The Aldous–Hoover Theorem in Categorical Probability (extended abstract)

Leihao Chen, Tobias Fritz, Tomáš Gonda, Andreas Klingler, Antonio Lorenzin

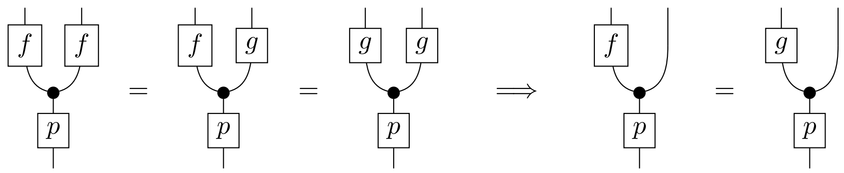
May 28, 2024

The theory of Markov categories [4, 7] is an abstract framework for probability theory, which has recently gained prominence. It is distinct from traditional measure-theoretic treatments and applicable also to other theories of uncertainty. A *Markov* category is one that satisfies several basic properties of concrete categories with Markov kernels as morphisms. The string-diagrammatic language available allows an explicit account of conditional independence of probabilistic processes. Adding additional axioms with clear interpretation (e.g. conditionals, representability, ...) provides all the structure needed to express and prove categorical versions of classical results in probability theory and statistics [1, 5, 8–11, 13, 14]. The fact that this approach axiomatizes how probabilities *behave* instead of defining what probabilities *are* makes it 'synthetic'.

In the present work, we extend the existing theory by stating and proving a synthetic version of the **Aldous–Hoover theorem**, which is arguably the deepest result of probability theory that has been developed synthetically so far. Before we turn to its statement, let us provide a bit of background first. The **de Finetti theorem** asserts that if the joint probability distribution of an infinite sequence of random variables remains unchanged under finite permutations, then these variables are conditionally independent given an appropriate random variable. This theorem holds substantial technical and philosophical significance: For example, it extends the law of large numbers to such permutation-invariant sequences; and it plays a significant role in the longstanding debate on the subjective vs. objective view of probability.

What happens when we have an infinite *matrix* of random variables instead of a mere infinite sequence? This is what the Aldous–Hoover theorem [2, 12] addresses. This more recent result of measure-theoretic probability is an analogue of the de Finetti theorem for matrices that display permutation invariance, now with respect to permuting rows and columns separately. If one thinks of the matrix entries as the colors of edges in an infinite complete bipartite graph, then the theorem characterizes these distributions as mixtures of distributions constructed as follows: assign to each vertex independently a random 'label', and then assign to each edge a color that depends only on the labels of its endpoints. This formulation also underlines connections to the theory of random graphs and networks [6].

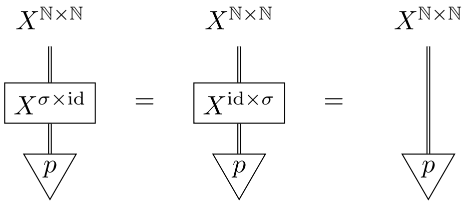
In this work, our main new result is a generalization of the Aldous–Hoover theorem to Markov categories, based on axioms related to those appearing in our earlier proof of the de Finetti theorem [8]. To state it, we first need to mention a new information flow axiom, which we also introduce in the present paper:



(1) {eq:spat_caus

For the Markov categories that describe measure-theoretic probability, this axiom reduces to saying that whenever $(f-g)^2$ is p-almost surely equal to 0, then f is p-almost surely equal to g, which is evidently true. Just like the classical Aldous–Hoover theorem deals with matrices of random variables that satisfy the relevant permutation invariance, our synthetic generalization applies to morphisms $p : I \to X^{\mathbb{N}\times\mathbb{N}}$

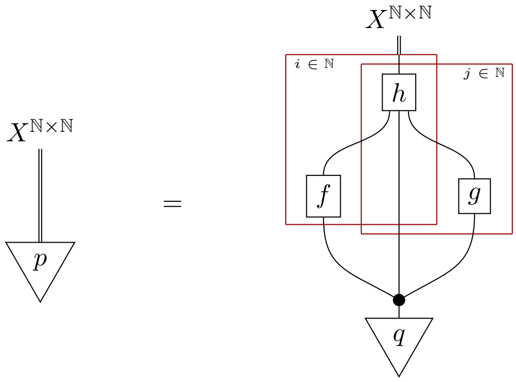
that are invariant under permutations of either of the two indexes:



(2)

for all finite permutations $\sigma : \mathbb{N} \to \mathbb{N}$. Here, $X^{\sigma\times\mathrm{id}}$ and $X^{\mathrm{id}\times\sigma}$ stand for the permutations of rows and columns respectively. We call such morphisms *row-and-column-exchangeable.* The result is now as follows:

ldous_hoover **Theorem 1** (synthetic Aldous–Hoover theorem)**.** *Let* C *be a Markov category with conditionals, countable Kolmogorov products, and satisfying implication* (1)*. Then a morphism* $p : I \to X^{\mathbb{N}\times\mathbb{N}}$ *is row-and-column-exchangeable if and only if it can be written in the form*



(3) {eq:aldous_h

for some morphisms f*,* g*,* h *and* q *and suitable objects on the intermediate wires.*[1]

Here, the red rectangles represent **plate notation**, which is a shorthand for drawing infinite copies of wires and feeding each copy into a copy of the same subdiagram inside the box. This notation extends the existing plate notation for Bayesian networks [3] to string diagrams.[2] So in the string diagram above, f and g are copied $\mathbb{N}$ times (with one copy for each column and row respectively), while h is copied $\mathbb{N} \times \mathbb{N}$ times (with one copy for each entry of the matrix). In terms of the random graph picture mentioned above, f generates a label for each 'row' vertex i, while g generates a label for each 'column' vertex j, and h generates a color for each edge (i, j) based on the labels of its endpoints.

We prove this theorem by several suitable applications of our synthetic de Finetti theorem[3] together with (1) to derive a number of conditional independence relations, and then use standard implications between conditional independence relations in order to prove compatibility with the relevant causal structure by an argument reminiscent of the d-separation criterion [9]. When instantiated in measure-theoretic probability (meaning in the Markov category BorelStoch), our Theorem 1 recovers the classical Aldous–Hoover theorem.

References

randomgraphs [1] N. Ackerman, C. E. Freer, Y. Kaddar, J. Karkwowski, S. Moss, D. Roy, S. Staton, and H. Yang. Probabilistic programming interfaces for random graphs: Markov categories, graphons, and nominal sets. In *Proceedings of the ACM on Programming Languages*, volume 8, pages 1819–1849. ACM, 2024. arXiv:2312.17127.

[1] In the proof, we construct these intermediate objects, forming the codomain of f and g and the domain of h, also as Kolmogorov powers of X.

[2] Moreover, these plates are closely related to functorial boxes.

[3] For which we also present a new version of the proof from [8] where the assumption of representability is replaced by the new information flow axiom expressed as implication (1) here.

presentations [2] D. J. Aldous. Representations for partially exchangeable arrays of random variables. *Journal of Multivariate Analysis*, 11(4):581–598, 1981.

94operations [3] W. L. Buntine. Operations for learning with graphical models. *Journal of Artificial Intelligence Research*, 2:159–225, 1994.

s2019strings [4] K. Cho and B. Jacobs. Disintegration and Bayesian inversion via string diagrams. *Math. Structures Comput. Sci.*, 29:938–971, 2019. arXiv:1709.00322.

23evidential [5] E. Di Lavore and M. 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.

nis2008graph [6] P. Diaconis and S. Janson. Graph limits and exchangeable random graphs. *Rend. Mat. Appl. (7)*, 28(1):33–61, 2008. arXiv:0712.2749.

019synthetic [7] T. Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. *Adv. Math.*, 370:107239, 2020. arXiv:1908.07021.

021definetti [8] T. Fritz, T. Gonda, and P. Perrone. de Finetti's theorem in categorical probability. *J. Stoch. Anal.*, 2(4), 2021. arXiv:2105.02639.

2dseparation [9] T. Fritz and A. Klingler. The d-separation criterion in categorical probability. *J. Mach. Learn. Res.*, 24(46):1–49, 2023. arXiv:2207.05740.

tz2024hidden [10] T. Fritz, A. Klingler, D. McNeely, A. Shah-Mohammed, and Y. Wang. Hidden markov models and the bayes filter in categorical probability, 2024. arXiv:2401.14669.

l2019zeroone [11] T. Fritz and E. F. Rischel. The zero-one laws of Kolmogorov and Hewitt–Savage in categorical probability. *Compositionality*, 2:3, 2020. compositionality-journal.org/papers/compositionality-2-3.

oover1982row [12] D. N. Hoover. Row-column exchangeability and a generalized model for probability. *Exchangeability in probability and statistics (Rome, 1981)*, pages 281–291, 1982.

s2022ergodic [13] S. Moss and P. Perrone. A category-theoretic proof of the ergodic decomposition theorem. *Ergodic Theory Dynam. Systems*, 43(12):4166–4192, 2023. arXiv:2207.07353.

4martingales [14] P. Perrone and R. Van Belle. Convergence of martingales via enriched dagger categories, 2024. arXiv:2404.15191.