Absolute continuity, supports and idempotent splitting in Markov categories (extended abstract)

Tobias Fritz, Tomáš Gonda, Antonio Lorenzin, Paolo Perrone and Dario Stein

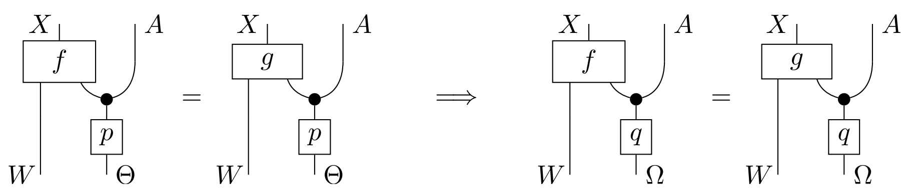
Consider a discrete probability distribution p on a finite set A. There is an unambiguous notion of the *support* of p—it is the set S of elements of A that are assigned non-zero probability by p. Given two distributions, p and q, one also says that q is *absolutely continuous* with respect to p, denoted $q \ll p$, if every property of elements of A that holds with probability 1 according to p also holds with probability 1 according to q. If one moves beyond the discrete case, however, it is not immediately clear how to define these notions. This applies even more so in the abstract setting of **Markov categories**, an abstract approach to probabilistic processes and information flow of increasing popularity [2–11].

In this work, we propose well-behaved definitions of these and explore their properties in the context of Markov categories. This comprises **absolute continuity** and **supports**, but also structured **idempotents**, which are closely linked. Our definitions of the first two concepts are notably improved over ones suggested in previous works, where absolute continuity had made a brief appearance in [5, Definition 2.8], and a candidate definition of supports was proposed in [3, Definition 13.20]. Thanks to these improvements, we now obtain desirable properties like compatibility with tensor products, and the existence of a structured functor to Rel, which generalizes the functor $\mathsf{FinStoch} \to \mathsf{Rel}$ assigning to every stochastic matrix its underlying relation. Moreover, all of our abstract notions recover the traditional ones from measure theory.

Our main result (Theorem 2 below) is a measure-theoretic one: It says that all idempotents in $\mathsf{BorelStoch}$, the category of standard Borel measurable spaces and Markov kernels, split. Furthermore, we expect splitting of idempotents to play a fundamental role in our ongoing work on ergodic theory [11].

We now sketch the main ideas in detail. The support S from above is characterized by the property that a general probability measure on A, seen as a stochastic matrix $q \colon I \to A$ from the singleton set I to A, factors through the inclusion $S \hookrightarrow A$ if and only if $q \ll p$ holds. This idea can be extended to a universal property of S, as follows.

Definition 1 (Absolute continuity and support)**.** *Given morphisms $p \colon \Theta \to A$ and $q \colon \Omega \to A$ in a Markov category C, the absolute continuity relation $\boldsymbol{q \ll p}$ holds if for any $f, g \colon W \otimes A \to X$ we have*



*Moreover, a **support** for p is a natural isomorphism*

$$\mathsf{C}(\_, S) \cong \mathsf{C}(\_, A)_{\ll p} := \{ q \colon \_ \to A \mid q \ll p \}$$

for a suitable representing object S, which additionally restricts to a bijection between the corresponding subsets of deterministic morphisms.

In words, $q \ll p$ says that p-a.s. equality implies q-a.s. equality. This notion of absolute continuity differs from [5, Definition 2.8] in that it allows for an additional input W. Thanks to W, we obtain compatibility of $\ll$ with the tensor product, and we have examples showing that such compatibility would fail without W. We also show that under mild assumptions on the Markov category, which often hold in practice, the definition is unchanged by omitting W. Further, this synthetic version of absolute continuity do in fact give the classical absolute continuity described at the beginning of this abstract.

Concerning supports, we prove that a support is characterized by a deterministic monomorphism $\mathsf{si}_p \colon S \to A$ satisfying $\mathsf{si}_p \ll p$ and such that for every $q \colon \Omega \to A$, $q \ll p$ implies that q factors through si_p. It follows that precomposition with si_p classifies equality almost surely, i.e. we have

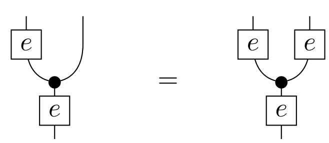
$$f =_p g \qquad \iff \qquad f \circ \mathsf{si}_p = g \circ \mathsf{si}_p.$$

However, abstract supports as in Definition 1 are not the only possible notion of support. For example, a support inclusion $S \hookrightarrow A$ in $\mathsf{FinStoch}$ is not just monic, but in fact split monic. Requiring that si_p splits gives rise to a stronger notion that we call a **split support**. We show that the split support of the tensor product is the product of their split supports and we develop a stronger universal property of split supports extending the original one of [3, Definition 13.20]. While Definition 1 gives a *mapping-in* universal property for the support, this one is surprisingly a *mapping-out* universal property.

We also show that in the Markov category $\mathsf{BorelStoch}$ of standard Borel spaces and Markov kernels, a probability measure has a support if and only if it is atomic. We recall that an atomic measure is a measure admitting elements with nonzero probability. Roughly speaking, the non-existence of a support for a non-atomic measure formalizes the idea that generic probability spaces do not have a minimal "region" in which all the mass is present. Furthermore, we prove that a non-atomic measure cannot even have a support in any Markov category containing $\mathsf{BorelStoch}$ as a Markov subcategory: This follows by showing that every morphism admitting a support satisfy a synthetic version of atomicity. In particular, we are able to extend this result to any causal Markov category.

The last part of our work studies idempotents in Markov categories. Our examples will show that these come up naturally in the context of conditioning, and they are also closely related to the notion of a support. As we will expand on in future work on the law of large numbers, idempotents are also especially relevant in the context of ergodic theory.

There are three basic ways in which an idempotent can interact with the Markov category structure, resulting in the definitions of **balanced idempotents**, **strong idempotents** and **static idempotents**. For example, a balanced idempotent is an idempotent $e \colon A \to A$ satisfying



The other two are characterized by similar-looking string diagrams. In the most standard examples of Markov categories, all idempotents are balanced, and this follows from a synthetic version of the Cauchy-Schwarz inequality.

In a positive Markov category C, any split idempotent $e = \iota \circ \pi$, with $\pi \circ \iota = \mathrm{id}$, is automatically balanced. Moreover, split idempotents are strong (resp. static) if and only if π (resp. ι) is deterministic.

Under positivity, we can also characterize a split support of a morphism p as equivalent to a static split idempotent e satisfying $e \circ p = p$ and equalizing all morphisms that are p-a.s. equal.

Finally, we show that if all static idempotents have a support under suitable assumptions on C, then all idempotents are split idempotents. This result can be applied to prove the following new measure-theoretic result.

Theorem 2. *Every idempotent Markov kernel on a standard Borel measurable space splits.*

Blackwell's historical result on idempotents [1, Theorem 7] can be derived from Theorem 2 above, while the converse is not immediate. In this sense, our result is stronger than Blackwell's theorem.

Acknowledgements. Tobias Fritz and Antonio Lorenzin acknowledge support from the Austrian Science Fund (FWF) through project P 35992-N. Paolo Perrone acknowledges support from the European Research Council through Sam Staton's grant "BLaSt – a Better Language for Statistics". Tomáš Gonda acknowledges support from the Austrian Science Fund (FWF) via the START Prize Y1261-N.

References

[1] David Blackwell. Idempotent Markoff chains. *Ann. of Math. (2)*, 43:560–567, 1942.

[2] Kenta Cho and Bart Jacobs. Disintegration and Bayesian inversion via string diagrams. *Math. Structures Comput. Sci.*, 29:938–971, 2019.

[3] Tobias Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. *Adv. Math.*, 370:107239, 2020.

[4] Tobias Fritz, Tomáš Gonda, Nicholas Gauguin Houghton-Larsen, Antonio Lorenzin, Paolo Perrone, and Dario Stein. Dilations and information flow axioms in categorical probability. arXiv:2211.02507.

[5] Tobias Fritz, Tomáš Gonda, Paolo Perrone, and Eigil Fjeldgren Rischel. Representable Markov categories and comparison of statistical experiments in categorical probability. *Theoretical Computer Science*, page 113896, 2023. arXiv:2010.07416.

[6] Tobias Fritz and Andreas Klingler. The d-separation criterion in categorical probability. *J. Mach. Learn. Res.*, 24(46):1–49, 2023. arXiv:2207.05740.

[7] Tobias Fritz and Wendong Liang. Free gs-monoidal categories and free Markov categories. *Appl. Categ. Structures*, 31(2):Paper No. 21, 2023. arXiv:2204.02284.

[8] Tobias Fritz and Eigil Fjeldgren Rischel. Infinite products and zero-one laws in categorical probability. *Compositionality*, 2:3, 2020.

[9] Bart Jacobs. Multinomial and hypergeometric distributions in Markov categories. In *Proceedings of the Thirty-Seventh Conference on the Mathematical Foundations of Programming Semantics (MFPS*, volume 351 of *Electron. Notes Theor. Comput. Sci.*, pages 98–115, 2021. arXiv:2112.14044.

[10] Sean Moss and Paolo Perrone. Probability monads with submonads of deterministic states. In *Proceedings of the 37th Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 1–13, 2022.

[11] Sean Moss and Paolo Perrone. A category-theoretic proof of the ergodic decomposition theorem. *Ergodic Theory Dynam. Systems*, pages 1–27, 2023. arXiv:2207.07353.