

A Categorical Account of the Metropolis–Hastings Algorithm

Rob Cornish   

College of Computing and Data Science, Nanyang Technological University, Singapore

Andi Q. Wang   

Department of Statistics, University of Warwick, Coventry, UK

Abstract

Metropolis–Hastings (MH) is a foundational Markov chain Monte Carlo (MCMC) algorithm. In this paper, we ask whether it is possible to formulate and analyse MH in terms of categorical probability, using a recent involutive framework for MH-type procedures as a concrete case study. We show how basic MCMC concepts such as invariance and reversibility can be formulated in Markov categories, and how one part of the MH kernel can be analysed using standard CD categories. To go further, we then study enrichments of CD categories over commutative monoids. This gives an expressive setting for reasoning abstractly about a range of important probabilistic concepts, including substochastic kernels, finite and σ -finite measures, absolute continuity, singular measures, and Lebesgue decompositions. Using these tools, we give synthetic necessary and sufficient conditions for a general MH-type sampler to be reversible with respect to a given target distribution.

2012 ACM Subject Classification Mathematics of computing $\rightarrow$ Bayesian computation; Theory of computation $\rightarrow$ Categorical semantics

Keywords and phrases Categorical probability, Metropolis–Hastings, Markov category, CD category

Digital Object Identifier 10.4230/LIPIcs.LICS.2026.32

Related Version *Full Version*: <https://arxiv.org/abs/2601.22911>

Funding *Rob Cornish*: Supported by the National Research Foundation, Singapore, under a National Research Foundation Fellowship in Artificial Intelligence (Award No.: NRFFIAI1-2024-0014).

Andi Q. Wang: Supported by a Discipline Hopping Award from the Prob_AI Hub (EP/Y028783/1).

Acknowledgements We are very grateful to Sam Power for many useful discussions throughout the course of the project. A big thanks to Sam Staton, Paolo Perrone, Tobias Fritz, Pedro Amorim, and Luke Ong for their support and helpful suggestions. Many thanks also to the three anonymous referees for their constructive comments which improved the manuscript.

1 Introduction

Markov kernels are a fundamental concept in probability theory. Recall that a Markov kernel $P : \mathcal{X} \rightarrow \mathcal{Y}$ encodes a family of probability distributions $P(x, dy)$ on the space $\mathcal{Y}$ indexed by points $x \in \mathcal{X}$. Recent work on *copy-discard (CD) categories* and *Markov categories* [7, 11] has sought to develop a synthetic framework for reasoning about systems of Markov kernels and their unnormalised counterparts algebraically. This approach has been applied successfully to a number of classical results from probability and statistics, yielding streamlined and more conceptual proofs than traditional measure-theoretic arguments [14, 29, 11, 27].

In applied settings, Markov kernels are often regarded as “stochastic functions” and used to model *conditional probability distributions*, which in turn have a huge number of applications in computational statistics, machine learning, and computer science generally. It is therefore natural to ask: to what extent are Markov and CD categories ready for reasoning about contemporary methodology in these domains? Are they usable “as-is”, or are additional structures or concepts needed to handle the complexities of modern algorithms?



© Rob Cornish and Andi Q. Wang;

licensed under Creative Commons License CC-BY 4.0

41st Annual Symposium on Logic in Computer Science (LICS 2026).

Editors: Claudia Faggian and Joost-Pieter Katoen; Article No. 32; pp. 32:1–32:26

Leibniz International Proceedings in Informatics



Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

We consider this question in the context of *Markov chain Monte Carlo* (MCMC) methods, and specifically the *Metropolis–Hastings* (MH) algorithm [26, 21], which is a foundational tool for sampling from complex probability distributions. We focus on a recent line of work which studies MH through the lens of *involutions* [28, 2, 16, 17, 9], building on earlier foundational work [35, 19]. This perspective has already lead to a range of new methodology of interest to computational statistics practitioners [5, 25, 38].

A major contribution to the involutive approach is the framework of [2]. This considers a Markov kernel $\Pi_{\text{MH}} : \mathcal{E} \rightarrow \mathcal{E}$ defined

$$\Pi_{\text{MH}}(\xi, d\xi') := \alpha(\xi) \delta_{\phi(\xi)}(d\xi') + (1 - \alpha(\xi)) \delta_{\xi}(d\xi'), \quad (1)$$

where $\alpha : \mathcal{E} \rightarrow [0, 1]$ is a measurable function regarded as an *acceptance probability*, $\phi : \mathcal{E} \rightarrow \mathcal{E}$ is a measurable *involution* (i.e. $\phi = \phi^{-1}$), and δ is the Dirac delta. Intuitively, given an input ξ , this kernel transitions to $\phi(\xi)$ with probability $\alpha(\xi)$, and otherwise remains at ξ . By combining Π_{MH} with a suitable *state-space augmentation* (see Section 2.4), it is shown in [2] that a very wide variety of MH-type algorithms can be recovered as special cases (see also Appendix B for two examples).

The main theoretical result of [2] is their Theorem 3, which gives sufficient conditions under which Π_{MH} is *reversible* with respect to a given target distribution. Reversibility in turn constitutes a key correctness guarantee for many MCMC algorithms. We state a summarised version of this result as follows (see Appendix A for our measure theory terminology and notation):

► **Theorem 1.** *Suppose μ is a finite measure on a measurable space $\mathcal{E}$, and $\phi : \mathcal{E} \rightarrow \mathcal{E}$ is a measurable involution. Then there exists a measurable $S \subseteq \mathcal{E}$ such that the restriction $\mu|_S$ and its pushforward $\mu|_S^\phi$ are equivalent as measures, and $\mu|_{S^c}$ and its pushforward $\mu|_{S^c}^\phi$ are singular. The kernel Π_{MH} (1) above is then $\mu|_S$ -reversible if*

$$\alpha(\xi) = \begin{cases} \alpha(\phi(\xi)) r(\xi) & \text{if } \xi \in S, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

where $r := \frac{d\mu|_S^\phi}{d\mu|_S}$ denotes the Radon–Nikodym derivative.

Summary of main results

Given the important implications of Theorem 1 for MH-type procedures, it serves as a natural test case for the applicability of categorical probability. Our goal in this paper is to obtain this result synthetically using Markov and CD categories. For this, we find that three levels of structure are needed:

- In Section 2, we use Markov categories to formulate generic aspects of MCMC algorithms, such as invariance, reversibility and state-space augmentation.
- In Section 3, we use CD categories to reason about the first summand in (1): we prove reversibility of this term, and recover part of the condition (2).
- However, to study the entire kernel Π_{MH} , we need additional structure: for example, we need a way to “add” morphisms, and we need a way to reason about decompositions of morphisms into disjoint “parts”. In Section 4, we therefore consider an *enriched* extension of CD categories. This allows us to provide a synthetic generalisation of Theorem 1 in the form of Theorems 52 and 68 below.

The appendix contains some additional background and proofs, as well as two additional examples of MCMC-related techniques (the systematic-scan Gibbs sampler, and importance sampling) within categorical probability, which may be of interest to practitioners.

We believe these results are interesting in their own right, as they involve arguments that are both more conceptual and more general than their measure-theoretic counterparts. We also hope our work illustrates how categorical tools may be used to analyse contemporary methodology in computational statistics and machine learning more generally.

Technical contributions

In order to consider the full MH kernel Π_{MH} , we study CD categories *enriched* over commutative monoids, which provides a minimal structure for reasoning about addition of morphisms. This enrichment is expressive, and permits formulating and proving results about abstract versions of the following important concepts from measure theory:

- *Substochastic kernels* and *probabilities* (i.e. values in $[0, 1]$).
- *Finite* and *σ -finite* kernels.
- Pointwise *absolute continuity* between kernels.
- *Singular measures* and *Lebesgue decompositions*.

The enrichment combines well with string diagrams, allowing for straightforward graphical reasoning. It also integrates naturally with existing categorical approaches to probabilistic choice: when suitable coproducts exist, the subcategory of normalised morphisms carries the structure of a *distributive Markov category* [1], relating our “algebraic” treatment of convex combinations to the “sampling” perspective.

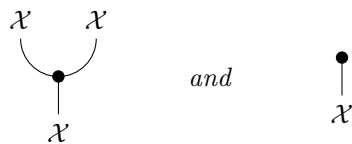
Importantly, we show that these tools are not merely descriptive, but may be used to obtain interesting results about probabilistic concepts. In the context of Theorem 1, this enrichment allows us to reason about decompositions of μ into parts supported on S and S^c , and to obtain the condition (2) via purely algebraic arguments. It also supplies a converse to Theorem 1, which was not given in [2], and permits a straightforward extension to the historically more complex *skew-reversible* setting (Corollary 54). We believe these results suggest this enrichment may be useful in applications beyond MCMC also.

2 MCMC foundations in Markov categories

2.1 Recap: Markov categories

We first recall the fundamentals of *Markov categories*, which will allow us to formulate key Markov chain concepts such as reversibility and invariance. We will assume familiarity with basic definitions from category theory and monoidal categories, as well as string diagrams [22, 32]. For readers unfamiliar with this material, we give some further details and references in Appendix C.

► **Definition 2.** A Markov category is a symmetric monoidal category $(\mathbf{C}, \otimes, \mathcal{I})$ where each object $\mathcal{X}$ in $\mathbf{C}$ is equipped with a commutative comonoid structure $(\text{copy}_{\mathcal{X}}, \text{del}_{\mathcal{X}})$. In string diagrams, we denote $\text{copy}_{\mathcal{X}}$ and $\text{del}_{\mathcal{X}}$ respectively as



These morphisms satisfy certain compatibility conditions with the monoidal structure (see Definition 73 in the Appendix), and the following requirement for all $P : \mathcal{X} \rightarrow \mathcal{Y}$ in $\mathbf{C}$:

$$\begin{array}{c} \bullet \\ | \\ \boxed{P} \\ | \\ \mathcal{X} \end{array} = \begin{array}{c} \bullet \\ | \\ \mathcal{X} \end{array} \quad (3)$$

► **Example 3.** For the purposes of obtaining Theorem 1, the most important example of a Markov category is the category **Stoch**. The objects of this category are measurable spaces, and the morphisms $P : \mathcal{X} \rightarrow \mathcal{Y}$ are *Markov kernels*, namely functions $P : \mathcal{X} \times \Sigma_{\mathcal{Y}} \rightarrow [0, 1]$ such that each function $x \mapsto P(x, A)$ is measurable and each function $A \mapsto P(x, A)$ is a probability measure. Composition is given by the Chapman–Kolmogorov equation:

$$(Q \circ P)(x, B) := \int P(x, dy) Q(y, B)$$

and may be regarded as *sequential composition* of Markov kernels. The monoidal unit $\mathcal{I}$ is the singleton measurable space $\{*\}$, and the monoidal product $\mathcal{X} \otimes \mathcal{Y}$ is given by the usual product of measurable spaces. The monoidal product of morphisms is defined as:

$$(R \otimes Q)((x, y), A \times B) := R(x, A) Q(y, B), \quad (4)$$

which may be regarded as *parallel composition*. The remaining structure of **Stoch** is given in Example 75 in the Appendix.

Morphisms $\pi : \mathcal{I} \rightarrow \mathcal{X}$ out of the monoidal unit $\mathcal{I}$ are regarded as having “no input” and appear in string diagrams as follows:

$$\begin{array}{c} \mathcal{X} \\ | \\ \triangle \pi \end{array}$$

A morphism $\phi : \mathcal{X} \rightarrow \mathcal{Y}$ is *deterministic* [11, Definition 10.1] if

$$\begin{array}{c} \mathcal{Y} \quad \mathcal{Y} \\ \curvearrowright \\ \bullet \\ | \\ \boxed{\phi} \\ | \\ \mathcal{X} \end{array} = \begin{array}{c} \mathcal{Y} \quad \mathcal{Y} \\ | \quad | \\ \boxed{\phi} \quad \boxed{\phi} \\ \curvearrowleft \\ \bullet \\ | \\ \mathcal{X} \end{array} \quad (5)$$

► **Example 4.** In **Stoch**, since $\mathcal{I}$ is the singleton space, a Markov kernel $\mu : \mathcal{I} \rightarrow \mathcal{X}$ encodes a single probability measure on $\mathcal{X}$, which we will denote simply as $\mu(dx)$. Likewise, every measurable function $\phi : \mathcal{X} \rightarrow \mathcal{Y}$ gives rise to a deterministic Markov kernel $\Phi : \mathcal{X} \rightarrow \mathcal{Y}$ defined as

$$\Phi(x, dy) := \delta_{\phi(x)}(dy), \quad (6)$$

where δ is the Dirac delta. We will often identify Φ and ϕ in what follows. This recovers the concept of *pushforwards* from measure theory: recall that the pushforward of a measure μ on $\mathcal{X}$ by a measurable function $\phi : \mathcal{X} \rightarrow \mathcal{Y}$ is the distribution μ^ϕ on $\mathcal{Y}$ with $\mu^\phi(B) := \mu(\phi^{-1}(B))$. In **Stoch**, this is just the composition $\mu^\phi = \phi \circ \mu$, where here the right-hand ϕ is regarded as a deterministic Markov kernel as in (6).

The following concept is fundamental in Bayesian statistics and machine learning, and will be useful for us in formulating *state space augmentation* in Section 2.4.

► **Definition 5** (Page 9, [7]). Let $P : \mathcal{X} \rightarrow \mathcal{Y}$ and $\pi : \mathcal{I} \rightarrow \mathcal{X}$ be morphisms in a Markov category $\mathbf{C}$. Then $P^\dagger : \mathcal{Y} \rightarrow \mathcal{X}$ is a Bayesian inverse of P with respect to π if

$$(7)$$

To explain the term “Bayesian inverse”, note that in **Stoch** the left-hand side of (7) is a Markov kernel we may sample from via

$$X \sim \pi(dx) \quad Y \sim P(X, dy) \quad \text{return } (X, Y).$$

The equation (7) then says that $P^\dagger(y, dx)$ is the conditional distribution of X given $Y = y$, which is equivalently the posterior distribution of X obtained via Bayes’ theorem.

2.2 Invariance and reversibility

Among the most fundamental concepts from Markov chain theory are *invariance* and *reversibility*. In the context of MCMC, these constitute key correctness properties that ensure an algorithm targets the intended distribution. We now show how these concepts can be formulated in Markov categories.

► **Definition 6.** Let $\mu : \mathcal{I} \rightarrow \mathcal{X}$ and $P : \mathcal{X} \rightarrow \mathcal{X}$ be morphisms in a Markov category $\mathbf{C}$. We say that P is μ -invariant if $P \circ \mu = \mu$. We say that P is μ -reversible if

$$(8)$$

Equivalently, this says that the left-hand side of (8) is invariant with respect to the swap isomorphism $\mathcal{X} \otimes \mathcal{X} \xrightarrow{\cong} \mathcal{X} \otimes \mathcal{X}$, or that P is its own Bayesian inverse with respect to μ .

► **Example 7.** Consider Definition 6 with $\mathbf{C} = \mathbf{Stoch}$. Here, $\mathcal{X}$ becomes a measurable space, μ a probability measure on $\mathcal{X}$, and $P : \mathcal{X} \rightarrow \mathcal{X}$ a Markov kernel. Then P is μ -invariant if and only if $\int \mu(dx) P(x, A) = \mu(A)$ for every measurable $A \subseteq \mathcal{X}$, recovering the usual definition of invariance, as observed by [27, Section 2.4]. Intuitively, invariance says that if a Markov chain with transition kernel P is initialised with distribution μ , then its marginal distribution will remain μ at all time steps afterwards. Likewise, (8) says that $\int_A \mu(dx) P(x, B) = \int_B \mu(dx') P(x', A)$ for all measurable $A, B \subseteq \mathcal{X}$, or more colloquially

$$\mu(dx) P(x, dx') = \mu(dx') P(x', dx).$$

This recovers the classical *detailed balance* condition from Markov chain theory, as observed by [13, Remark 4.1.11]. Intuitively, detailed balance says that if a Markov chain with transition kernel P is initialised with distribution μ , then the joint distribution of any two consecutive states is invariant under swapping them.

Invariance is usually the fundamental correctness property sought for MCMC algorithms, since it ensures that the Markov chain will converge to the desired distribution under mild conditions regardless of its initial state [30]. Reversibility is a sufficient condition for invariance that is often convenient to verify in practice. Markov categories are very well-suited for reasoning about these concepts, as the following result illustrates.

► **Proposition 8.** *Let $\mu : \mathcal{I} \rightarrow \mathcal{X}$ and $P, Q : \mathcal{X} \rightarrow \mathcal{X}$ be morphisms in a Markov category $\mathbf{C}$. If both P and Q are μ -invariant, then so is $P \circ Q$. Likewise, if P is μ -reversible, then it is also μ -invariant.*

Proof. For the first claim, by the definition of invariance, we have that $P \circ Q \circ \mu = P \circ \mu = \mu$. For the second claim, observe that:

where the second step uses reversibility, and the others use the basic Markov category axioms (see Appendix C). ◀

It is well known by MCMC practitioners that invariance does not imply reversibility in general. However, this does hold for *involutions*, where recall $\phi : \mathcal{X} \rightarrow \mathcal{X}$ is an involution if $\phi \circ \phi = \text{id}_{\mathcal{X}}$.

► **Proposition 9.** *Let $\mu : \mathcal{I} \rightarrow \mathcal{X}$ and $\phi : \mathcal{X} \rightarrow \mathcal{X}$ be morphisms in a Markov category $\mathbf{C}$, where ϕ is a deterministic involution. If ϕ is μ -invariant, then it is μ -reversible.*

Proof. If ϕ is μ -invariant, then we have the following

where the first step uses the involution property, the second uses determinism, and the third uses invariance. ◀

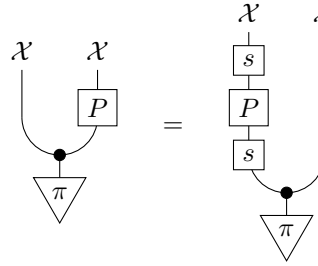
► **Remark 10.** The assumption of determinism in Proposition 9 is redundant in many practical settings. This is because an involution is by definition an isomorphism, and for the class of *positive* Markov categories [11, Definition 11.22], every isomorphism is necessarily deterministic [11, Remark 11.28]. Many Markov categories arising in practice are positive, including **Stoch** [11, Example 11.25].

2.3 Skew-reversibility

Reversible Markov chains often exhibit a diffusive “random walk” behaviour, and reversible MCMC algorithms can often perform poorly in practice as a result [10]. To address this, a variety of generalised notions of reversibility have been considered in the MCMC literature,

including *lifting* [6, 10] and *skew-reversibility* [37, 31, 34]. These all admit related formulations in terms of an involution acting on the state space. Recently, [3, Definition 2.4] gave a general notion of (π, s) -*reversibility* which subsumes these previous concepts, but at the cost of introducing heavier measure-theoretic machinery. We show how these concepts can be streamlined considerably in the context of Markov categories. Later on, this will yield a straightforward extension of our synthetic version of Theorem 1 to the skew-reversible setting (Corollary 54).

► **Definition 11.** Let $\pi : \mathcal{I} \rightarrow \mathcal{X}$ and $s, P : \mathcal{X} \rightarrow \mathcal{X}$ be morphisms in a Markov category $\mathbf{C}$, where s is a deterministic involution that is π -invariant. We say that P is (π, s) -reversible if the following equality holds:


(10)

► **Example 12.** In $\mathbf{C} = \mathbf{Stoch}$, in terms of measure theory, (10) says:

$$\pi(dx) P(x, dy) = \pi(dy) (s \circ P \circ s)(y, dx)$$

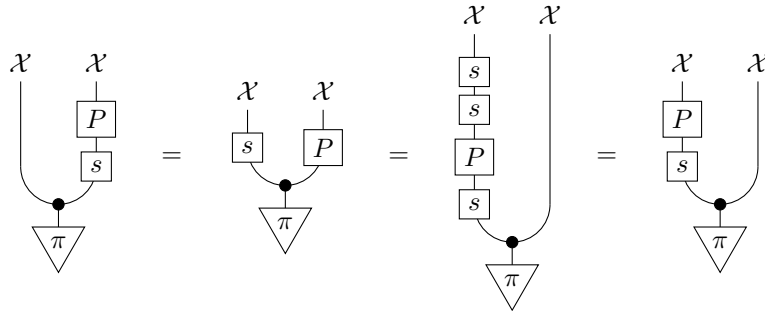
which recovers the definition of (π, s) -reversibility given in [3, Definition 2.4] (see also [2, Proposition 3, conclusion (c)]).

By a similar argument as Proposition 8, (π, s) -reversibility implies π -invariance, which explains its usefulness for MCMC. It is also clear that ordinary π -reversibility is the special case of $(\pi, \text{id}_{\mathcal{X}})$ -reversibility. Interestingly, the following result shows we can also define (π, s) -reversibility as a special case of π -reversibility, and recovers several results on this topic in the MCMC literature (e.g. [3, Proposition 2.5], [34, Equation (3)], [2, Corollary 2]).

► **Proposition 13.** With π , s , and P as in Definition 11, the following are equivalent:

- (1) P is (π, s) -reversible.
- (2) $P \circ s$ is π -reversible.
- (3) $s \circ P$ is π -reversible.
- (4) $P = Q \circ s = s \circ R$ for some π -reversible $Q, R : \mathcal{X} \rightarrow \mathcal{X}$.

Proof. (1) $\Leftrightarrow$ (2): Consider the following equalities:



The first step always holds by Proposition 9, and the third because s is an involution. The middle step holds if and only if P is (π, s) -reversible, as can be seen by applying s to both left-hand wires and using the involution property. Likewise, the left- and right-hand sides are equal precisely when $P \circ s$ is π -reversible, which gives the statement.

(1) $\Leftrightarrow$ (3) follows by a similar argument.

(2) $\Leftrightarrow$ (4) $\Leftrightarrow$ (3) now follows by taking $Q := P \circ s$ and $R := s \circ P$. $\blacktriangleleft$

2.4 Reversibility via augmented state spaces

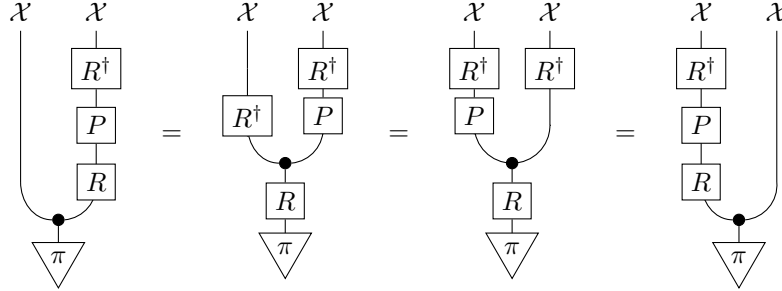
Given a target distribution on a space $\mathcal{X}$, a common MCMC technique is to consider a target on an *augmented space* $\mathcal{E}$ that is related to $\mathcal{X}$, but simpler to work with. Although widely used, augmentation techniques are often presented on a case-by-case basis for specific choices of the components involved. However, Markov categories permit a unified explanation of these techniques as follows.

► **Proposition 14.** *Let $\pi : \mathcal{I} \rightarrow \mathcal{X}$ be a morphism in a Markov category $\mathbf{C}$. Suppose we have a composition in $\mathbf{C}$ as follows:*

$$\mathcal{X} \xrightarrow{R} \mathcal{E} \xrightarrow{P} \mathcal{E} \xrightarrow{R^\dagger} \mathcal{X}. \quad (11)$$

If $R^\dagger$ is a Bayesian inverse of R with respect to π , and P is $(R \circ \pi)$ -reversible (resp. invariant), then (11) is π -reversible (resp. invariant)

Proof. We show the reversibility statement; the invariance statement follows similarly. By basic string diagram manipulations:



where the first and third steps use the definition of a Bayesian inverse (Definition 5), and the second uses reversibility of P . $\blacktriangleleft$

In the context of involutive MH, Andrieu et. al. [2] combine state space augmentation with their kernel (1) to recover a wide variety of existing MCMC algorithms. The scheme they consider involves introducing auxiliary random variables taking values in a space $\mathcal{Z}$, so that the augmented space has the form $\mathcal{E} = \mathcal{X} \otimes \mathcal{Z}$. The following result captures their approach in the setting of Markov categories.

► **Corollary 15.** *Let $\pi : \mathcal{I} \rightarrow \mathcal{X}$, $Q : \mathcal{X} \rightarrow \mathcal{Z}$, and $P : \mathcal{X} \otimes \mathcal{Z} \rightarrow \mathcal{X} \otimes \mathcal{Z}$ be morphisms in a Markov category $\mathbf{C}$, and denote:*

$$\begin{array}{c} \mathcal{X} \quad \mathcal{Z} \\ \downarrow \quad \downarrow \\ \triangle \mu \end{array} := \begin{array}{c} \mathcal{X} \quad \mathcal{Z} \\ \downarrow \quad \downarrow \\ \square Q \\ \downarrow \quad \downarrow \\ \triangle \pi \end{array} \quad \begin{array}{c} \mathcal{X} \\ \downarrow \\ \square \Pi \\ \downarrow \\ \mathcal{X} \end{array} := \begin{array}{c} \mathcal{X} \\ \downarrow \\ \square P \\ \downarrow \quad \downarrow \\ \square Q \\ \downarrow \quad \downarrow \\ \mathcal{X} \end{array} \quad (12)$$

If P is μ -reversible, then Π is π -reversible.

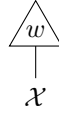
Proof. By defining

$$\begin{array}{c} \mathcal{X} \quad \mathcal{Z} \\ | \quad | \\ \boxed{R} \\ | \quad | \\ \mathcal{X} \end{array} := \begin{array}{c} \mathcal{X} \quad \mathcal{Z} \\ | \quad | \\ \bullet \\ | \quad | \\ \mathcal{X} \quad \mathcal{Z} \end{array} \quad \begin{array}{c} \mathcal{X} \\ | \\ \boxed{R^\dagger} \\ | \quad | \\ \mathcal{X} \quad \mathcal{Z} \end{array} := \begin{array}{c} \mathcal{X} \\ | \\ \mathcal{X} \end{array} \quad \begin{array}{c} \mathcal{Z} \\ | \\ \bullet \\ | \\ \mathcal{Z} \end{array} \quad (13)$$

monoidal product in this setting [33, Section 1.2]. A solution is to consider *s-finite* kernels instead [33] (see also Remark 72 for why *σ -finiteness* is not used). Recall that a kernel K is *s-finite* if it can be written as a countable sum $K = \sum_{n=1}^{\infty} K_n$ where each kernel K_n satisfies $\sup_{x \in \mathcal{X}} K_n(x, \mathcal{Y}) < \infty$ [23, Page 56]. It is shown in [7, Example 7.2] that measurable spaces and *s-finite* kernels form a CD category that we denote by **sfKern**, with its CD structure defined analogously to **Stoch** (Example 3).

► **Remark 19.** The axiom (3) that distinguishes CD categories from Markov categories may be regarded as a *normalisation* condition. For example, in **sfKern**, this condition holds for $K : \mathcal{X} \rightarrow \mathcal{Y}$ if and only if $K(x, \mathcal{Y}) = 1$ for all $x \in \mathcal{X}$. Accordingly, in a general CD category, we say that morphisms satisfying (3) are *normalised*.¹ Given any CD category **C**, the objects of **C** and its normalised morphisms form a Markov category that we call **C_{norm}** [7, Section 7].

In a CD category, morphisms of the form $w : \mathcal{X} \rightarrow \mathcal{I}$ play a special role. Following [7, Section 7], we refer to these morphisms as *effects*, and depict them in string diagrams as



► **Example 20.** Effects in **sfKern** correspond to nonnegative measurable functions [7, Example 7.2]. Explicitly, since $\mathcal{I} \cong \{*\}$ is the singleton space (Example 18), a function $w : \mathcal{X} \times \Sigma_{\mathcal{I}} \rightarrow [0, \infty]$ is a kernel if and only if $x \mapsto w(x, \mathcal{I})$ is measurable. In what follows, we will identify an effect with its corresponding measurable function to streamline notation (so $w(x, \mathcal{I})$ becomes simply $w(x)$).

Effects recover *Radon–Nikodym derivatives* from measure theory. This is made precise via the following definition, which is a special case of the *likelihoods* considered by [7, Definition 8.1].

► **Definition 21.** Let **C** be a CD category, and let $\pi, \mu : \mathcal{I} \rightarrow \mathcal{X}$ and $r : \mathcal{X} \rightarrow \mathcal{I}$ be morphisms in **C**. We call r a Radon–Nikodym derivative of π with respect to μ and write $r = \frac{d\pi}{d\mu}$ if

(15)

► **Example 22.** Consider Definition 21 in the case of **sfKern**. The condition (15) holds if and only if $\pi(A) = \int_A r(x) \mu(dx)$ for all measurable $A \subseteq \mathcal{X}$. This says precisely that r is a Radon–Nikodym derivative of π with respect to μ in the usual measure-theoretic sense.

► **Remark 23.** In Section 2, we gave definitions of *determinism*, *invariance*, and *reversibility* in the setting of Markov categories. In what follows, we will import these concepts directly into the setting of CD categories, noting that they continue to make sense without the assumption that all morphisms are normalised. For example, when we speak of a deterministic morphism in a CD category, we simply mean that the equation (5) holds.² We will also import Propositions 9 and 13, whose proofs do not require normalisation either.

¹ [7, Definition 3.2] refers to such morphisms as *causal*.

² [12] describes such a morphism as *copyable*, and defines “deterministic” to mean *both* copyable and normalised.

3.2 The first summand of Metropolis–Hastings

Recall that our overall goal is to understand when Π_{MH} defined in (1) is reversible with respect to a given target measure μ . In the proof given by Andrieu et al. [2, Theorem 3], considerable effort is spent showing μ -reversibility of the first summand (14). Their argument requires manipulating chains of integrals that do not necessarily shed light on the key underlying structure at play.

CD categories allow us to reason about the summand (14) in a more conceptual way. In string diagrams, this kernel may be written as follows:

$$(16)$$

CD categories also admit the following extension of Proposition 9. Notice this requires Radon–Nikodym derivatives (Definition 21), and so is not directly available in the Markov category setting.

► **Proposition 24.** *Suppose $\mu : \mathcal{I} \rightarrow \mathcal{E}$ and $\phi : \mathcal{E} \rightarrow \mathcal{E}$ are morphisms in a CD category $\mathbf{C}$, where ϕ is a deterministic involution. If a Radon–Nikodym derivative $r = \frac{d(\phi \circ \mu)}{d\mu}$ exists, then*

$$(17)$$

In other words, ϕ is μ -reversible “up to a correction factor of r ”.

Proof. This follows by the same argument as in (9), but now using the Radon–Nikodym derivative r in the third step. ◀

We can now give necessary and sufficient conditions for reversibility of the summand (14) entirely in terms of string diagrams.

► **Theorem 25.** *Suppose $\mu : \mathcal{I} \rightarrow \mathcal{E}$, $\phi : \mathcal{E} \rightarrow \mathcal{E}$, and $\alpha, r : \mathcal{E} \rightarrow \mathcal{I}$ are morphisms in a CD category $\mathbf{C}$, where ϕ is a deterministic involution and $r = \frac{d(\phi \circ \mu)}{d\mu}$. Then (16) is μ -reversible if and only if*

$$(18)$$

Proof. By basic string diagram manipulations, (16) is μ -reversible if and only if ϕ is reversible with respect to the left-hand side of (18). By Proposition 9, this holds if and only if ϕ leaves this morphism invariant, namely:

The diagram shows an equality between two string diagrams. On the left, a triangle labeled α is connected by a curved line to a dot, which is then connected by a straight line to a triangle labeled μ . A vertical line labeled $\mathcal{E}$ extends from the dot. On the right, the same structure is shown, but the vertical line $\mathcal{E}$ passes through a square box labeled ϕ before reaching the dot. The equation is labeled (19) on the right.

By postcomposing both sides of (17) with $\alpha \otimes \text{id}_{\mathcal{E}}$, Proposition 24 now shows that (19) is equivalent to (18). $\blacktriangleleft$

4 Semiadditive CD categories

In the previous section, we considered the first summand of the Metropolis–Hastings kernel (1). However, we did not have a way to reason about the full kernel Π_{MH} . One obvious difficulty is that Π_{MH} is written as a *sum* of two morphisms, which is not something we can express in a general CD category. It is also unclear how to express other concepts required for Theorem 1, such as the rejection probability $1 - \alpha(\mathcal{X})$, or the decomposition of $\mathcal{E}$ into a region S on which the target and its pushforward under ϕ are equivalent measures.

To proceed, we therefore consider a natural *enrichment* [24] of CD categories. Specifically, we consider CD categories whose hom-sets $\mathbf{C}(\mathcal{X}, \mathcal{Y})$ are equipped with the structure of a *commutative monoid*, which allows us to reason about “sums” of morphisms in an abstract way. Enrichment over commutative monoids arises naturally in a range of categorical settings. Previous work [18, 36] has used this enrichment to model summation of processes in classical probabilistic and quantum settings. Related enrichments have also appeared in the context of effectus theory, which considers enrichment over *partial* commutative monoids [8], as well as in cartesian differential categories [4], which may be characterised as *skew enriched* over commutative monoids [15]. In this work, we show that this simple enrichment captures a surprising number of concepts from classical probability theory, allowing us to reason about the full MH kernel and to give a self-contained reversibility proof that recovers Theorem 1 in **sfKern**. The enrichment also interacts well with string diagrams, enabling intuitive graphical reasoning (see Remark 28).

4.1 Basic definitions

Recall that a *commutative monoid* is a set $\mathcal{M}$ equipped with a commutative and associative binary operation, and a distinguished unit element. In other words, a commutative monoid is an abelian group without the requirement of inverses. Standard examples include $(\mathbb{N}, +, 0)$ and $(\mathbb{N}, \times, 1)$, the natural numbers under addition and multiplication respectively. A *homomorphism* between commutative monoids $\mathcal{M}$ and $\mathcal{M}'$ is a function $\varphi : \mathcal{M} \rightarrow \mathcal{M}'$ that preserves the binary operation and unit element. Commutative monoids and homomorphisms together form a category called **CMon**.

► **Definition 26.** A **CMon**-enrichment of a category $\mathbf{C}$ is a choice of commutative monoid structure $(+, 0)$ for each hom-set $\mathbf{C}(\mathcal{X}, \mathcal{Y})$ that makes composition of morphisms bilinear as follows:

$$\begin{aligned} P \circ 0 &= 0 & P \circ (Q + R) &= P \circ Q + P \circ R \\ 0 \circ P &= 0 & (Q + R) \circ P &= Q \circ P + R \circ P. \end{aligned} \tag{20}$$

When $\mathbf{C}$ has a monoidal product $\otimes$, we say that a $\mathbf{CMon}$ -enrichment is monoidal if the monoidal product is also bilinear:

$$\begin{aligned} P \otimes 0 &= 0 & P \otimes (Q + R) &= (P \otimes Q) + (P \otimes R) \\ 0 \otimes P &= 0 & (Q + R) \otimes P &= (Q \otimes P) + (R \otimes P). \end{aligned} \quad (21)$$

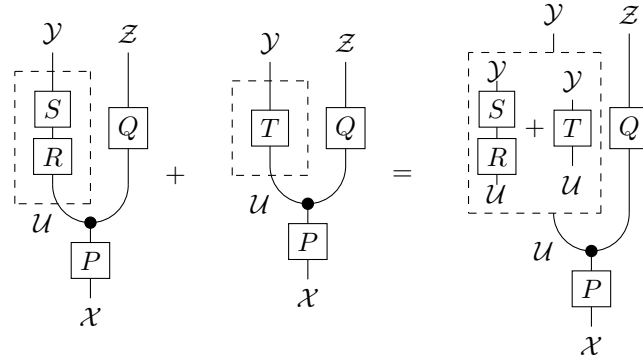
Here the morphisms in (20) and (21) range over all choices that make the expressions typecheck. Note also that we are using the same symbols $+$ and 0 to denote the addition operation and zero maps across different hom-sets.

We are especially interested in $\mathbf{CMon}$ -enrichments of CD categories. This leads to the following definition.

► **Definition 27.** A CD category is semiadditive³ if it is equipped with a monoidal $\mathbf{CMon}$ -enrichment.

We emphasise that for what follows, $\mathbf{CMon}$ -enrichment is the only additional structure we will require beyond that of a vanilla CD category. In particular, all the definitions and results below make sense without additional machinery beyond this.

► **Remark 28.** The combination of (20) and (21) has a useful interpretation in terms of string diagrams. The conditions involving zero maps mean that zero maps *annihilate* string diagrams: if any morphism of a string diagram is zero, then the entire diagram is also zero, regardless of the other morphisms it contains. The remaining additivity conditions mean that if we are given two string diagrams that are identical apart from one specific subdiagram, then their sum can be computed by replacing that subdiagram with the sum of those two subdiagrams. For example:



where here the dashed boxes indicate the subdiagrams that differ. This provides a convenient tool for reasoning about semiadditive CD categories that we will make use of below.

Categories of kernels (Example 18) admit natural enrichments over commutative monoids as follows.

► **Proposition 29.** Let $\mathcal{X}$ and $\mathcal{Y}$ be measurable spaces, and define the zero kernel $0 : \mathcal{X} \rightarrow \mathcal{Y}$ and the sum $P + Q : \mathcal{X} \rightarrow \mathcal{Y}$ of kernels $P, Q : \mathcal{X} \rightarrow \mathcal{Y}$ respectively as follows:

$$\begin{aligned} 0(x, dy) &:= 0 \\ (P + Q)(x, dy) &:= P(x, dy) + Q(x, dy). \end{aligned}$$

This constitutes a $\mathbf{CMon}$ -enrichment of $\mathbf{Kern}$, and a monoidal $\mathbf{CMon}$ -enrichment of $\mathbf{sfKern}$ (so in particular $\mathbf{sfKern}$ is semiadditive).

³ Note that our terminology does not imply the existence of *biproducts*, as is sometimes the case in the literature.

Proof. This follows straightforwardly by checking the commutative monoid axioms and the compatibility conditions (20) and (21). The key point is that kernels take values in $[0, \infty]$, which itself is a commutative monoid under addition. $\blacktriangleleft$

► **Remark 30.** Let $\mathbf{C}$ be a CD category. As observed by [12, Section 3.1], for any object $\mathcal{X}$, the effects $\mathbf{C}(\mathcal{X}, \mathcal{I})$ obtain an additional commutative monoid structure $(*, 1)$. Specifically, the operation $*$ is defined for effects $v, w : \mathcal{X} \rightarrow \mathcal{I}$ as

$$\begin{array}{c} \triangle v \\ | \\ \mathcal{X} \end{array} * \begin{array}{c} \triangle w \\ | \\ \mathcal{X} \end{array} := \begin{array}{c} \triangle v \quad \triangle w \\ \quad \bullet \\ | \\ \mathcal{X} \end{array}$$

and the unit is given by the deletion map, which will write suggestively in what follows as

$$1 := \text{del}_{\mathcal{X}} : \mathcal{X} \rightarrow \mathcal{I}. \quad (22)$$

(Note that, as for $+$ and 0 above, we do not denote the dependence of $*$ and 1 on $\mathcal{X}$ explicitly.)

In **sfKern**, the effect $v * w$ corresponds to the pointwise multiplication of the nonnegative functions associated to v and w (Example 20).

When $\mathbf{C}$ is semiadditive, effects $\mathbf{C}(\mathcal{X}, \mathcal{I})$ become a *commutative semiring* (i.e. a ring without additive inverses): their “addition” structure is obtained from the **CMon**-enrichment $(+, 0)$, and their “multiplication” structure is obtained from $(*, 1)$ just defined. The semiring axioms can be shown diagrammatically using Remark 28: for example, we obtain $0 * w = 0$ because zero maps annihilate general morphisms, and the distributive law $w * (v + u) = (w * v) + (w * u)$ holds because

$$\begin{array}{c} \triangle w \\ | \\ \bullet \\ | \\ \mathcal{X} \end{array} \begin{array}{c} \triangle v \\ | \\ \triangle u \end{array} = \begin{array}{c} \triangle w \quad \triangle v \\ \quad \bullet \\ | \\ \mathcal{X} \end{array} + \begin{array}{c} \triangle w \quad \triangle u \\ \quad \bullet \\ | \\ \mathcal{X} \end{array}$$

► **Remark 31.** Also as noted by [12, Section 3.1], in a CD category $\mathbf{C}$, the monoid of effects $\mathbf{C}(\mathcal{X}, \mathcal{I})$ acts on the hom-sets $\mathbf{C}(\mathcal{X}, \mathcal{Y})$. Specifically, given an effect $w : \mathcal{X} \rightarrow \mathcal{I}$ and a morphism $P : \mathcal{X} \rightarrow \mathcal{Y}$, we obtain a new morphism

$$\begin{array}{c} \mathcal{Y} \\ | \\ \boxed{w \cdot P} \\ | \\ \mathcal{X} \end{array} := \begin{array}{c} \mathcal{Y} \\ | \\ \triangle w \quad \boxed{P} \\ \quad \bullet \\ | \\ \mathcal{X} \end{array} \quad (23)$$

For example, the first MH term (16) may be expressed in terms of this action as $\alpha \cdot \phi$. In the case of **sfKern**, it is straightforward to check that the morphism $w \cdot K$ is kernel given by

$$(w \cdot P)(x, dy) = w(x) P(x, dy),$$

and so we think of $w \cdot P$ as being given by P *reweighted* according to w on the input side.

► **Remark 32.** It is a standard fact that a commutative monoid $\mathcal{M}$ admits a canonical *preorder* (i.e. a reflexive and transitive relation) $\leq$ defined by $a \leq b$ if there exists $c \in \mathcal{M}$ such that $a + c = b$. This extends to the morphisms of every **CMon**-enriched category $\mathbf{C}$ accordingly. It is also easy to check that this preorder is preserved by pre- and post-composition: if $P \leq Q$, then $R \circ P \circ S \leq R \circ Q \circ S$ for all compatibly typed morphisms R and S . From the perspective of enriched category theory, this means that every **CMon**-enriched category is also enriched over the category of preorders.

When $\mathbf{C}$ is monoidally **CMon**-enriched, this preorder is also compatible with the monoidal product: if $P \leq Q$, then some basic manipulations show that $R \otimes P \otimes S \leq R \otimes Q \otimes S$ holds for all morphisms R and S . It follows that the induced enrichment over preorders is also monoidal in this case.

► **Remark 33.** For a small number of results below, we will consider **CMon**-enrichments that satisfy certain additional compatibility conditions. For completeness, we give a list here (which can be skipped on a first reading). We will say that:

- $\mathbf{C}$ is *zero-sum-free* if $P + Q = 0$ implies $P = Q = 0$;
- $\mathbf{C}$ is *zero-monic* if $1 \circ P = 0$ implies $P = 0$ (where recall from (22) that $1 = \text{del}$);
- $\mathbf{C}$ has *no zero divisors* if $P \otimes Q = 0$ implies $P = 0$ or $Q = 0$.

Note that the last two conditions need additional structure (e.g. a CD category) to make sense. The zero-monic condition was previously introduced in the context of *effectus theory* [8, Definition 31].

It is straightforward to check that being zero-sum-free is equivalent to 0 being a bottom element with respect to the preorder from Remark 32. In other words, $P \leq 0$ implies $P = 0$. Some basic manipulations show that **sfKern** satisfies all three of these requirements, and **Kern** satisfies the first two.

4.2 Substochasticity, probabilities, and convex combinations

Semiadditive CD categories admit the following abstract definition of *substochastic kernels* and *probabilities*. Moreover, using the formalism from Remarks 30 and 32, this mirrors notationally the usual definitions from classical probability theory.

► **Definition 34.** Let $P : \mathcal{X} \rightarrow \mathcal{Y}$ be a morphism in a semiadditive CD category $\mathbf{C}$. Then P is *substochastic* if it holds that

$$1 \circ P \leq 1, \tag{24}$$

where recall that $1 = \text{dely}$ denotes the multiplicative unit in the commutative monoid of effects (Remark 30). A *probability* is defined as a substochastic effect.

► **Remark 35.** Since $\text{id}_{\mathcal{I}} = \text{del}_{\mathcal{I}} = 1$ in a CD category (Definition 71), a probability is equivalently an effect $\alpha : \mathcal{X} \rightarrow \mathcal{I}$ with $\alpha \leq 1$. By Remark 32, this in turn means equivalently that there exists some effect $\alpha^c : \mathcal{X} \rightarrow \mathcal{I}$ such that

$$\alpha + \alpha^c = 1.$$

Note that since $+$ is commutative, α^c is also a probability.

► **Example 36.** Recall from Example 20 that effects in **sfKern** correspond to nonnegative measurable functions. Given $P : \mathcal{X} \rightarrow \mathcal{Y}$ in **sfKern**, it is straightforward to check that the effect $1 \circ P$ corresponds to the measurable function $x \mapsto P(x, \mathcal{Y})$, and that $1 \circ P \leq 1$ if and only if $P(x, \mathcal{Y}) \leq 1$ holds for all $x \in \mathcal{X}$. In this way we recover the usual notion of a *substochastic kernel*. Likewise, a probability $\alpha : \mathcal{X} \rightarrow \mathcal{I}$ in **sfKern** corresponds to a measurable function $\mathcal{X} \rightarrow [0, 1]$, as the name suggests.

► **Remark 37.** Using probabilities, we can express *convex combinations* of morphisms in a semiadditive CD category. Specifically, given a probability $\alpha : \mathcal{X} \rightarrow \mathcal{I}$, we may regard the morphism $\alpha \cdot P + \alpha^c \cdot Q$ as an abstract convex combination of $P, Q : \mathcal{X} \rightarrow \mathcal{Y}$, where the action $\cdot$ is defined in (23). In string diagrams, this may be written as

$$\begin{array}{c} \mathcal{Y} \\ \triangle \alpha \quad \square P \\ \curvearrowright \\ \bullet \\ \downarrow \\ \mathcal{X} \end{array} + \begin{array}{c} \mathcal{Y} \\ \triangle \alpha^c \quad \square Q \\ \curvearrowright \\ \bullet \\ \downarrow \\ \mathcal{X} \end{array} \quad (25)$$

Intuitively, here P is weighted by α , while Q is weighted by α^c . In **sfKern**, by Remark 31, this corresponds to the kernel

$$\alpha(x) P(x, dy) + (1 - \alpha(x)) Q(x, dy), \quad (26)$$

which recovers the usual notion of convex combinations of kernels.

Before proceeding we note the following structural result that will be useful for us below.

► **Proposition 38.** *Let $\mathbf{C}$ be a semiadditive CD category. Then the objects and substochastic morphisms of $\mathbf{C}$ form a CD category, with its CD category structure inherited from $\mathbf{C}$.*

Proof. We first show substochasticity is preserved by composition. Let $P : \mathcal{X} \rightarrow \mathcal{Y}$ and $Q : \mathcal{Y} \rightarrow \mathcal{Z}$ be substochastic. Hence $1 \circ Q \leq 1$, and so precomposing both sides by P gives

$$1 \circ Q \circ P \leq 1 \circ P \leq 1$$

by Remark 32, where we use substochasticity of P in the second step. A similar argument (also using Remark 32) shows that substochastic morphisms are closed under monoidal products also.

Next, observe that every normalised (Remark 19) morphism $R : \mathcal{X} \rightarrow \mathcal{Y}$ (see [7, Definition 2.3]) in $\mathbf{C}$ by definition satisfies $1 \circ R = 1$ (since recall we defined $1 = \text{del}$ in Remark 30). Since $\leq$ is reflexive, every normalised morphism R is therefore substochastic. It is also standard to show that the structure morphisms of a CD category are all normalised, i.e., identities, copy and delete maps, and the symmetric monoidal structure maps [7, Section 7]. In this way, the required CD category structure is inherited from $\mathbf{C}$. ◀

4.2.1 Aside: connection with distributive Markov categories

Statisticians are trained automatically to associate a convex combination like (26) with a sampling procedure as follows:

$$B \sim \text{flip}(\alpha(x)) \quad Y \sim P(x, dy) \quad Y' \sim Q(x, dy) \quad \text{return } Y \text{ if } B = 1 \text{ else } Y' \quad (27)$$

Here $\text{flip}(a)$ denotes a Bernoulli distribution (i.e. a coin toss) with success probability a . However, while this makes sense in **sfKern**, it is not clear how to interpret (27) semantically inside a general semiadditive CD category, since in general these do not supply morphisms corresponding to flip or the “if-else” term here.

On the other hand, [1, Definition 2] recently introduced *distributive Markov categories*, which are Markov categories equipped with coproducts that distribute over the monoidal product (see Appendix D.2), and which do provide such morphisms in a canonical way.

However, since effects in Markov categories are always trivial, distributive Markov categories in general do not supply morphisms corresponding to individual probability *values* such as $\alpha : \mathcal{X} \rightarrow \mathcal{I}$. In turn, this means they cannot be used to interpret convex combinations like (26) directly.

For our goal of proving Theorem 1, it is sufficient to work entirely in semiadditive CD categories, as we do below. However, for general probability theory, it seems useful to have a setting that supports both these perspectives. This is provided by the following result, which we prove in Appendix D.2.

► **Proposition 39.** *Let $\mathbf{C}$ be a semiadditive CD category with an initial object and binary coproducts whose inclusions are normalised and deterministic. Then these coproducts equip $\mathbf{C}_{\text{norm}}$ (Remark 19) with the structure of a distributive Markov category.*

4.3 Cancellative and finite morphisms

In this subsection we build up to the definition of a *finitely cancellative* CD category (Definition 46), which we require for the converse part of Theorem 52 below. To begin, it is often useful to know that a commutative monoid $\mathcal{M}$ is *cancellative*. This property says that if $P + Q = P + R$ holds in $\mathcal{M}$, then also $Q = R$. Unfortunately, due to the possibility of obtaining infinite mass, kernels are not cancellative in general. For example, if $P(x, A) = \infty$ for all x and A , then $P + Q = P$ for any Q . This motivates the following elementwise definition instead:

► **Definition 40.** *Let $P : \mathcal{X} \rightarrow \mathcal{Y}$ be a morphism in a category $\mathbf{C}$ equipped with a $\mathbf{CMon}$ -enrichment. We say that P is cancellative if for all $Q, R : \mathcal{X} \rightarrow \mathcal{Y}$ such that $P + Q = P + R$, it holds that $Q = R$.*

Note that we define this purely in terms of a $\mathbf{CMon}$ -enrichment (not a CD structure). It therefore applies to the category $\mathbf{Kern}$ (Example 18) as well as to $\mathbf{sfKern}$.

The following basic result holds for cancellative morphisms in any $\mathbf{CMon}$ -enriched category.

► **Proposition 41.** *Let $P, Q : \mathcal{X} \rightarrow \mathcal{Y}$ be morphisms in a $\mathbf{CMon}$ -enriched category $\mathbf{C}$. If $P \leq Q$ and Q is cancellative, then so is P .*

Proof. By assumption, there is some $R : \mathcal{X} \rightarrow \mathcal{Y}$ such that $P + R = Q$. If we have $P + S = P + T$ for some $S, T : \mathcal{X} \rightarrow \mathcal{Y}$, then

$$Q + S = (P + R) + S = (P + R) + T = Q + T,$$

and so $S = T$ by cancellativity of Q . ◀

For kernels, cancellativity recovers the notion of σ -finiteness. Recall that a measure μ on a measurable space $\mathcal{X}$ is σ -finite if there exists disjoint measurable $A_1, A_2, \dots \subseteq \mathcal{X}$ such that $\mathcal{X} = \bigcup_{n=1}^{\infty} A_n$ and each $\mu(A_n) < \infty$. The following result is proven in Appendix D.3.

► **Proposition 42.** *Let $K : \mathcal{X} \rightarrow \mathcal{Y}$ be a kernel (i.e. a morphism in $\mathbf{Kern}$). If K is pointwise σ -finite (i.e. each $K(x, dy)$ is σ -finite), then K is cancellative. The converse holds if all singleton subsets of $\mathcal{X}$ are measurable.*

The following basic result would be tedious to show in terms of σ -finiteness, but is straightforward in terms of cancellativity.

► **Proposition 43.** *Cancellative morphisms in a $\mathbf{CMon}$ -enriched category are preserved under composition with isomorphisms.*

Proof. Let $P : \mathcal{X} \rightarrow \mathcal{Y}$ be cancellative, and let $\varphi : \mathcal{W} \rightarrow \mathcal{X}$ and $\psi : \mathcal{Y} \rightarrow \mathcal{Z}$ be isomorphisms. Suppose $Q, R : \mathcal{W} \rightarrow \mathcal{Z}$ satisfy

$$\psi \circ P \circ \varphi + Q = \psi \circ P \circ \varphi + R.$$

Bilinearity of composition (Definition 26) implies that

$$P + \psi^{-1} \circ Q \circ \varphi^{-1} = P + \psi^{-1} \circ R \circ \varphi^{-1},$$

and hence $\psi^{-1} \circ Q \circ \varphi^{-1} = \psi^{-1} \circ R \circ \varphi^{-1}$ by cancellativity of P . Applying ψ and φ on either side then gives $Q = R$. ◀

A closely related notion to cancellativity is *finiteness*. Our terminology here is motivated by the case of kernels (see Proposition 45 below, proven in Appendix D.3.2).

► **Definition 44.** *A morphism $P : \mathcal{X} \rightarrow \mathcal{Y}$ in a semiadditive CD category $\mathbf{C}$ is finite if the effect $1 \circ P$ is cancellative.*

► **Proposition 45.** *Let $P : \mathcal{X} \rightarrow \mathcal{Y}$ be a kernel (i.e. a morphism in $\mathbf{Kern}$). Then P is finite if and only if $P(x, \mathcal{Y}) < \infty$ for all $x \in \mathcal{X}$.*

In $\mathbf{Kern}$ (and hence $\mathbf{sfKern}$), it is clear that finiteness implies cancellativity (i.e. σ -finiteness) by Proposition 42. It is therefore natural to ask whether this relationship holds generally. We conjecture that it does not always hold, and so introduce the following definition to capture the cases where it does.

► **Definition 46.** *A semiadditive CD category $\mathbf{C}$ is finitely cancellative if all finite morphisms in $\mathbf{C}$ are cancellative.*

Finiteness interacts with substochasticity in a natural way, as the following result shows.

► **Proposition 47.** *Let $P : \mathcal{X} \rightarrow \mathcal{Y}$ and $Q : \mathcal{Y} \rightarrow \mathcal{Z}$ be morphisms in a semiadditive CD category $\mathbf{C}$. If P is finite and Q is substochastic, then $Q \circ P$ is finite.*

Proof. By substochasticity of Q and Remark 32, we have $1 \circ Q \circ P \leq 1 \circ P$. Since P is finite, the right-hand side is cancellative. By Proposition 41, the left-hand side is also cancellative, so $Q \circ P$ is finite. ◀

► **Remark 48.** Let $\mu : \mathcal{I} \rightarrow \mathcal{X}$ and $P, Q : \mathcal{X} \rightarrow \mathcal{Y}$ be morphisms in a CD category $\mathbf{C}$. Following [7, Definition 5.1], we will say that $P = Q$ μ -almost everywhere if

$$\begin{array}{c} \mathcal{X} \quad \mathcal{Y} \\ \downarrow \quad \downarrow \\ \text{---} \boxed{P} \text{---} \\ \uparrow \\ \bullet \\ \downarrow \triangle \mu \end{array} = \begin{array}{c} \mathcal{X} \quad \mathcal{Y} \\ \downarrow \quad \downarrow \\ \text{---} \boxed{Q} \text{---} \\ \uparrow \\ \bullet \\ \downarrow \triangle \mu \end{array} \quad (28)$$

Notice that this equation makes sense in any CD category. However, [7] consider this specifically in the context of Markov categories (or *affine CD categories* in their terminology). One of their reasons for this is that, in $\mathbf{sfKern}$, the condition (28) does not necessarily imply (for example) that

$$P(x, A) = Q(x, A) \quad \text{for } \mu\text{-almost all } x \in \mathcal{X}, \quad (29)$$

since pathologies can arise if μ is infinite. However, (29) *does* follow from (28) if μ is σ -finite. Accordingly, we will allow ourselves to consider the equation in full generality, knowing we can recover the expected meaning in $\mathbf{sfKern}$ by restricting to a cancellative μ .

4.4 The full MH kernel

Semiadditive CD categories allow us to reason about the full MH kernel Π_{MH} from (1), rather than just its first term. For this, simply observe that Π_{MH} is a convex combination of two kernels: one that applies the involution ϕ , and the identity kernel that leaves the input unchanged. Accordingly, suppose $\phi : \mathcal{E} \rightarrow \mathcal{E}$ is a deterministic involution and $\alpha : \mathcal{E} \rightarrow \mathcal{I}$ is a probability in a semiadditive CD category $\mathbf{C}$. The following morphism in $\mathbf{C}$ then constitutes an abstract version of Π_{MH} :

$$P_{\text{MH}} := \alpha \cdot \phi + \alpha^c \cdot \text{id}_{\mathcal{E}}.$$

In string diagrams, this is as follows:

$$(30)$$

► **Example 49.** Take $\mathbf{C} := \mathbf{sfKern}$, and suppose ϕ is obtained by lifting an involutive measurable function to a deterministic kernel $\delta_{\phi(\xi)}(d\xi')$ as in (6), where δ is the Dirac delta. Likewise, the identity $\text{id}_{\mathcal{E}}$ in $\mathbf{sfKern}$ becomes the kernel $\delta_{\xi}(d\xi')$. By Remark 37, we may therefore write

$$P_{\text{MH}}(\xi, d\xi') = \alpha(\xi) \delta_{\phi(\xi)}(d\xi') + (1 - \alpha(\xi)) \delta_{\xi}(d\xi'),$$

which recovers the original MH kernel (1) exactly.

We want to give conditions for our abstract MH kernel P_{MH} to be reversible. For this we require the following two results, which relate reversibility of a sum to reversibility of its summands.

► **Proposition 50.** *Let $\mu : \mathcal{I} \rightarrow \mathcal{X}$ and $P, Q : \mathcal{X} \rightarrow \mathcal{X}$ be morphisms in a semiadditive CD category $\mathbf{C}$. If P and Q are μ -reversible, then so is $P + Q$.*

Proof. If P and Q are μ -reversible, then

$$(31)$$

The first and third steps use Remark 28, and the second uses reversibility of P and Q . ◀

► **Proposition 51.** *Let $\mu : \mathcal{I} \rightarrow \mathcal{X}$ and $P, Q : \mathcal{X} \rightarrow \mathcal{X}$ be morphisms in a finitely cancellative semiadditive CD category $\mathbf{C}$, where μ is finite and Q is substochastic. If $P + Q$ and Q are both μ -reversible, then so is P .*

Proof. Suppose $P + Q$ and Q are μ -reversible. Then similar to (31), we have

$$\begin{aligned}
 & \begin{array}{c} \mathcal{X} \quad \mathcal{X} \\ | \quad | \\ \boxed{P} \\ | \quad | \\ \bullet \\ \diagup \quad \diagdown \\ \mu \end{array} + \begin{array}{c} \mathcal{X} \quad \mathcal{X} \\ | \quad | \\ \boxed{Q} \\ | \quad | \\ \bullet \\ \diagup \quad \diagdown \\ \mu \end{array} = \begin{array}{c} \mathcal{X} \quad \mathcal{X} \\ | \quad | \\ \boxed{P+Q} \\ | \quad | \\ \bullet \\ \diagup \quad \diagdown \\ \mu \end{array} = \begin{array}{c} \mathcal{X} \quad \mathcal{X} \\ | \quad | \\ \boxed{P+Q} \\ | \quad | \\ \bullet \\ \diagdown \quad \diagup \\ \mu \end{array} \\
 & = \begin{array}{c} \mathcal{X} \quad \mathcal{X} \\ | \quad | \\ \boxed{P} \\ | \quad | \\ \bullet \\ \diagdown \quad \diagup \\ \mu \end{array} + \begin{array}{c} \mathcal{X} \quad \mathcal{X} \\ | \quad | \\ \boxed{Q} \\ | \quad | \\ \bullet \\ \diagdown \quad \diagup \\ \mu \end{array} = \begin{array}{c} \mathcal{X} \quad \mathcal{X} \\ | \quad | \\ \boxed{P} \\ | \quad | \\ \bullet \\ \diagup \quad \diagdown \\ \mu \end{array} + \begin{array}{c} \mathcal{X} \quad \mathcal{X} \\ | \quad | \\ \boxed{Q} \\ | \quad | \\ \bullet \\ \diagup \quad \diagdown \\ \mu \end{array}
 \end{aligned}$$

where we use reversibility of $P + Q$ in the first step, and reversibility of Q in the second. But now the summand involving Q is a composition of a finite morphism (namely μ) and a substochastic morphism (using Proposition 38), and is therefore finite by Proposition 47. Since $\mathbf{C}$ is finitely cancellative, we may cancel this summand on both sides to obtain reversibility of P . $\blacktriangleleft$

We can now give necessary and sufficient conditions for reversibility of the full MH kernel P_{MH} from (30) in a finitely cancellative semiadditive CD category.

► **Theorem 52.** *Let $\mathbf{C}$ be a finitely cancellative semiadditive CD category, and suppose*

- $\mu : \mathcal{I} \rightarrow \mathcal{E}$ is finite
- $\phi : \mathcal{E} \rightarrow \mathcal{E}$ is a deterministic involution
- $\alpha : \mathcal{E} \rightarrow \mathcal{I}$ is a probability
- $r = \frac{d(\phi \circ \mu)}{d\mu}$ is a Radon–Nikodym derivative.

Then P_{MH} defined in (30) is μ -reversible if and only if the following “balancing condition” holds:

$$\alpha = (\alpha \circ \phi) * r \quad \mu\text{-almost everywhere.} \quad (32)$$

Proof. By the definition of $*$ (Remark 30) and μ -almost everywhere (Remark 48), the balancing condition (32) becomes precisely (18). From Theorem 25, it follows that the first summand $\alpha \cdot \phi$ of P_{MH} is μ -reversible if and only if the balancing condition holds. The second summand $\alpha^c \cdot \text{id}_{\mathcal{E}}$ is always μ -reversible, as can be seen from inspection of its string diagram in (30). By Proposition 38, it is also substochastic, since α^c is a probability (Remark 35). The result now follows by Propositions 50 and 51. $\blacktriangleleft$

Instantiating this result in **sfKern** now yields the following:

► **Corollary 53.** *Suppose μ is a finite measure on a measurable space $\mathcal{E}$, and $\phi : \mathcal{E} \rightarrow \mathcal{E}$ is a measurable involution such the pushforward measure μ^ϕ is absolutely continuous with respect to μ . Then the MH kernel Π_{MH} defined in (1) is μ -reversible if and only if*

$$\alpha(\xi) = \alpha(\phi(\xi)) r(\xi) \quad \text{for } \mu\text{-almost all } \xi \in \mathcal{E}, \quad (33)$$

where $r := \frac{d\mu^\phi}{d\mu}$ denotes the Radon–Nikodym derivative.

Proof. We apply Theorem 52 with $\mathbf{C} := \mathbf{sfKern}$. In this setting μ becomes a finite measure on a measurable space $\mathcal{E}$ (Proposition 45), and μ -reversibility corresponds to the usual measure-theoretic notion of reversibility (Example 7). Any choice of involutive measurable function $\phi : \mathcal{E} \rightarrow \mathcal{E}$ lifts to become a deterministic involutive kernel, and α and r correspond to measurable functions $\mathcal{E} \rightarrow [0, 1]$ and $\mathcal{E} \rightarrow [0, \infty]$ respectively (Examples 36 and 20). The condition $r = \frac{d(\phi \circ \mu)}{d\mu}$ says that r is a Radon–Nikodym derivative in the usual measure-theoretic sense (Example 22). The morphism P_{MH} becomes Π_{MH} (Example 49), while the balancing condition becomes (33) (Remarks 30 and 48), which gives the result. $\blacktriangleleft$

Corollary 53 fully recovers Theorem 1 in the case that μ and μ^ϕ are both absolutely continuous with respect to each other (which ensures the Radon–Nikodym derivative r exists). It also supplies a converse statement, which was not given in [2]. In Section 4.5 below, we show how we may also reason about absolute continuity within semiadditive CD categories, thereby recovering the rest of Theorem 1 as well.

4.4.1 Extension to skew-reversibility

We can combine Theorem 52 with our earlier Proposition 13 (see also Remark 23) to obtain necessary and sufficient conditions for the more complex property of skew-reversibility (Section 2.3). Consider the morphism $s \circ P_{\text{MH}}$ obtained by postcomposing our MH kernel with an involution $s : \mathcal{E} \rightarrow \mathcal{E}$. From bilinearity of composition, we have

$$s \circ P_{\text{MH}} = s \circ (\alpha \cdot \phi + \alpha^c \cdot \text{id}_{\mathcal{E}}) = \alpha \cdot (s \circ \phi) + \alpha^c \cdot (s \circ \text{id}_{\mathcal{E}}),$$

where the second step is easy to check using string diagrams. In the context of $\mathbf{sfKern}$, this recovers the following MH-like kernel considered by [2, Section 4]:

$$\Pi_{\text{sMH}}(\xi, d\xi') := \alpha(\xi) \delta_{s(\phi(\xi))}(d\xi') + (1 - \alpha(\xi)) \delta_{s(\xi)}(d\xi'). \quad (34)$$

When s is μ -invariant, Proposition 13 implies that P_{MH} is (μ, s) -reversible if and only if P_{MH} is μ -reversible. Together with Theorem 52, this gives directly the following:

► **Corollary 54.** *Given the setup of Corollary 53, suppose we also have a measurable involution $s : \mathcal{E} \rightarrow \mathcal{E}$ that is μ -invariant. Then Π_{sMH} is (μ, s) -reversible if and only if the balancing condition (33) holds.*

Combined with our results on absolute continuity in Section 4.5 below, this gives a more streamlined and more general version of [2, Proposition 3(c)], which gives sufficient conditions for Π_{sMH} to be reversible but does not supply a converse statement.

4.5 Absolute continuity

From Remark 32, the hom-sets of a $\mathbf{CMon}$ -enriched category $\mathbf{C}$ admit a preorder $\leq$. They also admit a separate preorder related to *absolute continuity*. Recall that a measure ν is absolutely continuous with respect to another measure μ if $\mu(A) = 0$ implies $\nu(A) = 0$. Intuitively, this says that ν assigns mass only to regions where μ does. When both are absolutely continuous with respect to each other, they are said to be *equivalent*. The following extends these ideas to arbitrary morphisms in a $\mathbf{CMon}$ -enriched category.

► **Definition 55.** *Let $P, Q : \mathcal{X} \rightarrow \mathcal{Y}$ be morphisms in a $\mathbf{CMon}$ -enriched category $\mathbf{C}$. We write $P \ll Q$ if for all compatible morphisms R and S in $\mathbf{C}$ we have*

$$S \circ Q \circ R = 0 \implies S \circ P \circ R = 0. \quad (35)$$

We write $P \equiv Q$ if both $P \ll Q$ and $Q \ll P$ hold.

Note that this does not require a CD structure: only a **CMon**-enrichment. In particular, we can ask what happens when $\mathbf{C} = \mathbf{Kern}$ (Example 18). This is answered by the following result, proven in Section D.4.1 of the Appendix.

► **Proposition 56.** *Let $P, Q : \mathcal{X} \rightarrow \mathcal{Y}$ be kernels (i.e. morphisms in $\mathbf{Kern}$). Then $P \ll Q$ if and only if $P(x, dy)$ is absolutely continuous with respect to $Q(x, dy)$ for all $x \in \mathcal{X}$. Accordingly, $P \equiv Q$ if and only if $P(x, dy)$ and $Q(x, dy)$ are equivalent as measures for all $x \in \mathcal{X}$.*

It is straightforward to check that $\ll$ is reflexive and transitive and is therefore a preorder. As for $\leq$ in Remark 32, it also induces an enrichment of $\mathbf{C}$ over preorders, as the following result shows. We omit the proof, which is immediate from Definition 55.

► **Proposition 57.** *$\ll$ is preserved by pre- and post-composition.*

The next two results establish basic facts about the preorder $\ll$. Proposition 58 shows that it has 0 as the unique bottom element, and Proposition 59 shows it is a less informative preorder than $\leq$ in many cases of interest.

► **Proposition 58.** *It holds that $0 \ll P$ for every morphism P . Also, if $P \ll 0$, then $P = 0$.*

Proof. We have directly from (20) that $S \circ 0 \circ R = 0$ for all compatible morphisms R and S . This immediately gives the first claim. If $P \ll 0$, then also $S \circ P \circ R = 0$ for all compatible R and S . Taking R and S to be identities shows $P = 0$, which gives the second claim. ◀

► **Proposition 59.** *Let $\mathbf{C}$ be a **CMon**-enriched category. If $\mathbf{C}$ is zero-sum-free (Remark 33), then $P \leq Q$ implies $P \ll Q$.*

Proof. Suppose $P \leq Q$, and $S \circ Q \circ R = 0$. Then $S \circ P \circ R \leq S \circ Q \circ R$ by Remark 32, and so $S \circ P \circ R = 0$ since $\mathbf{C}$ is zero-sum-free. ◀

A distinct notion of absolute continuity was previously introduced by [13, Definition 2.2.1] for Markov categories. Their definition applies to morphisms with potentially different domains: given $P : \mathcal{X} \rightarrow \mathcal{Y}$ and $Q : \mathcal{Z} \rightarrow \mathcal{Y}$, they say that P is *absolutely continuous with respect to Q* if Q -almost sure equality implies P -almost sure equality (where their notion of *almost sure equality* is an extension of (28); see [13, Definition 2.1.1]). In **Stoch**, this becomes:

$$Q(z, A) = 0 \text{ for all } z \in \mathcal{Z} \implies P(x, A) = 0 \text{ for all } x \in \mathcal{X}.$$

Note that this is not a “pointwise” condition, even when $\mathcal{X} = \mathcal{Z}$, and as such is generally distinct from the concept we consider (cf. Proposition 56 above). However, when $\mathcal{X} = \mathcal{Z} = \mathcal{I}$, the following result shows that their notion implies ours under mild assumptions on $\mathbf{C}$. See Section D.4.2 of the Appendix for the proof.

► **Proposition 60.** *Let $\mathbf{C}$ be a semiadditive CD category that is zero-monic and has no zero divisors (Remark 33). If $\mu, \nu : \mathcal{I} \rightarrow \mathcal{X}$ are morphisms in $\mathbf{C}$ such that ν is absolutely continuous with respect to μ in the sense of [13, Definition 2.2.1], then $\nu \ll \mu$ in our sense.*

4.6 Lebesgue decompositions

A fundamental concept in measure theory is the notion of a *Lebesgue decomposition*. Given measures ν and μ on the same measurable space, recall that a Lebesgue decomposition of ν with respect to μ is an expression of the form $\nu = \nu_{ac} + \nu_s$, where ν_{ac} is a measure that

is absolutely continuous with respect to μ , and ν_s is another measure that is *singular* with respect to μ . Here singular means there exists a measurable set S such that $\nu_s(S) = 0$ and $\mu(S^c) = 0$. Intuitively, ν_s and μ assign mass to disjoint regions of the space. It is a classical result that a Lebesgue decomposition always exists whenever ν and μ are σ -finite [20, Section 32]. We now establish an abstract version of this concept in the setting of **CMon**-enriched categories.

► **Definition 61.** Let $\mathbf{C}$ be a **CMon**-enriched category. A meet of morphisms $P, Q : \mathcal{X} \rightarrow \mathcal{Y}$ in $\mathbf{C}$ is a greatest lower bound of P and Q with respect to the preorder $\ll$.

Meets do not need always need to exist. Moreover, when they do, they are usually not unique, since $\ll$ is only a preorder.

► **Proposition 62.** Meets are preserved by pre- and post-composition with isomorphisms.

Proof. Suppose R is a meet of P and Q . Given compatible isomorphisms φ and ψ , Proposition 57 immediately gives that $\psi \circ R \circ \varphi$ is a lower bound of $\psi \circ P \circ \varphi$ and $\psi \circ Q \circ \varphi$ with respect to $\ll$. Now suppose that S is some other lower bound. Then $\psi^{-1} \circ S \circ \varphi^{-1}$ is a lower bound of P and Q , and so $\psi^{-1} \circ S \circ \varphi^{-1} \ll R$. Proposition 57 then gives $S \ll \psi \circ R \circ \varphi$ as required. ◀

Meets allow us to capture the notion of *singular measures* abstractly via the following definition. Proposition 64 then shows that this recovers the usual measure-theoretic notion of singularity in the case of kernels (see Section D.4.3 of the Appendix for the proof).

► **Definition 63.** Given morphisms $P, Q : \mathcal{X} \rightarrow \mathcal{Y}$ in a **CMon**-enriched category $\mathbf{C}$, we write $P \perp Q$ if 0 is a meet of P and Q .

► **Proposition 64.** Suppose $P, Q : \mathcal{X} \rightarrow \mathcal{Y}$ are kernels (i.e. morphisms in **Kern**). If $P(x, dy)$ and $Q(x, dy)$ are singular for all $x \in \mathcal{X}$, then $P \perp Q$. The converse also holds when all singleton subsets of $\mathcal{X}$ are measurable, and P and Q are pointwise σ -finite.

The following result shows that $\perp$ and $\ll$ interact as measure-theoretic intuitions (in terms of “assigning mass”) would suggest.

► **Proposition 65.** If $P \ll Q$ and $Q \perp R$, then $P \perp R$.

Proof. If we have S with $S \ll P$ and $S \ll R$, then $S \ll Q$ also, and so $S = 0$ since 0 is a meet of Q and R . ◀

The preceding definitions now allow us to give an abstract version of the notion of a Lebesgue decomposition in a **CMon**-enriched category as follows.

► **Definition 66.** Let $P, Q : \mathcal{X} \rightarrow \mathcal{Y}$ be morphisms in a **CMon**-enriched category $\mathbf{C}$. A Q -decomposition of P is a pair of morphisms $P_{ac}, P_s : \mathcal{X} \rightarrow \mathcal{Y}$ such that

- $P = P_{ac} + P_s$,
- P_{ac} is a meet of P and Q ,
- $P_s \perp Q$.

The following result shows that this definition recovers the classical notion of a Lebesgue decomposition in the case of kernels. See Section D.4.4 of the Appendix for the proof.

► **Proposition 67.** Suppose $P, Q, P_{ac}, P_s : \mathcal{X} \rightarrow \mathcal{Y}$ are kernels (i.e. morphisms in **Kern**). If $P(x, dy) = P_{ac}(x, dy) + P_s(x, dy)$ is a Lebesgue decomposition with respect to $Q(x, dy)$ for all $x \in \mathcal{X}$, then (P_{ac}, P_s) is a Q -decomposition of P . The converse holds when all singleton subsets of $\mathcal{X}$ are measurable, and P_{ac} and P_s are pointwise σ -finite.

4.6.1 Completing Theorem 1

We now wish to complete the proof of Theorem 1 using the tools developed above. It remains to show the first part of this result, which involves a decomposition of the target measure μ according to an involution ϕ . Lebesgue decompositions capture this abstractly via the following result.

► **Theorem 68.** *Let $\mathbf{C}$ be a zero-sum-free $\mathbf{CMon}$ -enriched category. Let $P : \mathcal{X} \rightarrow \mathcal{Y}$ and $\phi : \mathcal{Y} \rightarrow \mathcal{Y}$ be morphisms in $\mathbf{C}$, where ϕ is an involution, and suppose $(P_{\text{ac}}, P_{\text{s}})$ is a $(\phi \circ P)$ -decomposition of P . Then we have both of the following:*

- (1) $P_{\text{ac}} \equiv \phi \circ P_{\text{ac}}$
- (2) $P_{\text{s}} \perp \phi \circ P_{\text{s}}$

Proof. (1) By Proposition 62, $\phi \circ P_{\text{ac}}$ is a meet of $\phi \circ P$ and $P = \phi \circ \phi \circ P$. Since $(P_{\text{ac}}, P_{\text{s}})$ is a $(\phi \circ P)$ -decomposition, we also have $P_{\text{ac}} \ll \phi \circ P$ (by definition) and $P_{\text{ac}} \ll P$ (by Proposition 59, since $P_{\text{ac}} \leq P$). The meet property then implies $P_{\text{ac}} \ll \phi \circ P_{\text{ac}}$, which establishes one direction of (1). The other direction follows from Proposition 57, which gives

$$\phi \circ P_{\text{ac}} \ll \phi \circ \phi \circ P_{\text{ac}} = P_{\text{ac}}$$

and hence $\phi \circ P_{\text{ac}} \equiv P_{\text{ac}}$ as required.

(2) Suppose $R \ll P_{\text{s}}$ and $R \ll \phi \circ P_{\text{s}}$. Since $P_{\text{s}} \leq P$, we have $P_{\text{s}} \ll P$ and hence $\phi \circ P_{\text{s}} \ll \phi \circ P$ by Propositions 59 and 57 respectively. It follows that $R \ll \phi \circ P$. But now since 0 is a meet of P_{s} and $\phi \circ P$, we therefore have $R \ll 0$ as required. ◀

Instantiating this result in **sfKern** now gives the remaining part of Theorem 1:

► **Corollary 69.** *Suppose μ is a σ -finite measure on $\mathcal{E}$, and $\phi : \mathcal{E} \rightarrow \mathcal{E}$ is a measurable involution. Then there exists a measurable $S \subseteq \mathcal{E}$ such that the restriction $\mu|_S$ and its pushforward $\mu|_S^\phi$ are equivalent, and $\mu|_{S^c}$ and its pushforward $\mu|_{S^c}^\phi$ are singular.*

Proof. Since μ and μ^ϕ are both σ -finite measures, there exists a μ^ϕ -decomposition $\mu = \mu_{\text{ac}} + \mu_{\text{s}}$ [20, Section 32, Theorem C]. Regarding these as morphisms in **sfKern**, Theorem 68 gives $\mu_{\text{ac}} \equiv \mu_{\text{ac}}^\phi$ and $\mu_{\text{s}} \perp \mu_{\text{s}}^\phi$, and Proposition 65 gives $\mu_{\text{ac}} \perp \mu_{\text{s}}$. This shows μ_{ac} and μ_{ac}^ϕ are equivalent (Proposition 56) and μ_{s} and μ_{s}^ϕ are singular (Proposition 64). Similarly, μ_{ac} and μ_{s} are singular, which by definition just says there is a measurable $S \subseteq \mathcal{E}$ with $\mu_{\text{ac}}(S^c) = 0$ and $\mu_{\text{s}}(S) = 0$, and hence $\mu_{\text{ac}} = \mu|_S$ and $\mu_{\text{s}} = \mu|_{S^c}$ as required. ◀

5 Conclusion and future work

In this work we developed an enrichment of CD categories over commutative monoids, providing a minimal setting for reasoning about sums of morphisms, and used it to study the involutive MH framework of [2]. This gave a purely algebraic derivation of their “balancing condition” in Theorem 1, together with a converse result and an immediate extension to the skew-reversible setting.

An interesting direction for future work is to give general conditions under which Lebesgue decompositions exist in this setting, rather than only isolating the properties that such decompositions ought to satisfy. This appears to require a suitable version of the Radon–Nikodym theorem, which we conjecture may be attainable by assuming that the preorder $\leq$ induced by the enrichment is complete. We also expect that making this precise will involve an explicit notion of *predicates*, suggesting an interesting connection with *effectus theory* [8], which we leave to future work.

References

- 1 Nate Ackerman, Cameron E. Freer, Younesse Kaddar, Jacek Karwowski, Sean Moss, Daniel Roy, Sam Staton, and Hongseok Yang. Probabilistic programming interfaces for random graphs: Markov categories, graphons, and nominal sets. *Proceedings of the ACM on Programming Languages*, 8(POPL):1819–1849, January 2024. doi:10.1145/3632903.
- 2 Christophe Andrieu, Anthony Lee, and Sam Livingstone. A general perspective on the Metropolis-Hastings kernel. *arXiv preprint*, 2020. arXiv:2012.14881.
- 3 Christophe Andrieu and Samuel Livingstone. Peskun-Tierney ordering for Markovian Monte Carlo: Beyond the reversible scenario. *Ann. Statist.*, 49(4):1958–1981, 2021. doi:10.1214/20-AOS2008.
- 4 Richard F Blute, J Robin B Cockett, and Robert AG Seely. Cartesian differential categories. *Theory and Applications of Categories*, 22(23):622–672, 2009.
- 5 Nawaf Bou-Rabee, Bob Carpenter, and Milo Marsden. GIST: Gibbs self-tuning for locally adaptive Hamiltonian Monte Carlo. *arXiv preprint*, 2024. arXiv:2404.15253.
- 6 Fang Chen, Laszlo Lovasz, and Igor Pak. Lifting Markov chains to speed up mixing. In *The Annual ACM Symposium on Theory of Computing*, pages 275–281. ACM, 1999. doi:10.1145/301250.301315.
- 7 Kenta Cho and Bart Jacobs. Disintegration and Bayesian inversion via string diagrams. *Mathematical Structures in Computer Science*, 29(7):938–971, 2019. doi:10.1017/S0960129518000488.
- 8 Kenta Cho, Bart Jacobs, Bas Westerbaan, and Abraham Westerbaan. An introduction to effectus theory, 2015. arXiv:1512.05813.
- 9 Marco Cusumano-Towner, Alexander K. Lew, and Vikash K. Mansinghka. Automating involutive mcmc using probabilistic and differentiable programming, 2020. arXiv:2007.09871.
- 10 Persi Diaconis, Susan Holmes, and Radford M. Neal. Analysis of a nonreversible Markov chain sampler. *The Annals of Applied Probability*, 10(3):726–752, 2000. doi:10.1214/aoap/1019487508.
- 11 Tobias Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. *Advances in Mathematics*, 370:107239, 2020. doi:10.1016/j.aim.2020.107239.
- 12 Tobias Fritz, Fabio Gadducci, Paolo Perrone, and Davide Trotta. Weakly markov categories and weakly affine monads. *arXiv preprint*, 2023. arXiv:2303.14049.
- 13 Tobias Fritz, Tomáš Gonda, Antonio Lorenzin, Paolo Perrone, and Dario Stein. Absolute continuity, supports and idempotent splitting in categorical probability. *arXiv preprint*, 2023. doi:10.48550/arXiv.2308.00651.
- 14 Tobias Fritz, Tomáš Gonda, and Paolo Perrone. De Finetti’s Theorem in Categorical Probability. *Journal of Stochastic Analysis*, 2(4), 2021. doi:10.31390/josa.2.4.06.
- 15 Richard Garner and Jean-Simon Pacaud Lemay. Cartesian differential categories as skew enriched categories. *Applied Categorical Structures*, 29(6):1099–1150, 2021. doi:10.1007/s10485-021-09649-7.
- 16 Nathan Glatt-Holtz, Justin Krometis, and Cecilia Mondaini. On the accept-reject mechanism for Metropolis–Hastings algorithms. *Annals of Applied Probability*, 33(6B):5279–5333, 2023.
- 17 Nathan E Glatt-Holtz, Andrew J Holbrook, Justin A Krometis, and Cecilia F Mondaini. Parallel MCMC algorithms: theoretical foundations, algorithm design, case studies. *Transactions of Mathematics and Its Applications*, 8(2), 2024. doi:10.1093/imatrm/tnae004.
- 18 Stefano Gogioso and Carlo Maria Scandolo. Categorical probabilistic theories. *Electronic Proceedings in Theoretical Computer Science*, 266:367–385, February 2018. doi:10.4204/eptcs.266.23.
- 19 Peter J Green. Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. *Biometrika*, 82(4):711–732, 1995.
- 20 Paul R Halmos. *Measure Theory*, volume 18. Springer, 1974.

- 21 W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, April 1970. doi:10.1093/biomet/57.1.97.
- 22 André Joyal and Ross Street. The geometry of tensor calculus, I. *Advances in Mathematics*, 88(1):55–112, 1991. doi:10.1016/0001-8708(91)90003-P.
- 23 Olav Kallenberg. *Foundations of Modern Probability*, volume 99 of *Probability Theory and Stochastic Modelling*. Springer, 3 edition, 2021. doi:10.2307/2669881.
- 24 Gregory Maxwell Kelly. *Basic concepts of enriched category theory*, volume 64. CUP Archive, 1982.
- 25 Tiange Liu, Nikola Surjanovic, Miguel Biron-Lattes, Alexandre Bouchard-Côté, and Trevor Campbell. AutoStep: Locally adaptive involutive MCMC. *arXiv preprint*, 2024. doi:10.48550/arXiv.2410.18929.
- 26 Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- 27 Sean Moss and Paolo Perrone. A category-theoretic proof of the ergodic decomposition theorem. *Ergodic Theory and Dynamical Systems*, 43(12):4166–4192, 2023. doi:10.1017/ETDS.2023.6.
- 28 Kirill Neklyudov, Max Welling, Evgenii Egorov, and Dmitry Vetrov. Involutive MCMC: a Unifying Framework. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pages 7273–7282. PMLR, 2020. arXiv:2006.16653.
- 29 Eigil Fjeldgren Rischel and Tobias Fritz. Infinite products and zero-one laws in categorical probability. *Compositionality*, Volume 2 (2020)(3), 2020. doi:10.32408/COMPOSITIONALITY-2-3.
- 30 Gareth O. Roberts and Jeffrey S. Rosenthal. General state space markov chains and mcmc algorithms. *Probability Surveys*, 1(none), January 2004. doi:10.1214/154957804100000024.
- 31 Yuji Sakai and Koji Hukushima. Eigenvalue analysis of an irreversible random walk with skew detailed balance conditions. *Physical Review E*, 93(4):3–8, 2016. doi:10.1103/PhysRevE.93.043318.
- 32 Peter Selinger. A survey of graphical languages for monoidal categories. In *New structures for physics*, pages 289–355. Springer, 2010.
- 33 Sam Staton. Commutative semantics for probabilistic programming. In H Yang, editor, *Programming Languages and Systems*, volume 10201 of *Lecture Notes in Computer Science*, pages 855–879. Springer, Berlin, Heidelberg, 2017. doi:10.1007/978-3-662-54434-1_32.
- 34 Achille Thin, Nikita Kotelevskii, Christophe Andrieu, Alain Durmus, Eric Moulines, and Maxim Panov. Nonreversible mcmc from conditional invertible transforms: a complete recipe with convergence guarantees, 2021. arXiv:2012.15550.
- 35 Luke Tierney. A note on Metropolis-Hastings kernels for general state spaces. *The Annals of Applied Probability*, 8(1):1–9, 1998. doi:10.1214/aoap/1027961031.
- 36 Sean Tull. A categorical reconstruction of quantum theory. *Logical Methods in Computer Science*, Volume 16, Issue 1, January 2020. doi:10.23638/lmcs-16(1:4)2020.
- 37 Konstantin S. Turitsyn, Michael Chertkov, and Marija Vucelja. Irreversible Monte Carlo algorithms for efficient sampling. *Physica D: Nonlinear Phenomena*, 240(4-5):410–414, 2011. doi:10.1016/j.physd.2010.10.003.
- 38 Zuheng Xu and Trevor Campbell. Asymptotically exact variational flows via involutive MCMC kernels. *arXiv preprint*, 2025. arXiv:2506.02162.