Submitted to MFPS 2025

Order in Partial Markov Categories

Elena Di Lavore[a] Mario Román[a] Paweł Sobociński[b] Márk Széles[c]

[a] *University of Oxford, UK*
[b] *Tallinn University of Technology, Estonia*
[c] *Radboud University Nijmegen, The Netherlands*

Abstract

Partial Markov categories are a recent framework for categorical probability theory, providing an abstract account of partial probabilistic computation. In this article, we discuss two order relations on the morphisms of a partial Markov category. In particular, we prove that every partial Markov category is canonically enriched over the category of preordered sets and monotone maps. We show that our construction recovers several well-known order enrichments. We also demonstrate that the existence of codiagonal maps (comparators) is closely related to order properties of partial Markov categories.
We propose a synthetic version of the Cauchy–Schwarz inequality to facilitate inequational reasoning in partial Markov categories. We apply this new axiom to prove that updating a prior distribution with an evidence predicate increases the likelihood of the evidence in the posterior.

Keywords: Markov categories, preorder enrichment, string diagrams, probabilistic inference, copy-discard categories

1 Introduction

Markov categories [Fri20,CJ19] are a synthetic framework for probability theory. They allow us to reason about probabilistic processes using a few basic axioms that model key aspects of probabilistic computation. Morphisms in a Markov category can be composed both in sequence and in parallel, via a symmetric monoidal structure. Every object is equipped with a commutative comonoid, which allows for copying and discarding information. An important axiom is the existence of conditionals, which are an abstraction of conditional probability distributions. This simple setting has been successfully applied to various areas in probability theory [CFG+24,FK23,Jac21], semantics of probabilistic programming [Ste21,SS21,AFK+24] , causal inference [JKZ21,LT23,YZ22] and information theory [Per24].

Morphisms in Markov categories can be thought of as total probabilistic computations. Totality is encoded by an axiom – naturality of discarding – declaring that maps can be discarded without affecting the meaning of the process. However, many operations in probability theory are inherently partial. An important example is bayesian updating: if one tries to learn from evidence that is incompatible with the prior belief, updating is impossible. Markov categories lack the structure to express this partiality.

In order to express partiality, we need to drop naturality of discarding, and thus work with copy-discard (CD) categories [CJ19,CG99]. Morphisms in such copy-discard categories correspond to partial probabilistic computations. We would like to highlight two recent works following this approach. In [LT23], the authors start from a copy-discard category, and derive conditional distributions from equality comparison and normalisation of subprobability kernels. While the theory is elegant, it is mostly suited for discrete probability. In contrast, the framework of partial Markov categories [DR23] starts from a copy-discard category with conditionals. The advantage of this approach is the abundance of models. Examples of partial Markov categories include Kleisli-categories of 'subprobability monads', cartesian restriction categories [CL02,CL03,CL07], and cartesian bicategories of relations [CW87].

The main contribution of this paper is to show that every partial Markov category is enriched in the category of partial orders and monotone maps. That is, every hom-set in a partial Markov category is a preorder, and the operations of the symmetric monoidal category structure are monotone with respect to this preorder. The existence of such an order-enrichment makes the setting particularly appealing for developing the basics of inequational reasoning in synthetic probability theory. This is important, since some crucial properties of bayesian reasoning are inequational, such as the fact that updating a prior with evidence increases the likelihood of the evidence being true in the posterior.

This paper is structured as follows. In Section 2 we recall the basic theory of partial Markov categories. We also introduce some running examples. In Section 3, we define the preorder enrichment on partial Markov categories; we display several well-known order enrichments as instances of our construction. Section 4 is dedicated to the study of least conditionals in partial Markov categories. Our main results in this section are about the relationship of least conditionals and the existence of comparators (Theorems 4.4 and 4.6). Finally, Section 5 is an application of the conditional preorder. We derive a synthetic analogue of the validity-increase theorem, under suitable assumptions (Theorem 5.13).

Throughout this paper we will use string-diagrammatic notation with the diagrams written from left to right. For a detailed introduction to string-diagrammatic calculi, see [Sel10]. Occasionally, we will also apply equational reasoning. To ease notation, we will treat all monoidal categories as if they were strict.

Related work

Inequalities play an important role in the theory of cartesian restriction categories [CL02] and cartesian bicategories of relations [CW87]. In fact, restriction categories have a canonical ordering on morphisms [CL02, §2.1.4] and can be seen as enriched categories [CG14]. Similarly, morphisms in bicategories of relations have a canonical ordering that forms an enrichment [Nes24]. These canonical inequalities are instances of the ordering we introduce in this paper, see Propositions 3.11 and 3.12.

Order enrichments for categories of probabilistic maps have been considered before, in the context of effectus theory [CJWW15,Cho19,Jac18].

One of the anonymous reviewers pointed out the recent work on so-called quasi-Markov categories [FGL+25], which generalise cartesian restriction categories. Quasi-Markov categories are poset-enriched by the normalisation order. Investigation of the relationship between partial Markov categories and quasi-Markov categories is left for future work.

2 Preliminaries

Copy-discard categories (Definition 2.8) are theories of processes – symmetric monoidal categories – where resources can be copied and discarded. Partial Markov categories are copy-discard categories that additionally have conditionals: a factorization property often needed in categorical probability.

We go a step further and explicitly introduce *copy-discard-compare (CDC) categories* (Definition 2.8): theories of processes where we can assert the equality of two resources ("compare" them). Discrete partial Markov categories are copy-discard-compare categories with conditionals (Definition 2.9). Table 1 summarises the terminology related to variations copy-discard categories.

Definition 2.1 (Copy-discard category) A *copy-discard category* [CG99,Fri20] is a symmetric monoidal category $(\mathsf{C}, \otimes, I)$ in which every object X has a compatible commutative comonoid structure, consisting of a counit or *discard*, $\varepsilon_X : X \to I$; and a comultiplication or *copy*, $\delta_X : X \to X \otimes X$, satisfying the following axioms. 1.

(i) Comultiplication is associative, $\delta_X \fatsemi (\delta_X \otimes \mathrm{id}) = \delta_X \fatsemi (\mathrm{id} \otimes \delta_X)$.

(ii) Counit is neutral for comultiplication, $\delta_X \fatsemi (\varepsilon_X \otimes \mathrm{id}) = \mathrm{id} = \delta_X \fatsemi (\mathrm{id} \otimes \varepsilon_X)$.

(iii) Comultiplication is uniform, $\delta_{X\otimes Y} = (\delta_X \otimes \delta_Y) \fatsemi (\mathrm{id} \otimes \sigma \otimes \mathrm{id})$, and $\delta_I = \mathrm{id}$.

(iv) Counit is uniform, $\varepsilon_{X\otimes Y} = \varepsilon_X \otimes \varepsilon_Y$ and $\varepsilon_I = \mathrm{id}$.

(v) Comultiplication is commutative, $\delta_X \fatsemi \sigma_{X,X} = \delta_X$, where $\sigma_{X,X} : X \otimes X \to X \otimes X$ is the symmetry.

Structure	**Maps are**	**Comparators/caps**	**Conditionals**
Copy-discard category			
Copy-discard-compare category		✓	
Cartesian restriction category	deterministic		
Discrete cartesian restriction category	deterministic	✓	
Quasi-Markov category	self-normalising [FGL+25]		
Partial Markov category			✓
Discrete partial Markov category		✓	✓
Markov category	Maps are total		
Markov category with conditionals	Maps are total		✓

Table 1
In each row, the structure entails all properties listed (✓).

The counits and comultiplications are not required to form natural transformations.

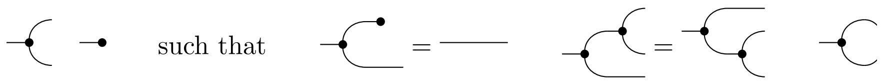


Fig. 1. Structure and axioms of a commutative comonoid.

The language of copy-discard categories can express when morphisms are deterministic, total and self-normalising.

Definition 2.2 (Deterministic, total) A morphism $f\colon X \to Y$ in a copy-discard category is *deterministic* if it commutes with the copy, $f \mathbin{⨟} \delta_Y = \delta_X \mathbin{⨟} (f \otimes f)$. It is *total* if it commutes with the discard, $f \mathbin{⨟} \varepsilon_Y = \varepsilon_X$.

Definition 2.3 (Normalisation) ([DR23,LT23]) Let C be a copy-discard category.

(i) We say that a map $g : X \to Y$ is a normalisation of $f : X \to Y$ if $f = \delta_X \mathbin{⨟} (g \otimes (f \mathbin{⨟} \varepsilon_Y))$. In this situation, we write $f \preceq g$. See Figure 2 for an illustration.

(ii) A map f is called self-normalising if $f \preceq f$. Clearly, deterministic maps are self-normalising.

The collection of self-normalising maps is partially ordered by normalisation [LT23, Lemma 98].

$f \preceq g$ if and only if

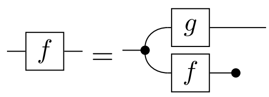


Fig. 2. The normalisation order

Definition 2.4 The conditional composition of two morphisms, $f\colon X \to A$ and $g\colon A \otimes X \to B$, of a copy-discard category, is the morphism defined by the equation in Figure 3. In other words, conditional composition passes a copy of the input and the output of the first morphism to the second morphism.

$$(f \triangleleft g) = \delta_X \mathbin{⨟} ((f \mathbin{⨟} \delta_A) \otimes \mathrm{id}_X) \mathbin{⨟} (\mathrm{id}_A \otimes g)$$

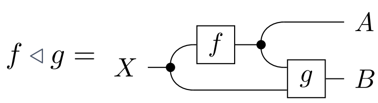


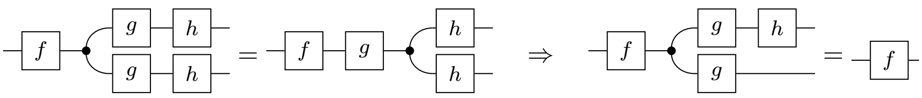
Fig. 3. Conditional composition.

Definition 2.5 (Conditionals) A copy-discard category has *conditionals* if every morphism $f\colon X \to A \otimes B$ can be decomposed as $f = (f \,⨾\, (\mathrm{id}_A \otimes \varepsilon_B)) \triangleleft c$, for some self-normalising $c\colon A \otimes X \to B$. In this situation we say that x is a conditional of f. For morphisms $f\colon X \to A$ and $g\colon A \to B$, a Bayesian inverse of g with respect to f is a conditional $g^{\dagger}_{f}\colon B \otimes X \to A$ of $f \,⨾\, \delta_A \,⨾\, (g \otimes \mathrm{id})$.

Definition 2.6 ([DR23]) A partial Markov category is a copy-discard category with conditionals.

We will later work with a particular subclass of copy-discard categories: those that are *balanced*. Balanced copy-discard categories cover a large class of examples, and they are particularly well-behaved with respect to the partial order we introduce in Definition 3.1. The name *balanced* is related to properties of idempotent morphisms: in a balanced copy-discard category every idempotent is balanced [FGL+23, Theorem 4.1.8.].

Definition 2.7 (Balanced copy-discard category [FGL+23]) A copy-discard category C is *balanced* if the following implication holds for all appropriately typed morphisms.



Being balanced was introduced as the *Cauchy–Schwarz property* [FGL+23, Definition A.5.1]; we avoid this name to avoid confusion with the synthetic version of the Cauchy–Schwarz inequality we introduce later. We moreover remark that there is another notion of *balanced category* in the literature [Joh14], unrelated to ours.

Definition 2.8 (Copy-Discard-Compare category) A copy-discard-compare category is a copy-discard category in which every object, X, has a compatible partial Frobenius structure, consisting of an additional commutative multiplication, or *comparator*, $\mu_X\colon X \otimes X \to X$, satisfying the following axioms.

(i) Multiplication is associative, $(\mu_X \otimes \mathrm{id}) \,⨾\, \mu_X = (\mathrm{id} \otimes \mu_X) \,⨾\, \mu_X$.

(ii) Multiplication is commutative, $\sigma_{X,X} \,⨾\, \mu_X = \mu_X$.

(iii) Multiplication is right inverse to comultiplication, $\delta_X \,⨾\, \mu_X = \mathrm{id}$.

(iv) Multiplication satisfies the Frobenius rule, $(\delta_X \otimes \mathrm{id}) \,⨾\, (\mathrm{id} \otimes \mu_X) = \mu_X \,⨾\, \delta_X = (\mathrm{id} \otimes \delta_X) \,⨾\, (\mu_X \otimes \mathrm{id})$.

(v) Multiplication is uniform, $\mu_{X \otimes Y} = (\mathrm{id} \otimes \sigma_{Y,X} \otimes \mathrm{id}) \,⨾\, (\mu_X \otimes \mu_Y)$, and $\mu_I = \mathrm{id}$.

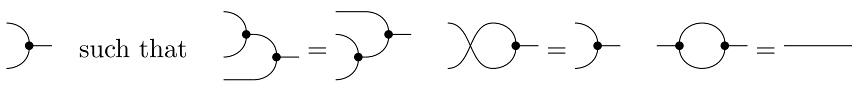


Fig. 4. Structure and axioms of comparators.

Definition 2.9 (Discrete Partial Markov category [DR23]) A discrete partial Markov category is a copy-discard-compare category with conditionals.

We provide an alternative characterization of copy-discard-compare categories via cap morphisms. This is the approach followed by Lorenz and Tull [LT23], where it was mentioned that the presentations via caps and comparators are equivalent. We provide a proof for completeness.

Definition 2.10 (Caps [LT23, Definition 6]) A copy-discard category is said to have caps if it is equipped with morphisms $\cap_X : X \otimes X \to I$, satisfying the following axioms.

(i) Caps are commutative, $\sigma_{X,X} \,⨾\, \cap_X = \cap_X$.

(ii) Caps interact with the comultiplication, $\delta_X \,⨾\, \cap_X = \varepsilon_X$.

(iii) Caps satisfy the Frobenius rule, $(\delta_X \otimes \mathrm{id}) \,⨾\, (\mathrm{id} \otimes \cap_X) = (\mathrm{id} \otimes \delta_X) \,⨾\, (\cap_X \otimes \mathrm{id})$.

(iv) Caps are uniform, $\cap_{X \otimes Y} = (\mathrm{id} \otimes \sigma_{Y,X} \otimes \mathrm{id}) \,⨾\, (\cap_X \otimes \cap_Y)$, and $\cap_I = \mathrm{id}$.

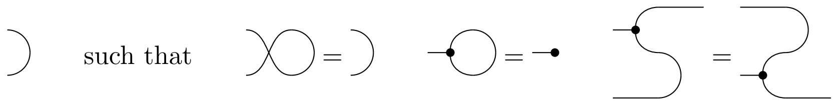


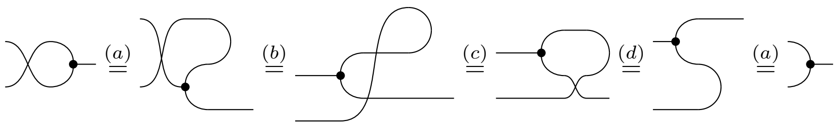
Fig. 5. Structure and axioms of caps.

Proposition 2.11 *A copy-discard category is a copy-discard-compare category if and only if it has caps.*

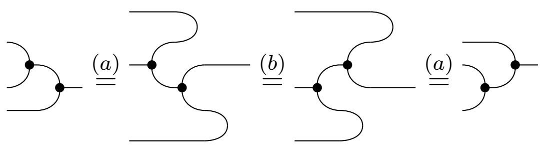
Proof. Assume that C is a copy-discard-compare category. We may define caps via $\cap_X = \mu_X \mathbin{⨟} \varepsilon_X$; the cap axioms follow immediately.

Conversely, assume that C has caps. Define the comparator as one of the two sides of the Frobenius rule (iii), $\mu_X = (\delta_X \otimes \mathrm{id}) \mathbin{⨟} (\mathrm{id} \otimes \cap_X) = (\mathrm{id} \otimes \delta_X) \mathbin{⨟} (\cap_X \otimes \mathrm{id})$.

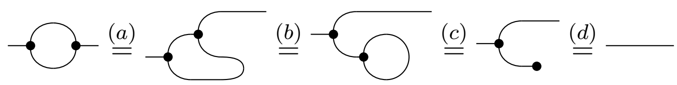
First, we show commutativity using (a) the definition of the compare map, (b) naturality of symmetries, (c) commutativity of the cap, and (d) commutativity of copy.



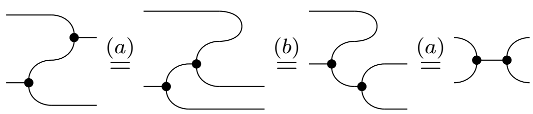
For associativity, we use (a) the definition of the compare map, and (b) associativity of copy.



Compare is right inverse to copy by (a) definition of the compare map, (b) associativity of copy, (c) axioms of caps, and (d) counitality of copy.



Finally, we show the Frobenius equation using (a) the definition of the compare map, and (b) associativity of copy.



□

2.1 Examples

We introduce some partial Markov categories as running examples.

Categories of subprobability kernels

The category $\mathsf{BorelStoch}_{\leq 1}$ of standard Borel spaces and measurable subprobability kernels is a discrete partial Markov category [DR23]. The subcategory $\mathsf{FinStoch}_{\leq 1} \hookrightarrow \mathsf{BorelStoch}_{\leq 1}$ of finite sets and substochastic matrices is also a discrete partial Markov category. One can see that both $\mathsf{BorelStoch}_{\leq 1}$ and $\mathsf{FinStoch}_{\leq 1}$ are balanced, by essentially the same argument as in [FGL⁺23, Proposition A.5.2.].

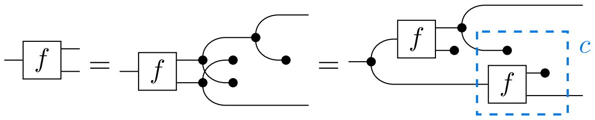
Cartesian restriction categories

Cartesian restriction categories [CL02,CL03,CL07] are an abstraction of the category Par of sets and partial functions. They are copy-discard categories where comultiplication is natural, i.e. all morphisms are deterministic but not necessarily total. Similarly, *discrete Cartesian restriction categories* are Cartesian restriction categories with a comparator.

We now show that cartesian restriction categories have conditionals. That is, cartesian restriction categories are partial Markov categories where all morphisms are deterministic.

Proposition 2.12 *All cartesian restriction categories are partial Markov categories.*

Proof. All morphisms in cartesian restriction categories are, by definition, deterministic. We use this to show that there are conditionals.



These conditionals are self-normalising because every morphism is deterministic. □

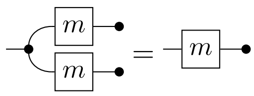
Since all maps are deterministic, cartesian restriction categories are balanced.

Cartesian bicategories of relations

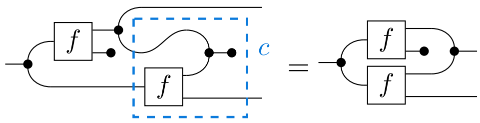
Cartesian bicategories of relations are an abstraction of the category Rel of sets and relations. They are copy-discard-compare categories with extra structure and axioms that give them the expressive power of regular logic (see Figure A.1). They are also an example class of partial Markov categories.

Proposition 2.13 *All cartesian bicategories of relations are partial Markov categories.*

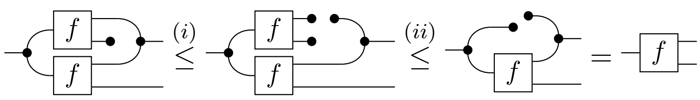
Proof. The compact closed structure and idempotency of convolution give self-normalising conditionals. We use the syntax of cartesian bicategories of relations [CW87]. Every morphism in a cartesian bicategory of relations is self-normalising because convolution is idempotent (see Figure A.1).



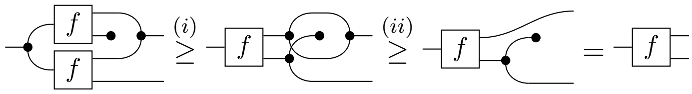
The equation for conditionals can be simplified using the Frobenius equation.



Cartesian bicategories of relations are poset-enriched (see Equation (1)). We use this partial order to bound the morphism above from above with f using adjointness of the discard and the codiscard (i), and lax naturality of the discard morphism (ii).



We bound the same morphism also from below with f using lax naturality of the copy morphism (i), and adjointness of the copy with the cocopy (ii).



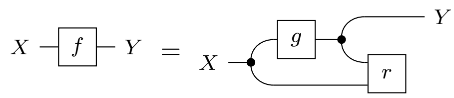
□

Cartesian bicategories of relations are not balanced in general. One can easily adapt [FGL+23, Example A.5.4.] to get a counterexample in Rel.

3 Conditional inequality

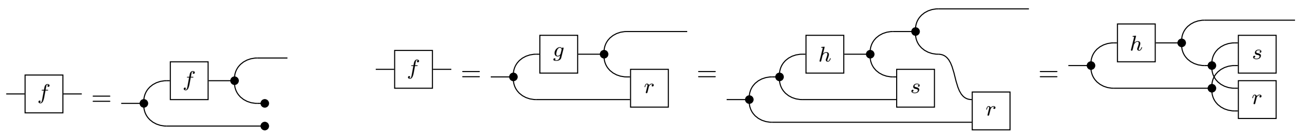
Let us endow partial Markov categories with a preorder-enrichment, given by the conditional inequality (Definition 3.1). We prove that this order is an enrichment in Theorem 3.3 and derive some basic inequalities.

Definition 3.1 (Conditional inequality). In any copy-discard category, conditional inequality relates two parallel morphisms, $f \sqsubseteq g$ for $f \colon X \to Y$ and $g \colon X \to Y$, if and only if there exists some morphism $r \colon Y \otimes X \to I$ such that $f = g \triangleleft r$. In this case, we say that r *witnesses* the inequality $f \sqsubseteq g$, and write $f \sqsubseteq^r g$.



Proposition 3.2 *Conditional inequality ($\sqsubseteq$) is a preorder.*

Proof. Reflexivity is witnessed by the discard morphism, $f \triangleleft \varepsilon = f$, as on the left below. Let us prove transitivity. If we assume $f \sqsubseteq^r g$ and $g \sqsubseteq^s h$, then by associativity of conditional composition, $f = g \triangleleft r = (h \triangleleft s) \triangleleft r = h \triangleleft (s \triangleleft r)$. Thus, $s \triangleleft r$ witnesses $f \sqsubseteq h$, as on the right below.

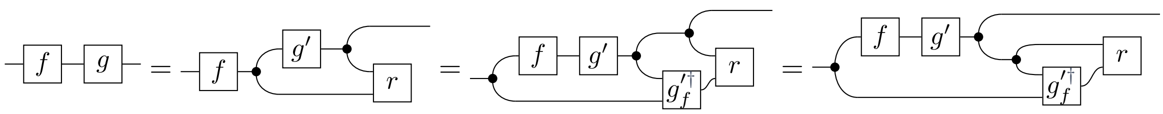


□

Theorem 3.3 *Any partial Markov category is a preorder-enriched symmetric monoidal category.*

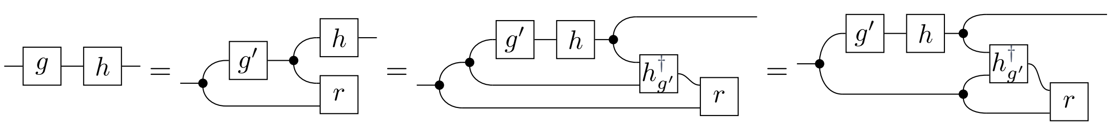
Proof. We show that the preorder ($\sqsubseteq$) on hom-sets is respected by composition and monoidal products.

Assuming that $g \sqsubseteq^r g'$, we can see that $f \fatsemi g \sqsubseteq f \fatsemi g'$ by $f \fatsemi g = f \fatsemi (g' \triangleleft r) = (f \fatsemi g') \triangleleft (g'^{\dagger}_{f} \triangleleft (r \otimes \varepsilon))$.



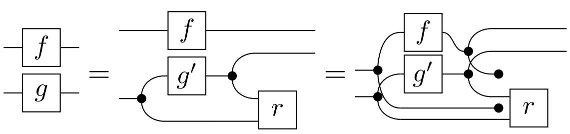
Assuming that $g \sqsubseteq^r g'$, we can see that $g \fatsemi h \sqsubseteq g' \fatsemi h$ by

$$g \fatsemi h = (g' \triangleleft r) \fatsemi h = (g' \fatsemi h) \triangleleft (h^{\dagger}_{g'} \triangleleft (\varepsilon \otimes r)).$$



Assuming that $g \sqsubseteq^r g'$, we can see that $f \otimes g \sqsubseteq f \otimes g'$ by

$$f \otimes g = f \otimes (g' \triangleleft r) = (f \otimes g') \triangleleft (\varepsilon \otimes r).$$



We proceed analogously to show that $f \sqsubseteq f'$ implies $f \otimes g \sqsubseteq f' \otimes g$. We showed that composition and tensors are monotone. Therefore, the conditional preorder is an enrichment. □

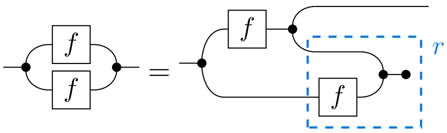
We collect a few elementary properties of the conditional preorder.

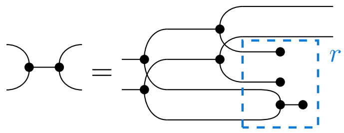
Proposition 3.4 *Let* C *be a discrete partial Markov category.*

(i) *For all* $f : X \to Y$*,* $\delta \fatsemi (f \otimes f) \fatsemi \mu \sqsubseteq f$

(ii) *Multiplication and comultiplication form an adjoint pair* $\delta \dashv \mu$ *in the sense of 2-categories [Lac10]. That is,* $\mathrm{id} \sqsubseteq \delta \fatsemi \mu$ *and* $\mu \fatsemi \delta \sqsubseteq \mathrm{id}$.

Proof. For point (i), observe that the equality below left follows from the partial Frobenius axioms. For point (ii), the first inequality holds by definition, while the second one can be derived from the partial Frobenius axioms.





□

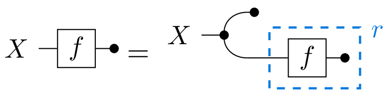
3.1 Subunital enrichments

One may be interested in partial Markov categories with additional logical structure. In such categories, effects $X \to I$ can be viewed as (fuzzy) predicates: e.g., the discard map, $\varepsilon_X : X \to I$, is interpreted as the 'true' predicate. When equipping a partial Markov category with a preorder enrichment, one would reasonably expect truth to be the largest predicate.

Call a preorder enrichment *subunital* if $p \leq \varepsilon_X$ for all effects $p : X \to I$.

Proposition 3.5 *In any partial Markov category, the conditional preorder is subunital.*

Proof. By counitality of copy.



□

Moreover, as the next proposition shows, the conditional preorder is the least subunital preorder enrichment on a partial Markov category.

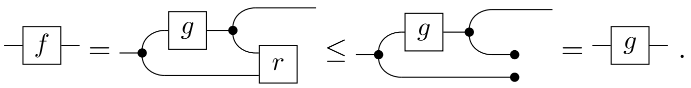
Proposition 3.6 *Let* $(\mathsf{C}, \leq)$ *be a preorder-enriched partial Markov category satisfying* $p \leq \varepsilon_X$ *for all effects* $p \colon X \to I$.

Then, for any two parallel morphisms, $f, g \colon X \to Y$*, the following implication holds:*

$$f \sqsubseteq g \implies f \leq g.$$

An immediate consequence of this is that, if $(\leq)$ *is a poset-enrichment, then the conditional preorder is also a poset, i.e., it is antisymmetric.*

Proof. Let $f \sqsubseteq g$ be witnessed by $r \colon Y \otimes X \to I$. Then,



□

3.2 Normalisation and the conditional preorder

In this subsection, we relate the conditional preorder to the normalisation order.

The proof of the following statement is immediate from the relevant definitions.

Lemma 3.7 *Let* C *be copy-discard category, and let* $f, g : X \to Y$ *be two self-normalising parallel morphisms. Then,* $f \preceq g$ *implies* $f \sqsubseteq g$.

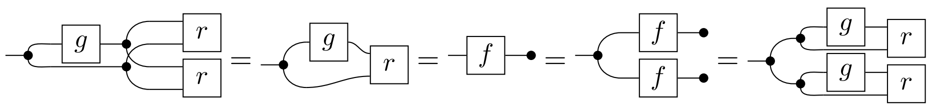
Proof. The conditional inequality $f \sqsubseteq g$ is witnessed by $f \fatsemi \varepsilon_Y$. □

The converse implication does not hold in general, see Example 3.13. However, under certain assumptions, we can derive it for self-normalising maps.

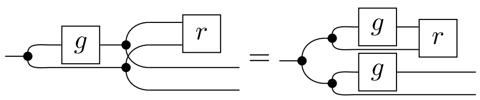
Definition 3.8 (Sharp witness property) A partial Markov category, C, satisfies the *sharp witness property* if each inequality, $f \sqsubseteq g$, between parallel self-normalising morphisms, $f, g : X \to Y$, is witnessed by a deterministic effect $r : Y \otimes X \to I$.

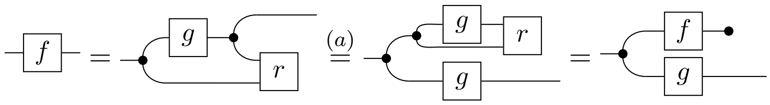
Proposition 3.9 *Let C be a balanced copy-discard category that satisfies the* sharp witness property*. Let $f, g \colon X \to Y$ be parallel self-normalising maps. Then, $f \sqsubseteq g$ if and only if $f \preceq g$.*

Proof. By Lemma 3.7, $f \preceq g$ implies $f \sqsubseteq g$. For the other direction, we use balance and sharp witnesses. Suppose $f \sqsubseteq g$ with a sharp witness r. Since f is self-normalising, we obtain:



By balance, we obtain the equality below left.





By, then, discarding the second output, we obtain the equality (a) below right, which shows that $f \preceq g$. □

3.3 Examples

We finish this section by instantiating the conditional preorder in our running examples.

Finitary subdistributions

Let $f, g \colon X \to \mathbf{D}_{\leq 1}(Y)$ be two parallel substochastic channels in $\mathsf{FinStoch}_{\leq 1}$. The inequality $f \sqsubseteq g$ holds exactly when f is pointwise dominated by g; that is, for all $y \in Y$ and $x \in X$, we have $f(y \mid x) \leq g(y \mid x)$.

TODO:sharp witness property.

Standard Borel spaces and subprobability kernels

One can define a pointwise order on parallel maps $f, g : (X, \sigma_X) \to (Y, \sigma_Y)$ in $\mathsf{BorelStoch}_{\leq 1}$ by setting $f \leq g$ if and only if $f(T \mid x) \leq g(T \mid x)$ for all points $x \in X$ and measurable sets $T \in \sigma_Y$. This pointwise order is a subunital enrichment, see e.g. [Cho19, Proposition 3.2.7. and Section 3.3.2.]. The next proposition shows that the conditional preorder coincides with the pointwise order.

Proposition 3.10 *For any two parallel morphisms $f, g : (X, \sigma_X) \to (Y, \sigma_Y)$ in $\mathsf{BorelStoch}_{\leq 1}$, $f \sqsubseteq g$ if and only if $f \leq g$.*

Proof. $(\Rightarrow)$ Immediate from Proposition 3.6.

$(\Leftarrow)$ Let $f \leq g$ pointwise. The Radon-Nikodým Theorem for kernels (see e.g. [VO18, Theorem 10]) states that there exists an $r : (Y, \sigma_Y) \times (X, \sigma_X) \to [0, \infty]$ such that $f(A \mid x) = \int_A r(y, x) \cdot g(dy|x)$ for all measurable sets $A \in \sigma_y$.

It only remains to show that one can pick r to be $[0, 1]$-valued. Fix an $x \in X$, and let $B_x = \{y \in Y :\mid r(y, x) > 1\}$. This B_x is clearly measurable. We show that $g(B_x|x) = 0$. By assumption,

$$\int_{B_x} r(y, x) \cdot g(dy \mid x) = f(B_x \mid x) \leq g(B_x \mid x) = \int_{B_x} 1 \cdot g(dy \mid x).$$

By definition of B_x, the converse inequality also holds, so $\int_{B_x} r(y,x) \cdot g(dy \mid x) = \int_{B_x} 1 \cdot g(dy \mid x)$. Let $q_x(y) = (r(y,x) - 1) \cdot \chi(y \in B_x)$. Then $q \geq 0$, and also $\int_{B_x} q_x(y) \cdot g(dy \mid x) = 0$. Therefore, q_x is $g(- \mid x)$ -almost-surely zero, and thus $r(-,x) \leq 1$ almost surely. We now set

$$p(y,x) = \begin{cases} r(y,x) & \text{if } r(y,x) \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

This p then satisfies $0 \leq p \leq 1$, and $f(A \mid x) = \int_A p(y,x) \cdot g(dy \mid x)$ for all $A \in \sigma_Y$.

Finally, we check that p is jointly measurable. This is easy, since for any measurable $T \subseteq [0,1]$ we have

$$p^{-1}(T) = \begin{cases} r^{-1}(T) & \text{if } 0 \notin T, \\ r^{1}(T) \cup r^{-1}((1,\infty)) & \text{otherwise.} \end{cases}$$

□

Cartesian restriction categories

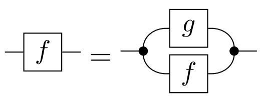
We now instantiate Definition 3.1 in cartesian restriction categories. Such categories have a canonical poset enrichment [CL02, §2.1.4], where $f \leq g$ if and only if $f \preceq g$. This means that f is a restriction of g on a smaller domain. Since all morphisms are deterministic, cartesian restriction categories satisfy the sharp witness property.

Proposition 3.11 *In a cartesian restriction category, the normalisation order and the restriction order coincide: $f \preceq g$ if and only if $f \sqsubseteq g$.*

Proof. Since cartesian restriction categories are balanced and satisfy the sharp witness property, this is immediate from Proposition 3.9. □

Cartesian bicategories of relations

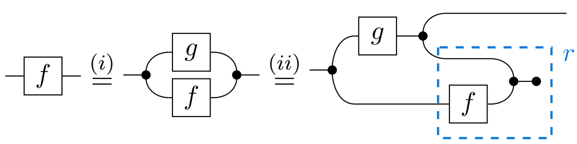
Let us instantiate Definition 3.1 in cartesian bicategories of relations. These also have a canonical poset enrichment where $f \leq g$ means that f is a 'subrelation' of g, in the following sense:

$f \leq g$ if and only if  . (1)

This presentation of the poset enrichment is equivalent to the usual one [Nes24, Lemma 4.1.5]. The next proposition relates the two orderings.

Proposition 3.12 *In a cartesian bicategory of relations, for two parallel morphisms $f, g \colon X \to A$, $f \leq g$ if and only if $f \sqsubseteq g$.*

Proof. ($\Rightarrow$) If $f \leq g$ (i), then we may take $r = (\mathrm{id} \otimes f) ⨟ \mu ⨟ \varepsilon$ and obtain that $f \sqsubseteq g$ by the Frobenius axioms (ii).



($\Leftarrow$) We apply Proposition 3.6 because ($\leq$) is also subunital. □

Every morphism in a cartesian bicategory of relations is self-normalising. However, since these categories are not balanced, the normalisation order may differ from the conditional preorder. The following example demonstrates that this is indeed the case.

Example 3.13 Let $R, S \subseteq \mathbb{N} \times \mathbb{N}$ be morphisms in Rel, given by $R = \{(1,2)\}$ and $S = \{(1,2),(1,3)\}$. Then $R \leq S$, but $R \not\preceq S$, since $\delta ⨟ (R \otimes (S ⨟ \varepsilon)) = R \neq S$.

4 Least conditionals

Asking partial Markov categories to have unique conditionals is too strong. In any partial Markov category with unique conditionals, $f \fatsemi \varepsilon_Y = g \fatsemi \varepsilon_Y$ implies $f = g$ for all parallel maps $f, g \colon X \to Y$ [DR23]. This is similar to how Markov categories with unique conditionals are posetal

However, once we have an order in partial Markov categories, there might be a canonical choice of conditional, namely the smallest among all conditionals. In this section, we consider least conditionals with respect to two orders on self-normalising maps: the conditional preorder $\sqsubseteq$ and the normalisation order $\preceq$. We demonstrate how least conditionals are related to the existence of comparators.

Definition 4.1 (Least conditionals) Let $f : X \to Y \otimes Z$ be a morphism in a (locally small) partial Markov category. Write $C_f = \{c \colon Z \otimes X \to Y \mid f = (f \fatsemi (\mathrm{id} \otimes \varepsilon)) \triangleleft c\}$ for the set of conditionals of f. Now, we say that a conditional, $g \in C_f$, is

(i) a $\sqsubseteq$-least conditional of f when it is a least element of C_f, ordered by the conditional preorder, $(\sqsubseteq)$;

(ii) or, a $\preceq$-least conditional of f if it is the least element of C_f ordered by the normalisation order $(\preceq)$.

By Lemma 3.7, every $\preceq$-least conditional is a $\sqsubseteq$-least conditional.

4.1 Least conditionals and copy-discard-compare categories

We start by characterising conditionals of copy and identity maps in a copy-discard category.

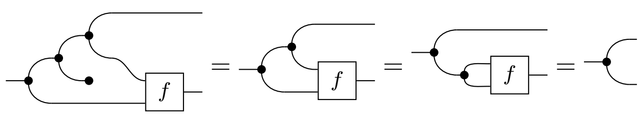
Lemma 4.2 *Let* C *be a copy-discard category.*

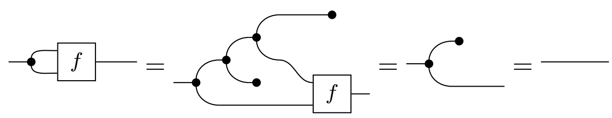
(i) *A morphism* $f \colon X \otimes X \to X$ *is a conditional of the copy* $\delta_X \colon X \to X \otimes X$ *if and only if* $\delta_X \fatsemi f = \mathrm{id}_X$.

(ii) *A morphism* $g \colon X \otimes X \to I$ *is a conditional of the identity* $\mathrm{id}_X \colon X \to X$ *if and only if* $\delta_X \fatsemi g = \varepsilon_X$.

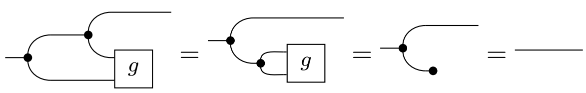
Note that the condition of Item (i) holds by definition for the multiplication μ_X. *Similarly, the definition of the cap* $(\cap_X)$ *contains the condition of Item (ii).*

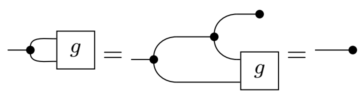
Proof. For point (i), suppose that $\delta_X \fatsemi f = \mathrm{id}_X$. Then, the defining equation of conditionals can be derived as below left. Conversely, if f is a conditional of δ_X, then the middle equality on the right holds by definition. Is then easy to see that $\delta_X \fatsemi f = \mathrm{id}_X$.





For point (ii), suppose that $\delta_X \fatsemi g = \varepsilon_X$. Then, the defining equation of conditionals can be derived as below left. Conversely, discarding the output on both sides of the equality on the left below shows $\delta_X \fatsemi g = \varepsilon_X$, as below right.

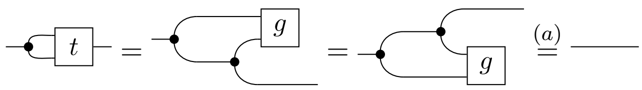


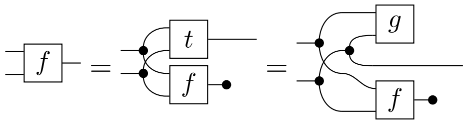


□

Lemma 4.3 *Let* C *be a copy-discard category. If* $f \colon X \otimes X \to X$ *is a conditional of the copy* $\delta_X \colon X \to X \otimes X$*, then* $f \fatsemi \varepsilon_X \colon X \otimes X \to I$ *is a conditional of the identity map* $\mathrm{id}_X \colon X \to X$. *Moreover, if* f *is the* $\preceq$*-least conditional of* δ_X *then* $f \fatsemi \varepsilon_X$ *is the* $\preceq$*-least conditional of* id_X.

Proof. By Lemma 4.2, $f \fatsemi \varepsilon_X$ is a conditional of the identity. Now assume that f is the $\preceq$-least conditional of the copy map, and let $g \colon X \otimes X \to I$ be a conditional of the identity. We have to show that $f \fatsemi \varepsilon_X \preceq g$. Let $t = (\mathrm{id}_X \otimes \delta_X) \fatsemi (g \otimes \mathrm{id}_X)$. Step (a) below left uses that g is a conditional of the identity to show that this t is a conditional of the copy map δ_X. Since f is the $\preceq$-least conditional of the copy map, δ_X, then $f \preceq t$ (below right).



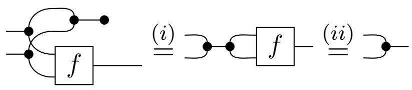


Discarding both sides of the above equation yields $f \fatsemi \varepsilon_X \preceq g$. □

We are ready to show that, if they exist, comparators and caps are least conditionals.

Theorem 4.4 *Let C be a copy-discard-compare category. The multiplication $\mu_X : X \otimes X \to X$ is the $\preceq$-least conditional of the copy map $\delta_X : X \to X \otimes X$. The cap $\cap_X : X \otimes X \to I$ is the least conditional of the identity $\mathrm{id}_X : X \to X$ with respect to the $\preceq$ order. By Lemma 3.7, μ_X and $\cap_X$ are also $\sqsubseteq$-least conditionals.*

Proof. By Lemma 4.2, μ_X and $\cap_X$ are conditionals of the copy and identity maps, respectively. We show that μ_X is the $\preceq$-least conditional of δ_X. Let $f : X \otimes X \to X$ be any conditional of the copy map. Using *(i)* the Frobenius equations and *(ii)* Lemma 4.2, we show that $f \preceq \mu_X$.



By Lemma 4.3, the cap is the $\preceq$-least conditional of identity. □

Since $\preceq$-least conditionals are unique, the following is an immediate consequence of the previous proposition.

Corollary 4.5 *A copy-discard category can admit at most one copy-discard-compare structure.*

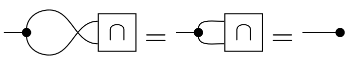
Next, we prove a converse to Theorem 4.4. We demonstrate that, under certain conditions, least conditionals of copier maps give rise to a copy-discard-compare category structure. As it turns out, one needs to consider the normalisation order $\preceq$ to prove the following proposition.

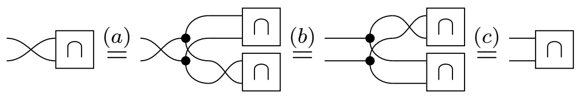
Theorem 4.6 *Let C be a copy-discard category, and assume the following.*

- (i) *All copy maps $\delta_X : X \to X \otimes X$ have a $\preceq$-least conditional $\mu_X : X \otimes X \to X$. By Lemma 4.3, the identity maps then also have $\preceq$-least conditionals $\cap_X = \mu_X \fatsemi \varepsilon_X : X \otimes X \to I$.*
- (ii) *The $\cap_X$ maps satisfy Item (iv) of Definition 2.10.*

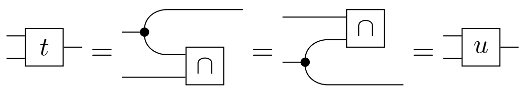
In this situation, the μ_X maps make C a copy-discard-compare category.

Proof. We show that the equations (i)-(iii) of Definition 2.10 are satisfied. The equality (ii) $\delta_X \fatsemi \cap_X = \varepsilon_X$ holds by Lemma 4.2. We show commutativity. Using commutativity of the copy map and equality (i), we obtain that $\delta_X \fatsemi \sigma_{X,X} \fatsemi \cap_X = \varepsilon_X$ also holds (below left). By Lemma 4.2, this implies that $\sigma_{X,X} \fatsemi \cap_X$ is also a conditional of id_X. Therefore, $\cap_X \preceq \cap_X \circ \sigma_{X,X}$. We use this fact in steps (a) and (c) below right. Step (b) applies naturality of the symmetries.

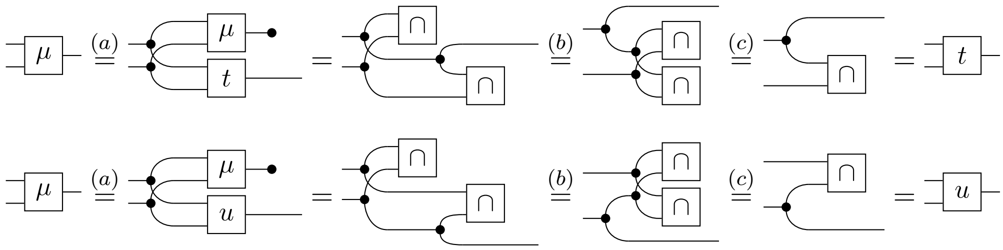




At last, we need to show equation (iii), whose two sides we name t and u.



Using coassociativity of copy, equality (i) and counitality of copy, one can easily show that $\delta ⨾ t = \mathrm{id} = \delta ⨾ u$. By Lemma 4.2, we obtain that both t and u are conditionals of the copy $\delta_X : X \to X \otimes X$. We now prove that $t = \mu_X = u$. Steps marked with (a) use minimality of μ_X. Equations (b) use associativity and commutativity of the copier. Steps (c) use that $\cap_X$ is self-normalising.



This shows that C is a copy-discard category with caps, which, by Proposition 2.11, is also a copy-discard-compare category. □

The following is an immediate consequence of the previous theorem and Proposition 3.9.

Corollary 4.7 *Let C be a balanced partial Markov category that satisfies the sharp witness property. Assume that the copier and identity maps have $\sqsubseteq$-least conditionals μ_X and $\cap_X$, and that the $\cap_X$ maps satisfy Item (iv) of Definition 2.10. Then C is a discrete partial Markov category.*

4.2 Examples

Proposition 4.8 *The category $\mathsf{FinStoch}_{\leq 1}$ is balanced and satisfies the sharp witness property. Therefore, $\preceq$-least conditionals are the same as $\sqsubseteq$-least conditionals.*

Given $f : X \to A \otimes B$, write $\|f(- \mid x)\| = \sum_{b' \in B} f(a, b' \mid x)$ for the total mass of the subdistribution $f(- \mid x)$. The least conditional $c_0 : A \otimes X \to B$ f f is given by

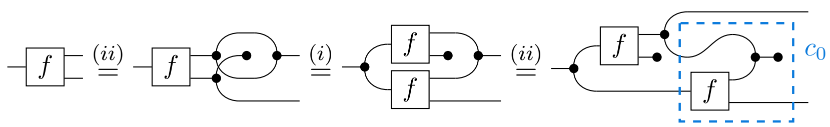
$$c_0(b \mid a, x) = \begin{cases} \frac{f(a,b|x)}{\sum_{b' \in B} f(a,b'|x)} & \text{if } \|f(- \mid x)\| \neq 0, \\ 0 & \text{if } \|f(- \mid x)\| = 0. \end{cases}$$

Let $x \in X$ such that $\|f(- \mid x)\| \neq 0$. Then for any conditional $c : A \otimes X \to B$ of f, the value $c(b \mid a, x)$ is forced to be equal to $c_0(b \mid a, x)$. If $\|f(- \mid x)\| = 0$, then $c(- \mid a, x)$ could be any subdistribution. The least conditional is the solution that contains no 'junk information'.

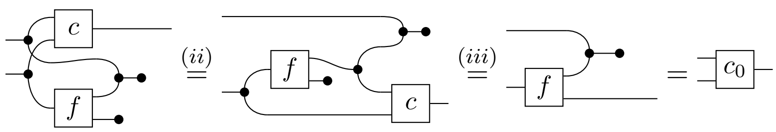
Proposition 4.9 *By Proposition 3.11, $\preceq$-least conditionals are the same as $\sqsubseteq$-least conditionals in a discrete cartesian restriction category.*

Any discrete cartesian restriction category has least conditionals. Given a map $f : X \to A \otimes B$, the least conditional of f is given by $c_0 = (\mathrm{id} \otimes f) ⨾ ((\mu ⨾ \varepsilon) \otimes \mathrm{id})$.

Proof. We check that this is indeed a self-normalising conditional. It is self-normalising because every morphism is deterministic (i). We check the conditional equation using the partial Frobenius axioms (ii).



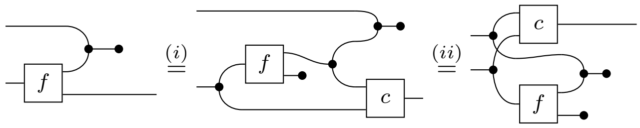
Suppose there is another conditional $c : A \otimes X \to B$ of f (iii).



Then, $c_0 \leq c$. □

Proposition 4.10 *Any cartesian bicategory of relations has $\preceq$-least conditionals.*

Proof. By Proposition 2.13, the morphism $c_0 = (\mathrm{id} \otimes f) ⨾ ((\mu ⨾ \varepsilon) \otimes \mathrm{id})$ is a self-normalising conditional of $f\colon X \to A \otimes B$. Suppose there is another self-normalising conditional, $c\colon A \otimes X \to B$, of f.



We use that (i) c is a conditional of f, and (ii) the Frobenius axioms of cartesian bicategories of relations to show that $c_0 \preceq c$. □

5 Updates increase validity

In its most general form, Bayesian updating involves a prior distribution and a (fuzzy) evidence predicate. From these inputs, one can compute a posterior distribution by incorporating the evidence into the prior.

Given a distribution σ and a predicate p, the validity of p in σ is a real number in the unit interval $[0, 1]$. This validity expresses the probability that the predicate p holds for a random sample from σ.

A key property of Bayesian updating is that the validity of the evidence is greater or equal in the posterior than in the prior. A proof of this fact for discrete probability can be found in [Jac19, Theorem 2.1]. Ultimately, the result is a consequence of the well-known Cauchy–Schwarz inequality. The *Cauchy–Schwarz inequality* states that the square of the inner product of two vectors $u, v \in \mathbb{R}^n$ is smaller than the product of their squared norms:

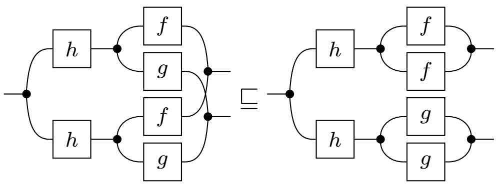
$$\left(\sum\nolimits_i u_i \cdot v_i\right)^2 \leq \left(\sum\nolimits_i u_i^2\right) \cdot \left(\sum\nolimits_j v_j^2\right).$$

The inequality also holds if we allow parameters x, y and z, thus considering families of vectors $h : X \to \mathbb{R}^n$ $f : Y \to \mathbb{R}^n$, and $g : Z \to \mathbb{R}^n$. If we write $f_i(x)$ for $f(x)_i$ (similarly for g and h) and let $u_i = \sqrt{h_i(x)} \cdot f_i(y)$ and $v_j = \sqrt{h_j(x)} \cdot g_j(z)$, the Cauchy–Schwarz inequality becomes:

$$\left(\sum\nolimits_i h_i(x) \cdot f_i(y) \cdot g_i(z)\right)^2 \leq \left(\sum\nolimits_i h_i(x) \cdot f_i(y)^2\right) \cdot \left(\sum\nolimits_j h_j(x) \cdot g_j(z)^2\right). \tag{2}$$

We consider this general form and express it synthetically, using the conditional preorder.

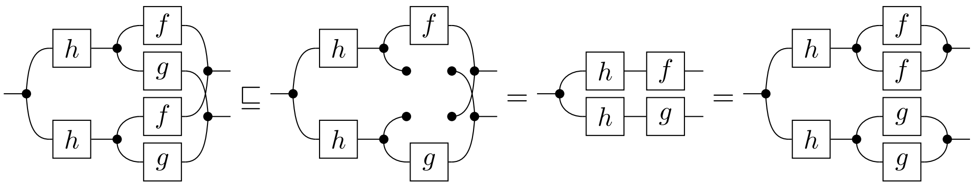
Definition 5.1 (Cauchy–Schwarz inequality) A discrete partial Markov category satisfies the Cauchy–Schwarz inequality if, for all triples of morphisms $h\colon X \to A$, $f\colon A \to Y$, and $g\colon A \to Z$, the following inequality holds.



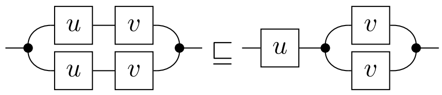
Example 5.2 Instantiating the above definition in $\mathsf{FinStoch}_{\leq 1}$ yields precisely the usual *Cauchy–Schwarz inequality* (2). Therefore, $\mathsf{FinStoch}_{\leq 1}$ satisfies the synthetic Cauchy–Schwarz inequality.

Example 5.3 The Cauchy–Schwarz inequality holds as equality in discrete cartesian restriction categories.

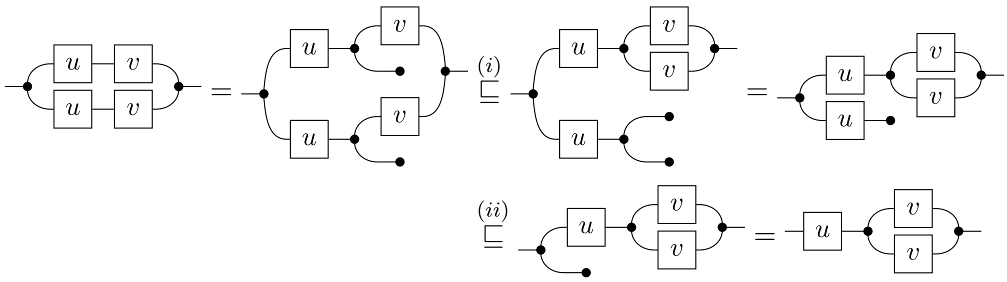
Example 5.4 Cartesian bicategories of relations satisfy the Cauchy–Schwarz inequality.



Proposition 5.5 *A partial Markov category that satisfies the Cauchy–Schwarz inequality also satisfies the inequality below, called the means inequality, for all* $u\colon X \to A$ *and* $v\colon A \to Y$.



Proof. Let $h = u$, $f = v$ and $g = \varepsilon$. We apply *(i)* the Cauchy–Schwarz inequality and *(ii)* Proposition 3.5.



□

Remark 5.6 (QM-AM inequality) The means inequality owes its name to the inequality between the weighted arithmetic and quadratic means of a vector of positive real numbers, $x_1, \dots, x_n$ with weights $w_1 + \cdots + w_n = 1$.

$$\left(\sum_{i \in I} w_i \cdot x_i\right)^2 \leq \sum_{i \in I} w_i \cdot x_i^2.$$

Example 5.7 The partial Markov category of substochastic channels, $\mathsf{FinStoch}_{\leq 1}$, satisfies the means inequality. In particular, if f is a state $\sigma : I \to X$, and g is a predicate $p : X \to I$, then the means inequality gives the well-known inequality between expectations of random variables, usually written as $E[X]^2 \leq E[X^2]$.

With this setup, we can work towards a synthetic analogue of the validity-increase theorem. We start by providing the relevant definitions.

Definition 5.8 (Validity) In a partial Markov category, the *validity* of a predicate $p\colon X \to I$ with respect to a prior $\sigma\colon I \to X$ is the scalar obtained by composition, $(\sigma \mathbin{⨟} p)$.

Definition 5.9 (Bayesian update) Let $\sigma : I \to X$, $p : X \to I$. The *bayesian update* of σ with evidence p is defined to be the bayesian inverse $p^{\dagger}_{\sigma} : I \to X$.

We first prove a special case of the validity increase theorem for deterministic evidence.

Proposition 5.10 *Let* $\sigma : I \to X$, $p : X \to I$ *in a partial Markov category such that* p *is deterministic. Then updating a prior state with a predicate increases the predicate's validity in the posterior. That is,*

$$\sigma \mathbin{⨟} p \sqsubseteq p^{\dagger}_{\sigma} \mathbin{⨟} p.$$

Proof. The conditional equation for the Bayesian inverse is $\sigma \mathbin{⨟} \delta_X \mathbin{⨟} (p \otimes \mathrm{id}_X) = (\sigma \mathbin{⨟} p) \otimes p^{\dagger}_{\sigma}$. Therefore, $\sigma \mathbin{⨟} \delta_X \mathbin{⨟} (p \otimes \mathrm{id}_X) \sqsubseteq p^{\dagger}_{\sigma}$; and, as a consequence, $\sigma \mathbin{⨟} p = \sigma \mathbin{⨟} \delta_X \mathbin{⨟} (p \otimes p) \sqsubseteq p^{\dagger}_{\sigma} \mathbin{⨟} p$. □

For the case where p is not necessarily deterministic, we will need the means inequality and the following notion of zero scalars.

Definition 5.11 Let C be a copy-discard category.

(i) A scalar $s : I \to I$ is called a *zero scalar*, if it satisfies $s \otimes f = s \otimes g$ for all $f, g : X \to Y$. Any scalar that does not satisfy this property is said to be *non-zero*.

(ii) We say that the non-zero scalars in C are *cancellative* if $s \otimes f = s \otimes g$ implies $f = g$ for all $f, g : X \to Y$ and non-zero scalar $s : I \to I$.

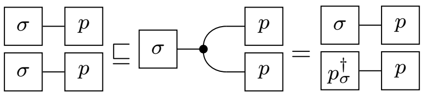
Lemma 5.12 *Let C be a copy-discard category whose non-zero scalars are cancellative. Then, the following cancellation property holds for the conditional inequality. If $s : I \to I$ is a non-zero scalar and $f, g : X \to Y$, then $s \otimes f \sqsubseteq s \otimes g$ implies $f \sqsubseteq g$.*

Proof. Let $s \otimes f \sqsubseteq_r s \otimes g$. Then, by definition $f \otimes s = (s \otimes g) \triangleleft r = s \otimes (g \triangleleft r)$. By cancellativity, $f = g \triangleleft r$, and thus $f \sqsubseteq_r g$. □

Theorem 5.13 *Let C be a partial Markov category that satisfies the means inequality and whose non-zero scalars are cancellative. Then, updating a prior state with a predicate increases the predicate's validity in the posterior. That is, for all $\sigma : I \to X$ and $p : X \to I$,*

$$\sigma \fatsemi p \sqsubseteq p^{\dagger}_{\sigma} \fatsemi p.$$

Proof. We apply the means inequality in the following calculation.



Since the non-zero scalars are cancellative, we obtain that $\sigma \fatsemi p \sqsubseteq p^{\dagger}_{\sigma} \fatsemi p$, because the scalar is either zero or cancellative. □

6 Concluding remarks

We introduced a novel preorder enrichment in the general setting of partial Markov categories; we showed that it generalizes the canonical preorder enrichments of two categories of probability kernels, cartesian restriction categories, and cartesian bicategories of relations. We showed how comparators relate to least conditionals of copier maps. We axiomatized the Cauchy–Schwarz inequality and the quadratic-arithmetic mean inequality. From these, we abstractly derived the fact that Bayesian updating increases the validity of the evidence in the posterior.

Further exploration of the basic theory of the conditional inequality is warranted. One interesting question is about antisymmetry of the order, which is related to inverses of scalars. An inverse of $s : I \to I$ is a scalar $s^{-1} : I \to I$ such that $s \otimes s^{-1} = \varepsilon_I$. Clearly, conditional inequality is antisymmetric only when the only scalar with an inverse is ε_I; it is an open question whether the converse statement holds. It would also be interesting to find necessary conditions for the conditional preorder to form an enrichment. Relatedly, the synthetic formulation of the Cauchy–Schwarz inequality and the means inequality opens the way for a more refined axiomatization: can we derive both from some more fundamental principle?

Finally, a potential application of the synthetic means inequality lies in mixing it with the additive structure of some partial Markov categories, like effectuses with copiers [CJWW15]. In that setting it may be possible to develop a synthetic theory of variance and covariance.

Acknowledgements

The authors want to thank Dario Stein for providing the example of standard Borel spaces (§3.3). We thank the anonymous reviewers for the detailed feedback.

Funding: Mario Román was supported by the Air Force Office of Scientific Research (AFOSR) award number FA9550-21-1-0038. Elena Di Lavore, Mario Román, and Paweł Sobociński were supported by the Advanced Research + Invention Agency (ARIA) Safeguarded AI Programme. Paweł Sobociński was additionally supported by European Union and Estonian Research Council via grants PRG1215 and TEM-TA5 and the Estonian Center of Excellence in Artificial Intelligence.

References

[AFK+24] Nathanael L. Ackerman, Cameron E. Freer, Younesse Kaddar, Jacek Karwowski, Sean K. Moss, Daniel M. Roy, Sam Staton, and Hongseok Yang. Probabilistic programming interfaces for random graphs: Markov categories, graphons, and nominal sets. *Proc. ACM Program. Lang.*, 8(POPL):1819–1849, 2024.

[CFG+24] Leihao Chen, Tobias Fritz, Tomás Gonda, Andreas Klingler, and Antonio Lorenzin. The Aldous-Hoover theorem in categorical probability. *CoRR*, abs/2411.12840, 2024.

[CG99] Andrea Corradini and Fabio Gadducci. An Algebraic Presentation of Term Graphs, via GS-Monoidal Categories. *Applied Categorical Structures*, 7(4):299–331, 1999.

[CG14] J. Robin B. Cockett and Richard Garner. Restriction categories as enriched categories. *Theor. Comput. Sci.*, 523:37–55, 2014.

[Cho19] Kenta Cho. Effectuses in categorical quantum foundations. *CoRR*, abs/1910.12198, 2019.

[CJ19] Kenta Cho and Bart Jacobs. Disintegration and Bayesian Inversion via String Diagrams. *Mathematical Structures in Computer Science*, pages 1–34, March 2019.

[CJWW15] Kenta Cho, Bart Jacobs, Bas Westerbaan, and Abraham Westerbaan. An introduction to effectus theory. *CoRR*, abs/1512.05813, 2015.

[CL02] J. Robin B. Cockett and Stephen Lack. Restriction categories I: categories of partial maps. *Theoretical Computer Science*, 270(1-2):223–259, 2002.

[CL03] J. Robin B. Cockett and Stephen Lack. Restriction categories II: partial map classification. *Theoretical Computer Science*, 294(1-2):61–102, 2003.

[CL07] Robin Cockett and Stephen Lack. Restriction categories III: colimits, partial limits and extensivity. *Mathematical Structures in Computer Science*, 17(4):775–817, 2007.

[CW87] A. Carboni and R.F.C. Walters. Cartesian bicategories i. *Journal of Pure and Applied Algebra*, 49(1):11–32, 1987.

[DR23] Elena Di Lavore and Mario Román. Evidential decision theory via partial Markov categories. In *2023 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–14, 2023.

[FGL+23] Tobias Fritz, Tomás Gonda, Antonio Lorenzin, Paolo Perrone, and Dario Stein. Absolute continuity, supports and idempotent splitting in categorical probability. *CoRR*, abs/2308.00651, 2023.

[FGL+25] Tobias Fritz, Tomáš Gonda, Antonio Lorenzin, Paolo Perrone, and Areeb Shah Mohammed. Empirical measures and strong laws of large numbers in categorical probability, 2025.

[FK23] Tobias Fritz and Andreas Klingler. The d-separation criterion in categorical probability. *J. Mach. Learn. Res.*, 24:46:1–46:49, 2023.

[Fri20] Tobias Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. *Advances in Mathematics*, 370:107239, 2020.

[Jac18] Bart Jacobs. From probability monads to commutative effectuses. *J. Log. Algebraic Methods Program.*, 94:200–237, 2018.

[Jac19] Bart Jacobs. Learning along a channel: the expectation part of expectation-maximisation. *Electronic Notes in Theoretical Computer Science*, 347:143–160, 2019. Proceedings of the Thirty-Fifth Conference on the Mathematical Foundations of Programming Semantics.

[Jac21] Bart Jacobs. Multinomial and hypergeometric distributions in markov categories. In Ana Sokolova, editor, *Proceedings 37th Conference on Mathematical Foundations of Programming Semantics, MFPS 2021, Hybrid: Salzburg, Austria and Online, 30th August - 2nd September, 2021*, volume 351 of *EPTCS*, pages 98–115, 2021.

[JKZ21] Bart Jacobs, Aleks Kissinger, and Fabio Zanasi. Causal inference via string diagram surgery: A diagrammatic approach to interventions and counterfactuals. *Mathematical Structures in Computer Science*, 31(5):553–574, 2021.

[Joh14] Peter T Johnstone. *Topos theory.* Courier Corporation, 2014.

[Lac10] Stephen Lack. *A 2-Categories Companion*, pages 105–191. Springer New York, New York, NY, 2010.

[LT23] Robin Lorenz and Sean Tull. Causal models in string diagrams. *CoRR*, abs/2304.07638, 2023.

[Nes24] Chad Nester. *Partial and Relational Algebraic Theories.* PhD thesis, Tallinn University of Technology, 2024.

[Per24] Paolo Perrone. Markov categories and entropy. *IEEE Trans. Inf. Theory*, 70(3):1671–1692, 2024.

[Sel10] Peter Selinger. A survey of graphical languages for monoidal categories. In *New structures for physics*, pages 289–355. Springer, 2010.

[SS21] Dario Stein and Sam Staton. Compositional semantics for probabilistic programs with exact conditioning. In *2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–13. IEEE, 2021.

[Ste21] Dario Maximilian Stein. Structural foundations for probabilistic programming languages. *University of Oxford*, 2021.

[VO18] Matthijs Vákár and Luke Ong. On s-finite measures and kernels, 2018.

[YZ22] Yimu Yin and Jiji Zhang. Markov categories, causal theories, and the do-calculus, 2022.

A Additional diagrams

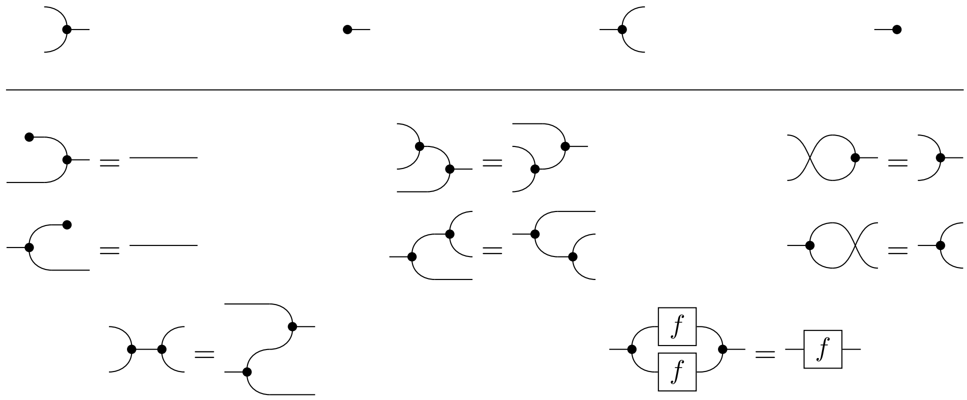


Fig. A.1. Structure and axioms of cartesian bicategories of relations.