Master Internship proposal

# Graph Sampling in Graph Machine Learning

In the last few years, **Graph Machine Learning** [1] went from a subfield of Machine Learning (ML) to a primary citizen in the ML landscape. It encompasses many applications in chemistry, biology, or recommender systems, to name a few, and many algorithms and models, from early spectral approaches to complex Graph Neural Networks.

Many methods in Graph ML use *node sampling* as an intermediate process [3], mainly to mitigate the high memory footprint and time complexity due to the graphs' ever-increasing sizes. In particular, many classical Graph Neural Networks (GNNs) use some kind of node sampling within their architecture or during training.

However, node sampling is far from being a simple endeavor in general [2], as nodes are connected and interdependent. Thus, many methods exist to sample nodes of a graph, from simple uniform sampling to determinantal point processes [4] to more heuristic methods or even learning-based methods [5]. Despite this, there are few, if any, rigorous evaluation of the effect of different sampling processes at the various stages of graph ML algorithms.

In this internship, potentially followed by a PhD, we will thoroughly evaluate different graph sampling methods for graph ML and GNNs, both empirically and, when possible, theoretically. We will in particular study non-uniform repulsive point processes such as determinantal point processes, whose strong properties have not yet been explored in the context of GNNs. The balance between theoretical and empirical studies may depend on the candidate.

**Infos.** Location: IRISA, Rennes, France. Starting date in 2026. Funded by the MALAGA ERC Starting Grant: `https://nkeriven.github.io/malaga`

**Contact.** Please send CV and short cover letter to `nicolas.keriven@cnrs.fr`

## References

[1] Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv:2104.13478*, 2021.

[2] Pili Hu and Wing Cheong Lau. A survey and taxonomy of graph sampling. pages 1–34, 2013.

[3] Xin Liu, Mingyu Yan, Lei Deng, Guoqi Li, Xiaochun Ye, and Dongrui Fan. Sampling methods for efficient training of graph convolutional networks: A survey. *IEEE/CAA Journal of Automatica Sinica*, 9:205–234, 2022.

[4] Nicolas Tremblay, Pierre Olivier Amblard, and Simon Barthelmé. Graph sampling with determinantal processes. *25th European Signal Processing Conference, EUSIPCO 2017*, 2017-Janua:1674–1678, 2017.

[5] Taraneh Younesian, Daniel Daza, Emile van Krieken, Thiviyan Thanapalasingam, and Peter Bloem. Grapes: Learning to sample graphs for scalable graph neural networks. 2023.