Graph Fourier Transform for Samples of Structured Graphons

Mahya Ghandehari (University of Delaware, Department of Mathematical Sciences),
Jeannette Janssen (Dalhousie University, Department of Mathematics and Statistics)

November 5 –12, 2023

1 Overview of the Field

Due to the large amount of networked data that is being collected, the recent field of Graph Signal Processing (GSP) has attracted the attention of many data scientists (in mathematics and engineering), and has turned into a fast-growing and vibrant research area. In classical signal processing, the Fourier transform and its variants play a crucial role in analysis, modifying, and synthesizing signals. This fundamental concept has been generalized to graph signals using spectral features of the underlying graph [15]. As a result, processing of a graph signal rigidly depends on the underlying network; this is a major drawback, as the underlying graph of a signal may sustain minor variations due to error or the natural evolution of the network. Thus, it becomes vital to establish instance-independent graph signal processing methods.

The theory of graph limits and graphons was introduced only in 2006 [12], and the use of graphons to develop a common scheme for signal processing on large dynamic graphs (i.e. instance-independent graph signal processing schemes) was first proposed in 2020 [14]. A *graphon* is a symmetric measurable function on $[0,1]^2$. Graphons offer a non-parametric approach to network modeling, which is very valuable when studying stochastic networks. Indeed, graphons represent random processes that generate networks, and networks produced by the same graphon have similar large-scale features. For this reason, we think of graphons as the large-scale blueprint of any graphs that they generate. Naturally, to study networks, we only need to study the (underlying) graphon, rather than the individual networks generated by that graphon [2, 11]. For recent applications of graph limit theory in developing novel GSP methods and robustness/consistency analysis of GSP, see for example [1, 3, 10, 13, 8, 9, 16].

2 Recent Developments and Open Problems

In [14], Ruiz, Chamon and Ribeiro used graphon theory to prove robustness of graph Fourier transform. Namely, they proved that for a sequence of graph signals $\{(G_n, f_n)\}$ that converge to a graphon signal (w, f), the graph Fourier transform of (G_n, f_n) converges to the graphon Fourier transform of (w, f) if the following conditions on w and f are satisfied:

(i) the graphon signal f is bandlimited,

(ii) w is a non-derogatory graphon (i.e. the integral operator associated with w does not have repeated nonzero eigenvalues.)

This important convergence result lays the ground for developing instance-independent graph signal processing, where similarity of graphs is formalized in the context of graph limit theory. We generalized this result in [7], and dropped both conditions which were imposed on w and f, allowing for a convergence theorem applicable to all graphons. The significance of this generalization lies in the fact that many important examples of graphons, including Cayley graphons on non-Abelian groups, have multi-dimensional eigenspaces. In order to obtain the general convergence result, we refined the definition of the graph/graphon Fourier transform. Inspired by Fourier analysis of non-Abelian groups, we replaced the concept of "Fourier coefficients" by projections onto eigenspaces of the shift operator. The results of [14, 7] pave the way for a compelling initiative in the development of important instance-independent GSP tools (such as preferred orthonormal basis, frames, filters, etc.) built upon the limiting graphon. This RIT workshop was focused on developing suitable GSP frameworks for graphs sampled from graphons with particular spatial structures (e.g. Cayley graphons or spatial graphons). Such graphons are usually suitable models for large networks arising from a variety of applications.

3 Scientific Progress Made

During our RIT week, we focused on the following two problems:

I. Identifying samples of multi-dimensional spatial graphons. A challenging question regarding networks generated according to spatial models is to uncover the hidden spatial layout of the network. The classical *seriation problem* is the 1-dim spatial layout problem for graphs. The objective of the seriation problem is to obtain a *Robinson ordering* for vertices of a spatial graph, i.e., to order vertices so that adjacent ones are precisely those that are placed close to each other. Although the error-free seriation problem is solved in polynomial time, its perturbed version turned out to be very difficult. The goal of our research programme in this area is to use the theory of graph limits for the approximate seriation of *almost Robinson graphs* and their higher dimensional analogues. One can think of almost Robinson graphs as samples of Robinson graphons. These are graphons whose values increase when moving towards the diagonal along any horizontal or vertical ray. In [4, 6], the PIs and collaborators introduced a function Γ on the space of graphons, and showed that it successfully formalizes the Robinson property. Namely, Γ satisfies the following essential properties on graphons:

1. **(Recognition)** $\Gamma(w) = 0$ *iff* w is a Robinson graphon [4].
2. **(Continuity)** If w is $\|\cdot\|_\square$-close to a Robinson graphon then $\Gamma(w)$ is close to 0 [4].
3. **(Stability)** If $\Gamma(w)$ is close to 0 then w is $\|\cdot\|_\square$-close to a Robinson graphon [5].

This is the first known function that allows the recognition of samples of Robinson graphons. That is, for a dense graph sequence $\{G_n\}$ converging to a graphon w, we have $\Gamma(w) \to 0$ *iff* the graphon w is Robinson.

During our stay at BIRS, we managed to extend our 1-dim results to higher-dim cases as follows. Firstly, we obtained a suitable generalization of the concept of 1-dim spatial graphs to higher dimensions. We call this new class of graphs the *spatial graphs on* $([0,1]^n, \|\cdot\|_p)$. Next, we addressed the following central question:

> *Estimate how close a network G is to being spatial on $([0,1]^n, \|\cdot\|_p)$, and find a network R that is spatial on $([0,1]^n, \|\cdot\|_p)$ and approximates G.*

To handle this question, we follow the same philosophy as in the 1-dim case. Namely, we identify the appropriate definition of *n-dim spacial graphons* with respect to norm $\|\cdot\|_p$, and define a graphon parameter Λ that acts as a gauge of spatiality on $([0,1]^n, \|\cdot\|_p)$; i.e., a function that *recognizes* graphons spatial on $([0,1]^n, \|\cdot\|_p)$, is *continuous* on the space of graphons, and provides a *stable* measurement of the spatial property. To be precise, we define Λ as follows:

Definition. *Let $w : [0,1]^n \times [0,1]^n \to [0,1]$ be a graphon. For $0 \leq d \leq 1$, define $U(d) := \{(X,Y) \in \mathbb{R}^n \times \mathbb{R}^n : \|X-Y\|_p \geq d\}$ and $L(d) := \{(X,Y) \in \mathbb{R}^n \times \mathbb{R}^n : \|X-Y\|_p \leq d\}$. Define*

$$\Lambda(w) = \sup_{0\leq d\leq 1} \sup_{\substack{A_1\times A_2\subseteq U(d)\\ B_1\times B_2\subseteq L(d),\\ |A_1\times A_2|=|B_1\times B_2|}} \left[\iint_{A_1\times A_2} w - \iint_{B_1\times B_2} w\right].$$

Roughly speaking, the parameter Λ takes the largest amount of average violation to the definition of graphon spatiality over blocks (rather than points). Using this parameter, we propose a method for obtaining spatial approximations of graphons. The spatial approximation of a given graphon w will be a graphon τ_w that is spatial on $([0,1]^n, \|\cdot\|_p)$, and is defined using certain block averages of w where the block size is informed by $\Lambda(w)$. We summarize our results below.

New results. *Let $1 \leq p \leq \infty$, and $w : [0,1]^n \times [0,1]^n \to [0,1]$ be a graphon. Then,*

(i) **(Recognition)** *w is spatial on $([0,1]^n, \|\cdot\|_p)$ a.e. if and only if $\Lambda(w) = 0$.*

(ii) **(Continuity)** *Λ is continuous with respect to cut-norm.*

(iii) **(Stability)** *There exists $\alpha > 0$ s.t. for any graphon $w : [0,1]^n \times [0,1]^n \to [0,1]$, there exists a graphon τ_w, called the* spatial approximation *of w on $([0,1]^n, \|\cdot\|_p)$, satisfying*

$$\|w - \tau_w\|_\square \leq \Lambda(w)^\alpha.$$

The significance of (iii) lies in the fact that the error of approximation is bounded in terms of cut-norm, which is the suitable norm when studying large or dynamic networks. Luckily, it turns out that the approximation τ_w is easy to formulate. However, the proof of stability is rather technical, as it needs to combine careful counting arguments with delicate norm estimates. Our results pave the way for addressing the question of how to describe almost spatial graphons in dimension n. To be more precise, we plan to prove a result of the following type:

(Recognition of samples) Suppose $\{G_n\}$ is a sequence of graphs converging to a graphon $w : [0,1] \times [0,1] \to [0,1]$. Then we have the following:

$$\exists \text{ meas.-pres. } \phi \text{ s.t. } w^\phi \text{is spatial on } ([0,1]^n, \|\cdot\|_p) \quad \Leftrightarrow \quad \exists \text{ labeling of } V(G_n) \text{ s.t. } \Lambda(G_n) \to 0.$$

II. Developing signal processing on almost–structured graphs. We dedicated a portion of our time at BIRS to the question of developing GSP methods on samples of spatial graphons. In particular, we investigated how the Robinson structure of the underlying graphon can be exploited to guide the graph signal processing of its samples. This line of research ties together our earlier work in [7] with our new result regarding the recognition of near-spatial networks. Taking the underlying structure of the graphon into account, we proposed new methods for developing signal processing tools (such as filters and frames) for the class at hand. This work is in preliminary stages, and we plan to pursue it in near future.

4 Outcome of the Meeting

We found our time in BIRS extremely fruitful and productive. We made a substantial progress extending our earlier results on 1-dim spatial graphs/graphons to the multi-dimensional case. In addition, our new results allow the use of any L^p norm for $1 \leq p \leq \infty$. Currently, we are in the process of writing a paper containing our new results, which we plan to submit to a high-quality journal for publication. In addition, we initiated a new project regarding developing new GSP methods on spatial graphons. Finally, we used our time at BIRS to make further progress on a book manuscript, which will be published by *Applied and Numerical Harmonic Analysis (ANHA)* book series in 2025.

References

[1] A. Bonato, J. Janssen, and P. Prałat, Geometric protean graphs, *Internet Mathematics*, **8** (2012), 2–28.

[2] C. Borgs and J. Chayes, Graphons: A nonparametric method to model, estimate, and design algorithms for massive networks, In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17, page 665672, New York, NY, USA, 2017. Association for Computing Machinery.

[3] M. Bradonjic, A. Hagberg, and A. G. Percus. The structure of geographical threshold graphs, *Internet Mathematics*, 4:113–139, 2009.

[4] H. Chuangpishit, M. Ghandehari, M. Hurshman, J. Janssen, and N. Kalyaniwalla. Linear embeddings of graphs and graph limits, *J. Combin. Theory Ser. B*, 113:162–184, 2015.

[5] M. Ghandehari and J. Janssen. Graph sequences sampled from Robinson graphons, Manuscript, see arXiv https://arxiv.org/abs/2005.05253.

[6] M. Ghandehari and J. Janssen. Graph sequences sampled from Robinson graphons, *European J. Combin.*, 116:Paper No. 103859, 31, 2024.

[7] M. Ghandehari, J. Janssen, and N. Kalyaniwalla. A noncommutative approach to the graphon Fourier transform, *Appl. Comput. Harmon. Anal.*, 61:101–131, 2022.

[8] P. D. Hoff. Bilinear mixed-effects models for dyadic data, *J. Amer. Statist. Assoc.*, 100(469):286–295, 2005.

[9] P. D. Hoff, A. E. Raftery, and M. S. Handcock. Latent space approaches to social network analysis, *Journal of the American Statistical Association*, 97(460):1090, 2002.

[10] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and Ma. Boguñá. Hyperbolic geometry of complex networks, *Phys. Rev. E (3)*, 82(3):036106, 18, 2010.

[11] L. Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[12] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.

[13] N. Pržulj, D. G. Corneil, and I. Jurisica. Modeling interactome: Scale-free or geometric? *Bioinformatics*, 20(18):3508–3515, 2004.

[14] L. Ruiz, L. F. O. Chamon, and A. Ribeiro. Graphon signal processing, 69:4961–4976, 2021. IEEE Trans. Signal Processing.

[15] D. I. Shuman, S. K. Narang, A. Ortega P. Frossard, and P. Vandergheynst. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains, *IEEE Signal Processing Magazine.*, 30(5):83–98, 2013.

[16] M. Tang, D. L. Sussman, and C. E. Priebe. Universally consistent vertex classification for latent positions graphs, *Ann. Statist.*, 41(3):1406–1430, 2013.