 1\times 1}"] & T1 \times T1 \arrow[d, "c_{1,1}"] \\
& T(1 \times 1) \times T1 \arrow[r, "c_{1\times 1, 1}"] \arrow[d, "\cong"'] & T(1 \times 1 \times 1) \arrow[r, "\cong"] \arrow[d, "\cong"] & T(1 \times 1) \arrow[d, "\cong"] \\
& T1 \times T1 \arrow[r, "c_{1,1}"'] & T(1 \times 1) \arrow[r, "\cong"'] & T1
\end{tikzcd}$$

and the commutativity shows that ι satisfies the equations making it the inversion map for a group structure. ◀

▶ **Example 4.9.** In the Kleisli category of the measure monad Kl_M (which is not weakly affine) consider the following diagram

$$\begin{tikzcd}
MX \times MY \times MZ \arrow[r, "\mathrm{id} \times c_{Y,Z}"] \arrow[d, "c_{X,Y} \times \mathrm{id}"'] & MX \times M(Y \times Z) \arrow[d, "c_{X, Y \times Z}"] \\
M(X \times Y) \times MZ \arrow[r, "c_{X \times Y, Z}"] & M(X \times Y \times Z)
\end{tikzcd}$$

In the top-right corner $MX \times M(Y \times Z)$, take the pair $(0, p)$ where p is any non-zero measure on $Y \times Z$, and similarly, in the bottom-left corner take the pair $(q, 0)$ where q is any non-zero measure on $X \times Y$. Following the diagram, both pairs are mapped to the zero measure in the bottom-right corner. If the diagram was a pullback, we would be able to express the top-right and bottom-left corners as coming from the same triple in $MX \times MY \times MZ$, that is, there would exist a measure m on Y such that $m \cdot 0 = p$ and $0 \cdot m = q$. Since p and q are non-zero, this is not possible.

▶ **Remark 4.10.** It is worth noting that the pullback condition on the associativity square is not equivalent to the localised independence property of Lemma 4.8: recall that a zero measure always exhibits conditional independence of all its outputs (Example 4.4). Therefore, for zero measures, the localised independence property is always trivially valid, and hence the Kleisli category of the measures monad M satisfies it in general. However, the example above shows explicitly that the pullback property fails.

For now it is an open question whether the localised independence property for a Kleisli category is reflected by an equivalent condition on the monad.

5 Conclusions and future work

Our paper introduces weak Markov categories and weakly affine monads and explore their relationship. More explicitly, our main result (Theorem 4.7) establishes a tight correspondence between the algebraic properties of $T1$ and the universal properties of certain commutative squares given by the structural arrows of T for a commutative monad T on a cartesian category. We believe that this theorem suggests at least two potential directions for future research, namely

- generalising the statement to weakly affine monads on weakly Markov categories;
- generalising other Markov-categorical notions, such as the positivity axiom, to weakly Markov or even general gs-monoidal categories.

We will provide further details on these potential directions in the following.

Regarding possible generalisations. In Theorem 4.7, we provide a characterisation of weakly affine monads on cartesian monoidal categories. Taking inspiration from the case of affine monads on Markov categories [9, Corollary 3.2], it seems natural to consider whether our main result can be extended to commutative monads on *weakly Markov categories*.

However, solving this problem is non-trivial and requires clever adjustments to the main definitions. The crucial point is that, in general, the structure of the internal group of $T1$ and the structure of the group $\mathcal{D}(X, T1)$ are not necessarily related in the current definitions. One approach could be to introduce a form of *compatibility* for $T1$ and $\mathcal{D}(X, T1)$ by defining a weakly affine monad on a weakly Markov category as a commutative monad such that $T1$ is an internal group and $\mathcal{D}(X, T1)$ is a group with the composition and units induced by those of $T1$. With this change, for example, Proposition 3.6 would work for any weakly Markov category, but Theorem 4.7 would likely fail as its proof involves the universal property of products.

On the positivity axiom. Recall that a strong monad T on a cartesian monoidal category is *strongly affine* [15] if for every pair of objects X and Y the following diagram is a pullback

$$\begin{tikzcd} X \times TY \arrow[r, "s"] \arrow[d, "\pi_1"] & T(X \times Y) \arrow[d, "T\pi_1"] \\ X \arrow[r, "\eta"] & TX \end{tikzcd}$$

where s denotes the strength and η denotes the unit of the monad. Every strongly affine monad is affine. The corresponding condition on the Markov category Kl_T has recently been characterised as an information flow axiom called *positivity* [11, Section 2].

For a generic commutative monad, the diagram above may even fail to commute (take for example the measure monad M, and start with $(x, 0)$ in the top left corner). One can however consider the following diagram, which reduces to the one above (up to isomorphism) in the affine case

$$\begin{tikzcd}
X \times TY \arrow[r, "s"] \arrow[d, "\mathrm{id} \times T(\mathrm{del}_Y)"] & T(X \times Y) \arrow[d, "T(\mathrm{id} \times \mathrm{del}_Y)"] \\
X \times T1 \arrow[r, "s"] & T(X \times 1) \cong TX
\end{tikzcd}$$

and which always commutes by naturality of the strength. One can then call the monad T *positive* if this second diagram is a pullback. Upon defining *positive gs-monoidal categories* analogously to positive Markov categories, one may conjecture that T is positive if and only if Kl_T is positive. This would generalise the existing result for Markov categories.

References

1. Kenta Cho and Bart Jacobs. Disintegration and Bayesian inversion via string diagrams. *Mathematical Structures in Computer Science*, 29(7):938–971, 2019.
2. Bob Coecke, Bill Edwards, and Robert W. Spekkens. Phase groups and the origin of non-locality for qubits. In Bob Coecke, Prakash Panangaden, and Peter Selinger, editors, *QPL@MFPS 2009*, volume 270 of *ENTCS*, pages 15–36. Elsevier, 2009.
3. Andrea Corradini and Fabio Gadducci. A 2-categorical presentation of term graph rewriting. In Eugenio Moggi and Giuseppe Rosolini, editors, *CTCS 1997*, volume 1290 of *LNCS*, pages 87–105. Springer, 1997.
4. Andrea Corradini and Fabio Gadducci. An algebraic presentation of term graphs, via gs-monoidal categories. *Applied Categorical Structures*, 7(4):299–331, 1999.
5. Andrea Corradini and Fabio Gadducci. Rewriting on cyclic structures: equivalence between the operational and the categorical description. *RAIRO - Theoretical Informatics and Applications - Informatique Théorique et Applications*, 33(4-5):467–493, 1999.
6. Andrea Corradini and Fabio Gadducci. A functorial semantics for multi-algebras and partial algebras, with applications to syntax. *Theoretical Computer Science*, 286(2):293–322, 2002.
7. Dion Coumans and Bart Jacobs. Scalars, monads, and categories. In Chris Heunen, Mehrnoosh Sadrzadeh, and Edward Grefenstette, editors, *Quantum Physics and Linguistics - A Compositional, Diagrammatic Discourse*, pages 184–216. Oxford University Press, 2013.
8. Thomas Fox. Coalgebras and cartesian categories. *Communications in Algebra*, 4(7):665–667, 1976.
9. Tobias Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. *Advances in Mathematics*, 370:107239, 2020.
10. Tobias Fritz, Fabio Gadducci, Davide Trotta, and Andrea Corradini. Lax completeness for gs-monoidal categories. arXiv:2205.06892.
11. Tobias Fritz, Tomáš Gonda, Nicholas Gauguin Houghton-Larsen, Paolo Perrone, and Dario Stein. Dilations and information flow axioms in categorical probability. arXiv:2211.02507.
12. Tobias Fritz and Andreas Klingler. The d-separation criterion in categorical probability. *Journal of Machine Learning Research*, 24(46):1–49, 2023.
13. Tobias Fritz and Wendong Liang. Free gs-monoidal categories and free Markov categories. *CoRR*, abs/2204.02284, 2022.
14. Bart Jacobs. Semantics of weakening and contraction. *Annals of Pure and Applied Logic*, 69(1):73–106, 1994.
15. Bart Jacobs. Affine monads and side-effect-freeness. In Ichiro Hasuo, editor, *CMCS 2016*, volume 9608 of *LNCS*, pages 53–72. Springer, 2016.
16. Anders Kock. Bilinearity and cartesian closed monads. *Mathematica Scandinavica*, 29(2):161–174, 1971.
17. Anders Kock. Commutative monads as a theory of distributions. *Theory and Applications of Categories*, 26, 2012.

18 Peter Selinger. A survey of graphical languages for monoidal categories. In Bob Coecke, editor, *New Structures for Physics*, volume 813 of *Lecture Notes in Physics*, pages 289–355. Springer, 2011.