Compression and Efficient Processing of Graph Data: From Signal Processing to Deep Learning

▶ **Research theme:** Graph signal processing, graph machine learning, geometric deep learning

▶ **Keywords:** Graph signals, compression, graph coarsening

▶ **Research groups:** Télécom Paris, LTCI, Télécom SudParis, SAMOVAR, Institut Polytechnique de Paris, University of Southern California, STAC Lab

▶ **Advisors:** Jhony H. Giraldo (jhony.giraldo@telecom-paris.fr),
Aref Einizade (aref.einizade@telecom-paris.fr), Antonio Ortega (aortega@usc.edu)

▶ **Starting date and duration**: 5-6 months starting in Spring 2026 (March/April)

Context

Graph signal processing (GSP) [1] and graph machine learning (GML) [2] are research fields that aim to generalize classical concepts from signal processing and machine learning to data defined on graphs. For example, in GSP, classical operations such as filtering, sampling, and reconstruction are extended to graph signals, *i.e.*, functions defined over the vertices of a graph [3]. In GML, neural network architectures are, for example, adapted to graph-structured data, leading to models such as graph neural networks [4], message passing neural networks [5], and their physics-informed geometric counterparts [6,7].

GSP and GML are crucial because graph-structured data appear in numerous applications, including social networks, the web, recommender systems, biological networks, and knowledge graphs, among others [8,9]. A common challenge across these domains is the massive scale of the underlying graphs, which often contain millions or even billions of nodes and edges [10]. Therefore, it becomes essential to compress both the graph topology and the signals defined on it, which also enables efficient learning and processing on the compressed data.

In the GSP literature, a large body of work has addressed the problems of sampling and reconstruction of graph signals in a principled manner [3, 11, 12]. These works provide theoretical guarantees for reconstructing graph signals from samples based on spectral properties of the graph, guiding the design of optimal or near-optimal sampling sets [13, 14]. Concurrently, in the GML community, autoencoder-based architectures have been developed to learn compressed latent representations of graph data, such as graph autoencoders, variational graph autoencoders, and differentiable graph pooling methods [15,16].

Despite these advances, the problem of *jointly compressing and decoding both the graph topology and the graph signals* remains a major open challenge. This task has broad implications, ranging from graph compression and storage to efficient processing of relational datasets in domains such as social networks, the web, and biological systems. This internship aims to develop robust and mathematically sound methodologies for the joint compression of graph data and topology. Our ultimate goal is to achieve more efficient GSP and GML by developing new methods to reduce the size of the graph and graph signal as much as possible without significantly affecting the task outcome compared to directly using the original graph and signals.

Candidate profile

We are looking for candidates:

- Currently pursuing an M2 in engineering, data science, computer science, applied mathematics, signal processing, statistics, or equivalent, with a strong background in signal processing and machine learning. The student should have a genuine interest in working in graph signal processing and geometric deep learning.
- Have strong programming skills in Python (including PyTorch).
- Have a genuine interest in understanding the mathematics behind graph signal processing and geometric deep learning (this is a strong requirement).
- Have good communication skills.

Outstanding candidates may be considered for continuation into a PhD on the same topic, subject to performance and funding availability. This internship serves as an initial step toward developing more efficient geometric deep learning models for relational data.

Team and location

Télécom Paris and Télécom SudParis are premier engineering schools in France and constituent members of Institut Polytechnique de Paris. The University of Southern California is a leading research university in Los Angeles, California, United States. Both universities are consistently ranked among the best universities worldwide (Shanghai Ranking, QS Ranking). Télécom Paris and Télécom SudParis are located on the outskirts of Paris (around 45 minutes by train from the center of Paris) at the center of the Paris-Saclay cluster—a fast-growing research and industrial ecosystem. The internship position is part of the ongoing project *DeSNAP – Deep Simplicial Neural Networks for Advanced Geometry Processing*, funded by ANR (French National Research Agency). The student will be integrated within the MM Team at LTCI lab.

How to apply

Please send your application material (PDF format; in English) by email to Jhony H. Giraldo, Aref Einizade, and Antonio Ortega, including the following:

- A full CV.
- A motivation letter explaining your interest in the position (max 1 page).
- Transcript of records (grades).

The applications will be reviewed on a rolling basis until the position is filled.

References

[1] A. Ortega *et al.*, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[2] M. M. Bronstein *et al.*, "Geometric deep learning: going beyond Euclidean data," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 18–42, 2017.

[3] S. Chen *et al.*, "Discrete signal processing on graphs: Sampling theory," *IEEE Transactions on Signal Processing*, vol. 63, no. 24, pp. 6510–6523, 2015.

[4] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *International Conference on Learning Representations (ICLR)*, 2017.

[5] J. Gilmer *et al.*, "Neural message passing for quantum chemistry," in *International Conference on Machine Learning (ICML)*, vol. 70, 2017, pp. 1263–1272.

[6] A. Einizade, F. Malliaros, and J. H. Giraldo, "Continuous product graph neural networks," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 37, pp. 90 226–90 252, 2024.

[7] A. Einizade, D. Thanou, F. D. Malliaros, and J. H. Giraldo, "Continuous simplicial neural networks," *arXiv preprint arXiv:2503.12919, Accepted in NeurIPS 2025*, 2025.

[8] M. E. J. Newman, *Networks: An Introduction*. Oxford University Press, 2010.

[9] R. Ying *et al.*, "Graph convolutional neural networks for web-scale recommender systems," in *ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, 2018, pp. 974–983.

[10] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford large network dataset collection," http://snap.stanford.edu/data, June 2014.

[11] M. Tsitsvero, S. Barbarossa, and P. D. Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Transactions on Signal Processing*, vol. 64, no. 18, pp. 4845–4860, 2016.

[12] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Sampling of graph signals with successive local aggregations," *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1832–1843, 2016.

[13] A. Anis, A. Gadde, and A. Ortega, "Efficient sampling set selection for bandlimited graph signals using graph spectral proxies," *IEEE Transactions on Signal Processing*, vol. 64, no. 14, pp. 3775–3789, 2016.

[14] G. Puy, N. Tremblay, R. Gribonval, and P. Vandergheynst, "Random sampling of bandlimited signals on graphs," *Applied and Computational Harmonic Analysis*, vol. 44, no. 2, pp. 446–475, 2018.

[15] T. N. Kipf and M. Welling, "Variational graph auto-encoders," arXiv preprint arXiv:1611.07308, 2016.

[16] R. Ying *et al.*, "Hierarchical graph representation learning with differentiable pooling," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 31, 2018, pp. 4800–4810.