# Intership proposal:
# Exploring the functoriality of approximate posteriors for sheaf-structured models

Grégoire Sergeant-Perthuis*

Graphical models are widely used families of probability distributions that capture conditional independence relations between features, i.e., random variables $X_i$, $i \in I$, of a signal (see Chapter 16 [1] or [2], Chapter 8 [3]). Celebrated examples are Hidden Markov models and Bayesian Networks [4]. Graphical models and their extensions are especially well suited for situations in which the modeler possesses prior knowledge about the structure of interactions among the features. This is why they are widely used to analyze time series [5], for expert systems [6], structured molecules [7,8]. Inference on graphical models and their generalizations, such as factor graphs [9, 10], can be expressed as optimizing a free energy over a contravariant functor [11–15] assimilated to a sheaf (see Chapter 2 [16]). Its source is a partially ordered set, and its target is **FinStoch** (see [17]), the category of finite sets with stochastic matrices as morphisms. The algorithms that perform approximate inference on these sheaves are message-passing algorithms, one example of which is the Belief Propagation algorithm.

The subject of the internship falls within recent efforts to exploit geometric and topological structures of signals and models to improve processing and to further theoretical understanding of such processing (see geometric deep learning [18, 19], sheaves in data science [11, 13–16, 20–23], and categorical approaches in machine learning and probability [14, 17, 24–30]).

The aim of the internship is to understand the impact that transformations and deformations of the underlying sheaf have on the local minima of the free energy, as well as on the associated message-passing algorithm. For a presentation of the approach and recent results in this direction, see the slides of the presentation at *Commutative Geometry & Higher Structures* [31]. The internship will take place at Sorbonne Université under the supervision of Grégoire Sergeant-Perthuis. The duration of the internship is 6 months and the remuneration will be in line with the standard rates for internships. We are looking for a motivated candidate with a background in mathematics or

*LCQB, Sorbonne Université

theoretical computer science, who is eager to apply their knowledge of discrete structures (graphs, posets), geometry, or category theory to machine learning.

## References


[1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.

[2] K. Gimpel and D. Rudoy, *Statistical inference in graphical models*. Massachusetts Institute of Technology, Lincoln Laboratory, 2008.

[3] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[4] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, 1988.

[5] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[6] A. P. Dawid, "Applications of a general propagation algorithm for probabilistic expert systems," *Statistics and Computing*, vol. 2, pp. 25–36, 1992.

[7] Z. Smith and P. Tiwary, "Making High-Dimensional Molecular Distribution Functions Tractable through Belief Propagation on Factor Graphs," *The Journal of Physical Chemistry B*, vol. 125, pp. 11150–11158, Oct. 2021. Publisher: American Chemical Society.

[8] P. Tian, "Factorization in Molecular Modeling and Belief Propagation Algorithms," preprint, PHYSICAL SCIENCES, Oct. 2021.

[9] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Generalized belief propagation," in *Proceedings of the 13th International Conference on Neural Information Processing Systems*, NIPS'00, (Cambridge, MA, USA), p. 668–674, MIT Press, 2000.

[10] J. Yedidia, W. Freeman, and Y. Weiss, "Constructing Free-Energy Approximations and Generalized Belief Propagation Algorithms," *IEEE Transactions on Information Theory*, vol. 51, pp. 2282–2312, July 2005.

[11] O. Peltre, "Message passing algorithms and homology," 2020. Ph.D. thesis, Link.

[12] O. Peltre, "A homological approach to belief propagation and Bethe approximations," in *International Conference on Geometric Science of Information*, pp. 218–227, Springer, 2019.

[13] G. Sergeant-Perthuis, "Regionalized optimization," 2022. `https://arxiv.org/abs/2201.11876`.

[14] G. Sergeant-Perthuis, "Compositional statistical mechanics, entropy and variational free energy (Extended abstract)," in *7th International Conference on Applied Category Theory (ACT 7)*, (Oxford, UK, France), David Jaz Myers and Michael Johnson, June 2024.

[15] G. Sergeant-Perthuis and N. Ruet, "Inference on diagrams in the category of Markov kernels (Extended abstract)," in *7th International Conference on Applied Category Theory (ACT 7)*, (Oxford (UK), United Kingdom), David Jaz Myers and Michael Johnso, June 2024.

[16] J. Curry, *Sheaves, cosheaves and applications*. PhD thesis, The University of Pennsylvania, 2013. arXiv:1303.3255.

[17] T. Fritz, "A synthetic approach to markov kernels, conditional independence and theorems on sufficient statistics," *Advances in Mathematics*, vol. 370, p. 107239, 2020.

[18] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, "Geometric deep learning: Grids, groups, graphs, geodesics, and gauges," 2021.

[19] G. Sergeant-Perthuis, J. Maier, J. Bruna, and E. Oyallon, "On non-linear operators for geometric deep learning," in *Advances in Neural Information Processing Systems* (S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, eds.), vol. 35, pp. 10984–10995, Curran Associates, Inc., 2022.

[20] J. Hansen and R. Ghrist, "Distributed optimization with sheaf homological constraints," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 565–571, 2019.

[21] J. Hansen and R. Ghrist, "Toward a spectral theory of cellular sheaves," *Journal of Applied and Computational Topology*, vol. 3, pp. 315 – 358, 2018.

[22] C. Bodnar, F. D. Giovanni, B. P. Chamberlain, P. Lio, and M. M. Bronstein, "Neural sheaf diffusion: A topological perspective on heterophily and oversmoothing in GNNs," in *Advances in Neural Information Processing Systems* (A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho, eds.), 2022.

[23] O. Peltre, "A homological approach to belief propagation and bethe approximations," in *Geometric Science of Information* (F. Nielsen and F. Barbaresco, eds.), (Cham), pp. 218–227, Springer International Publishing, 2019.

[24] T. Fritz and P. Perrone, "A probability monad as the colimit of spaces of finite samples," *Theory and Applications of Categories*, 2019.

[25] P. Baudot and D. Bennequin, "The homological nature of entropy," *Entropy*, vol. 17, no. 5, pp. 3253–3318, 2015.

[26] Y. Jia, G. Peng, Z. Yang, and T. Chen, "Category-theoretical and topos-theoretical frameworks in machine learning: A survey," 2024.

[27] J. P. Vigneaux, "Information structures and their cohomology," *Theory and Applications of Categories*, vol. 35, no. 38, pp. 1476–1529, 2020.

[28] M. Marcolli, "Gamma spaces and information," *Journal of Geometry and Physics*, 2019.

[29] Y. Manin and M. Marcolli, "Homotopy Theoretic and Categorical Models of Neural Information Networks," *Compositionality*, vol. Volume 6 (2024), Sept. 2024.

[30] G. Sergeant-Perthuis, "Compositional statistical mechanics, entropy and variational inference," in *Twelfth Symposium on Compositional Structures (SYCO 12)*, 2024.

[31] G. Sergeant-Perthuis, "Functoriality of inference on diagrams in the category of Markov kernels.," in *Non Commutative Geometry & Higher Structures*, (Thessaloniki, Greece), P. Batakidis and F. Bonechi and A. S. Cattaneo and N. Ciccoli and F. D' Andrea and D. Fiorenza and M. Jotz and F. Petalidou and M. Schiavina and P. Xu, Sept. 2024. `https://hal.sorbonne-universite.fr/hal-04706267v2/file/hal-presentation.pdf`.