THE GRAPHON FOURIER TRANSFORM

Luana Ruiz†*, Luiz F. O. Chamon*† *and Alejandro Ribeiro*†

† Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, USA

ABSTRACT

In many network problems, graphs may change by the addition of nodes, or the same problem may need to be solved in multiple similar graphs. This generates inefficiency, as analyses and systems that are not transferable have to be redesigned. To address this, we consider graphons, which are both limit objects of convergent graph sequences and random graph models. We define graphon signals and introduce the Graphon Fourier Transform (WFT), to which the Graph Fourier Transform (GFT) is shown to converge. This result is demonstrated in two numerical experiments where, as expected, the GFT converges, hinting to the possibility of centralizing analysis and design on graphons to leverage transferability.

Index Terms— graphons, convergent graph sequences, graph filters, graph Fourier transform, graph signal processing

1. INTRODUCTION

System *transferability* is a prevalent problem in signal processing, statistics and machine learning [1, 2]. For instance, when designing information processing architectures on networks that are bound to grow, we want to avoid making adjustments every time a new node is added to the network. This is the case, for instance, of streaming services that get thousands of new users every day and whose recommendation algorithms run on user-similarity networks [3, 4]. Another example is reproducing a certain type of low-dimensional feature analysis on multiple instances of the same type of graph, eg., quantifying air pollution dispersion spectra on air quality sensor networks in different cities (cf. Section 5); it can be inefficient to have to recalculate the parameters of the transformation for each network where the analysis is repeated.

There are many reasons why redesigning a system or readjusting a transformation's parameters on different networks can be challenging. To begin with, in graphs whose number of nodes can grow, it becomes increasingly hard to measure the graph: it is either too big to obtain full measurements, or changes too frequently for any isolated measurement to be meaningful. This is a significant drawback in graph signal processing [5–8], where problems like graph filter design usually assume full knowledge of the underlying graph [9, 10]. Very large graphs are also difficult to visualize [11] and store, especially when they are not sparse. Even though this can be mitigated by dimensionality reduction techniques based on the graph's spectral decomposition [12], computing interior eigenvalues of large matrices is itself a costly operation [13, 14]. Meanwhile, problems that involve replicating an analysis or design to a set of graphs with common characteristics suffer from similar issues: when there are many of these graphs, measuring, storing, visualizing and decomposing them can quickly become an expensive task.

To avoid these issues, we propose a signal processing framework that enables a centralized and transferable approach to network systems based on *graphons*. Graphons are infinite-dimensional representations of graphs that are at once limit objects of convergent graph sequences and random graph models [15]. The former makes them good representations of large graphs; the latter provides a probability model to which a family of graphs can be tied. The use of graphons is justified by the fact that, in practical problems involving graphs that grow or families of graphs, it is often the case that a certain large scale structure is retained (eg., the edges are always connected according to the same set of rules). This is exemplified in Figure 1, where the structural similarities between different geometric graphs of same size (left and center) and of different sizes (center and right) can be seen by noticing that there only exist edges between nodes positioned within a fixed radius of one another. Graphons have been studied in multiple fields, and are the object of works that include accounting for uncertainty in the estimation of graph models [16, 17] and computing properties such as node centrality [18], clustering coefficients [19] and network game equilibria [20] in very large networks. In this paper, we lay the ground for a graphon signal processing framework by defining graphon signals as the limit objects of sequences of graph signals and by introducing the main building block of the framework — the Graphon Fourier Transform — in Def. 2. We also show that, for sequences of graphs converging to graphons, the Graph Fourier Transform converges to the Graphon Fourier Transform (Theorem 1), as suggested by the analysis done in [21] for some specific random graph models. This is an important result, as it hints to the possibility of centralizing the analysis and design of network systems on the graphon associated with a sequence or family of graphs.

In Section 5, Theorem 1 is illustrated in two numerical experiments. In the first, we draw two n-node air pollution sensor networks from a common graphon model for growing n and compare the GFTs of signals generated from the same pollution model. This experiment alludes to situations where we want to compare the spectra of the same type of signal on different graphs, in order to find similarities that are not noticeable on the vertex domain. This demonstrates the limit behavior described by Theorem 1, and, for large enough n, evidences the possibility of designing analyses and systems on one network (or on the graphon itself) and applying them to other graphs associated with the same graphon. In the second experiment, we compare the GFTs of movie ratings on small and large user networks to illustrate a more practical consequence of our results on graphs that, although inherently similar, are built from model-free data and are not related by any common generating graphon.

2. GRAPHS AND GRAPHONS

Symmetric graphs are triplets $G = (V, E, A)$ where V is a set of $n = |V|$ nodes, $E \subseteq V \times V$ is a set of edges and $A : E \to I \subseteq \mathbb{R}$ is a weight function assigning real-valued weights $A(i, j) = A(j, i)$ in the set I to edges $(i, j) \in E$. A graph is unweighted if $I = \{0, 1\}$, in which case we omit A and simply write $G = (V, E, A) = (V, E)$. Graphons are bounded, symmetric and measurable functions $\mathcal{W} : [0, 1]^2 \to [0, 1]$. If we interpret points u and v of the unit line as nodes, $\mathcal{W}(u, v)$ can be seen as the weight of the edge connecting u and v. We can thus think of $\mathcal{W}$ as a graph with an uncountable number of nodes that will serve as the limit of a sequence of graphs $\{G_n\}$ with growing number of nodes n.

To study convergence of graph sequences to graphons consider arbitrary unweighted and undirected reference graphs $F = (V', E')$. We define homomorphisms of F into G as adjacency preserving maps $\beta : V' \to V$ in which $(i, j) \in E'$ implies $(i, j) \in E$. There are a total of $|V|^{|V'|} = n^{n'}$ maps from V' to V but only some of them are adjacency preserving homomorphisms. We therefore define $\hom(F, G) =$

Supported by USA NSF CCF 1717120 and ARO W911NF1710438.

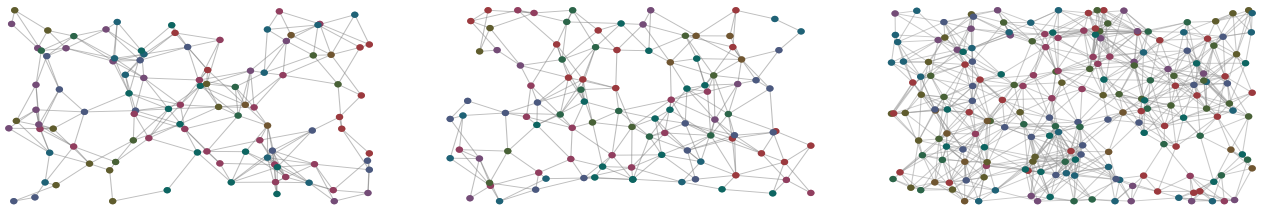

Fig. 1: Examples of graphs with $n = 50$ (left, center) and $n = 100$ (right) drawn from a random geometric graph model with radius 0.25.

$\sum_\beta \prod_{(i,j)\in E'} A(\beta(i),\beta(j))$ as a weighted sum of the total number of homomorphisms that map F into G and the density of homomorphisms as the ratio [15],

$$t(F,G) := \frac{\hom(F,G)}{n^{n'}} := \frac{\sum_\beta \prod_{(i,j)\in E'} A(\beta(i),\beta(j))}{n^{n'}}. \quad (1)$$

The interpretation of (1) is easiest when G is unweighted. In that case $\prod_{(i,j)\in E'} A(\beta(i),\beta(j)) = 1$ and we are just counting the total number of ways in which the reference graph F can be mapped into G. We can think of F as a motif and of $t(G,F)$ as how often that motif appears in G relative to the maximum number of times the motif could appear.

The homomorphism density in (1) can be generalized to graphons if we replace the sum by an integral. Thus, for a given reference motif F we define the homomorphism density of F into the graphon $\mathcal{W}$ as

$$t(F,\mathcal{W}) := \int_{[0,1]^{V'}} \prod_{(i,j)\in E'} \mathcal{W}(u_i,u_j) \prod_{i\in V'} du_i. \quad (2)$$

Given eqs. 1 and 2 we say that the graph sequence $\{G_n\}$ converges to the graphon $\mathcal{W}$ if for any given motif F it holds

$$\lim_{n\to\infty} t(F,G_n) = t(F,\mathcal{W}). \quad (3)$$

To understand the meaning of (3), it is instructive to consider graphs G_n sampled from a graphon $\mathcal{W}$, which are called $\mathcal{W}$*-random graphs*. These graphs have labels $u_i \sim U[0,1]$ drawn uniformly and independently at random from $[0,1]$, and edge sets such that $(u_i,u_j) \in E$ with probability $\mathcal{W}(u_i,u_j)$. It is possible to see that (3) holds with probability 1 for sequences of $\mathcal{W}$-random graphs. The convergence mode in (3) allows for more general generative models asides from these randomly sampled graphs. We point out that every graphon is the limit object of a sequence of convergent graphs and, conversely, that every convergent graph sequence converges to a graphon [15] — i.e., if the limit $\lim_{n\to\infty} t(F,G_n)$ exists it can be written as an integral of the form of (2).

3. GRAPH SIGNALS, GRAPHON SIGNALS AND THEIR FOURIER TRANSFORMS

We define a graph signal as the pair $(G,\mathbf{x})$ where G is the graph on which the signal is supported and $\mathbf{x} \in \mathbb{R}^n$ is such that $[\mathbf{x}]_i$ stands for the value of this signal at node i of G. The vector $\mathbf{x}$ can be seen as a *vertex representation* of data, but graph signals also admit a *spectral representation* obtained by application of the Graph Fourier Transform (GFT) [7]. This is defined with respect to a shift operator $\mathbf{S}$, which is a stand-in for a matrix representation of the graph G. Since we are working with undirected graphs we have that $\mathbf{S}$ is symmetric and that it therefore accepts an eigenvector decomposition of the form $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$ with $\mathbf{V}$ being the orthonormal matrix of eigenvectors and $\mathbf{\Lambda}$ a diagonal matrix containing the eigenvalues of $\mathbf{S}$. The GFT of $(G,\mathbf{x})$ is defined as

$$\hat{\mathbf{x}}_G = \text{GFT}\{(G,\mathbf{x})\} := \mathbf{V}^{\mathsf{H}}\mathbf{x} \quad (4)$$

which is equivalent to decomposing $(G,\mathbf{x})$ in the spectral basis of G. Where there is no ambiguity, we may omit the subscript G. Because $\mathbf{V}^{\mathsf{H}}$ is orthonormal, an inverse transformation can also be defined. We write the inverse Graph Fourier Transform (iGFT) of $\hat{\mathbf{x}}_G$ as $\text{iGFT}\{\hat{\mathbf{x}}_G\} := \mathbf{V}\hat{\mathbf{x}}_G = \mathbf{x}$. The GFT allows switching from the vertex to the spectral domain, and the iGFT provides the way back. The goal of this section is to define graphon analogous of graph signals, GFTs, and iGFTs.

3.1. Graphon Signals and Limits of Graph Signal Sequences

A graphon signal is a pair $(\mathcal{W},X)$, where $\mathcal{W}$ is a graphon and $X : [0,1] \to \mathbb{R}$ is a function mapping points $u \in [0,1]$ onto real numbers $X(u)$. We only consider finite energy graphon signals $X \in L^2$. Graphon signals can be seen as the continuous counterpart of graph signals. It will again be important to consider X in conjunction with $\mathcal{W}$.

Every graph signal $(G,\mathbf{x})$ induces a graphon signal $(\mathcal{W}_G,X_G)$. This "continuous" representation of a graph signal will be useful in the derivations carried out in Section 4. Explicitly, let I_j, $j = 1,\dots,N$, be a partition of the unit interval. Then, $(\mathcal{W}_G,X_G)$ is defined as

$$\mathcal{W}_G(u,v) = [\mathbf{S}]_{jk} \times \mathbb{I}(u \in I_j)\,\mathbb{I}(v \in I_k) \quad (5)$$

$$X_G(v) = [\mathbf{x}]_k \times \mathbb{I}(v \in I_k). \quad (6)$$

On sequences of graphs $\{G_n\}$ converging to a graphon $\mathcal{W}$, sequences of graph signals $\{(G_n,\mathbf{x}_n)\}$ converging to graphon signals $(\mathcal{W},X)$ can always be defined. We characterize their convergence in Def. 1.

Definition 1 (Convergent sequences of graph signals)**.** *A sequence of graph signals $\{(G_n,\mathbf{x}_n)\}$ is said to converge to the graphon signal $(\mathcal{W},X)$ if there exists a sequence of permutations $\{\pi_n\}$ such that*

$$t(F,G_n) \to t(F,\mathcal{W}) \quad (7)$$

for any unweighted and undirected graph F and

$$\|X_{\pi_n(G_n)} - X\|_{L^2} \to 0 \quad (8)$$

where $(\mathcal{W}_{\pi_n(G_n)},X_{\pi_n(G_n)})$ denotes the step function graphon signal induced by $(\pi_n(G_n),\pi_n(\mathbf{x}_n))$.

In short, we say that $\{(\mathbf{x}_n,G_n)\}$ converges to $(X,\mathcal{W})$ when: (i) G_n converges to $\mathcal{W}$ in the homomorphism density sense, and (ii) for some sequence of permutations $\{\pi_n\}$ of the labels of each G_n, the graphon signal induced by $\mathbf{x}_n$ converges to X in L^2. Notice that, similarly to how graphons model graph families, graphon signals can also be interpreted as models for network phenomena. A graph signal can be sampled from the graphon signal at the same locations u_i that a $\mathcal{W}$-random graph is sampled from the graphon. Since $\mathcal{W}$-random graphs converge to the graphon in probability, as long as the graphon signal is in L^2 any sequence of graph signals generated in this way can be shown to converge to the graphon signal (in the sense of Def. 1) with probability 1.

3.2. Graphon Fourier Transform

Because $\mathcal{W}$ is bounded, it defines a Hilbert-Schmidt integral operator $T_\mathcal{W} : L^2 \to L^2$ on $(\mathcal{W},X)$,

$$(T_\mathcal{W}X)(v) := \int_0^1 \mathcal{W}(u,v)X(u)du \quad (9)$$

which we call the *graphon shift operator* (WSO). Since $\mathcal{W}$ is symmetric, $T_{\mathcal{W}}$ is a self-adjoint operator that can be decomposed as $\mathcal{W}(u,v) \sim \sum_{i=0}^{\infty} \sigma_i \varphi_i(u)\varphi_i(v)$, where $\sigma_i \in [-1,1]$ and $\varphi_i : [0,1] \to \mathbb{R}$ are countable. Splitting between positive and negative eigenvalues, we can reorder the σ_i and φ_i as σ_j and φ_j with $j \in \mathbb{Z} \setminus \{0\}$ such that $1 \geq \sigma_1 \geq \sigma_2 \geq \ldots \geq 0 \geq \ldots \geq \sigma_{-2} \geq \sigma_{-1} \geq -1$. In this ordering, $\sigma_j \to 0$ as $|j| \to \infty$ due to compactness of $T_{\mathcal{W}}$; zero is the only possible point of accumulation, which means that all $\sigma_j \neq 0$ have finite multiplicities [22, Chapter 28, Theorem 3]. The eigenfunctions φ_j form the *graphon spectral basis*, which is a complete orthonormal basis of $L^2[0,1]$. The inner products $\langle X, \varphi_j \rangle$ thus yield a complete representation of $(\mathcal{W}, X)$ that we call the Graphon Fourier Transform.

Definition 2 (Graphon Fourier Transform)**.** *Consider the graphon signal $(\mathcal{W}, X)$ with $\mathcal{W} : [0,1]^2 \to [0,1]$ and $X : [0,1] \to \mathbb{R}$. Let $\{\sigma_j\}$ and $\{\varphi_j\}$ form the spectral decomposition of $T_{\mathcal{W}}$. Then, the Graphon Fourier Transform (WFT) of $(\mathcal{W}, X)$, denoted $\hat{X} = \text{WFT}\{(\mathcal{W}, X)\}$, is*

$$[\hat{X}]_j = \hat{X}(\sigma_j) = \langle X, \varphi_j \rangle = \int_0^1 X(u)\varphi_j(u)du \tag{10}$$

and the inverse Graphon Fourier Transform (iWFT) of $\hat{X}$ is $\text{iWFT}\{\hat{X}\} := \sum_j \hat{X}(\sigma_j)\varphi_j = X$.

Existence of the iWFT is guaranteed by orthonormality of the $\{\varphi_j\}$. Using the WFT, we can define graphon signals that are *bandlimited*.

Definition 3 (Bandlimited graphon signals)**.** *A graphon signal $(\mathcal{W}, X)$ is said to be c-bandlimited if, for some $c \in (0,1)$, the WFT coefficients $\hat{X}(\sigma_j)$ for which $j \in \{k \text{ s.t. } |\sigma_k| < c\}$ are zero. In particular, $\hat{X}$ has finite dimension for bandlimited graphon signals $(\mathcal{W}, X)$.*

It is important to point out that Def. 2, as well as the definition of a graphon signal, are not realizable in the way that graph signals and the GFT are. Unlike graphs, graphons are intangible objects, but their value lies in providing a complete and concise representation of multiple graphs belonging to the same "class" (eg. [16, 23]), or of *very large* graphs whose properties converge towards those of the graphon, as extensively demonstrated in [15]. Similarly to [18], where the notion of centrality was extended to graphons, the definitions in this section should be interpreted as merely a parallel to graph signal processing concepts, intended to provide a comparison with the case of discrete graphs and to enable the convergence analysis carried out in Section 4.

4. THE GFT CONVERGES TO THE WFT

The WFT only depends on the graphon through its eigenfunctions, whereas the GFT only depends on the graph's eigenvectors. Hence, for a convergent sequence of graph signals (Def. 1), convergence of the GFT to the WFT is equivalent to the graph's eigenvectors converging to the graphon's eigenfunctions. When graph signals converge to a graphon signal that is *bandlimited*, and when the underlying graphon is *non-derogatory*, this is indeed the case (cf. Lemma 2). This is stated in Theorem 1, but first we define non-derogatory graphons in Def. 4.

Definition 4 (Non-derogatory graphons)**.** *A graphon $\mathcal{W}$ is said to be non-derogatory if $\sigma_j \neq \sigma_k$ for $j \neq k$.*

Theorem 1 (GFT convergence)**.** *Let $\{(G_n, \mathbf{x}_n)\}$ be a sequence of graph signals converging to the c-bandlimited graphon signal $(\mathcal{W}, X)$, where $\mathcal{W}$ is non-derogatory. Then, there exists a sequence of permutations π_n such that $\text{GFT}\{(\pi_n(G_n), \pi_n(\mathbf{x}_n))\} \to \text{WFT}\{(\mathcal{W}, X)\}$ in the sense that $[\hat{\mathbf{x}}_n]_j/\sqrt{n} \to [\hat{X}]_j$ for all $j \in \mathbb{Z} \setminus \{0\}$ as $n \to \infty$.*

This theorem has three main takeaways. First, it allows drawing insights about the general spectral behavior of a graph signal without knowing the graph and/or the signal exactly, so long as the "class" to which the graph belongs (its generating graphon) and the generative model of the graph signal (the graphon signal) are known. Secondly, it links the Fourier transform on the "discrete" domain of graphs to the Fourier transform on the "continuous" domain of graphons. In the limit, this relationship resembles that of the discrete Fourier transform (DFT) with the continuous Fourier transform (CFT) for time signals. Finally, Theorem 1 also works in the opposite direction, meaning that if we sample a graph signal from the graphon signal, its GFT can be sampled from the WFT. This follows from the fact that sampled sequences of graph signals converge to the generating graphon signal with probability 1. We also point out that the requirement that the graphon be non-derogatory is not very restrictive because the space of non-derogatory graphons is dense in the space of graphons (cf. Prop. 4 in the extended version in [24]), and so, for every derogatory graphon, there exists a non-derogatory graphon arbitrarily close (in the appropriate metric) for which Theorem 1 holds.

We now proceed to proving Theorem 1.

Proof of Theorem 1. Although the GSOs $\mathbf{S}_n$ of the graphs G_n have a finite number of eigenvalues λ_j, we maintain the convention of associating the eigenvalue sign with its index and ordering the eigenvalues in decreasing order of absolute value. Hence, the indices j are now defined on some finite set $L \subseteq \mathbb{Z} \setminus \{0\}$. It will be useful to consider the continuous representation of the graph signals $(G_n, \mathbf{x}_n)$ provided by their induced graphon signals $(\mathcal{W}_{G_n}, X_{G_n})$ (eq. (5)). It is fundamental that $(\mathcal{W}_{G_n}, X_{G_n})$ retains the spectral properties of $(G_n, \mathbf{x}_n)$, which is guaranteed by Lemma 1 below.

Lemma 1. *Let $(\mathcal{W}_G, X_G)$ be the graphon signal induced by the n-node graph signal $(G, \mathbf{x})$ as defined in* (5)*. Then, for $j \in L$ we have*

$$\sigma_j(T_{\mathcal{W}_G}) = \frac{\lambda_j(\mathbf{S})}{n}$$
$$\varphi_j(T_{\mathcal{W}_G})(u) = [\mathbf{v}_j]_k \times \sqrt{n}\mathbb{I}\,(u \in I_k)$$
$$[\hat{X}_G]_j = \frac{[\hat{\mathbf{x}}]_j}{\sqrt{n}}$$

For $j \notin L$, $\sigma_j(T_{\mathcal{W}_G}) = [\hat{X}_G]_j = 0$ and $\varphi_j(T_{\mathcal{W}_G}) = \phi_j$, such that $\{\varphi_j(T_{\mathcal{W}_G})\} \cup \{\phi_j\}$ forms an orthonormal basis of $L^2([0,1])$.

Proof. Refer to the appendices of the extended version in [24].

It then suffices to show $\text{WFT}(\mathcal{W}_{\pi_n(G_n)}, \pi_n(X_{G_n})) \to \text{WFT}(\mathcal{W}, X)$, since $L \to \mathbb{Z}\setminus\{0\}$ as $n \to \infty$. We will leave the dependence on $\pi_n(G_n)$ implicit and write $\mathcal{W}_n = \mathcal{W}_{\pi_n(G_n)}$ and $X_n = \pi_n(X_{G_n})$. We proceed by using the following lemma, whose proof can be found in the appendices of the extended version [24]. Our result then follows from the fact that inner products are continuous in the product topology they induce.

Lemma 2. *Let $C = \{j \in \mathbb{Z} \setminus \{0\} \mid |\sigma_j(T_{\mathcal{W}})| \geq c\}$ be the set of indices of the non-vanishing eigenvalues and $\mathcal{E}$ denote the subspace spanned by the eigenfunctions $\{\varphi_j(T_{\mathcal{W}})\}_{j \notin C}$. Then, $\varphi_j(T_{\mathcal{W}_n}) \to \varphi_j(T_{\mathcal{W}})$ weakly for $j \in C$ and $\varphi_j(T_{\mathcal{W}_n}) \to \phi \in \mathcal{E}$ for $i \notin C$.*

Proof. Refer to the appendices of the extended version [24].

Start by the eigenvectors with indices in C. For any $\epsilon > 0$, it holds from Lemma 2 and from $X_n \to X$ in L^2, that there $\exists\, N_1, N_2$ such that

$$\|\varphi_j(T_{\mathcal{W}_n}) - \varphi_j(T_{\mathcal{W}})\|_{L^2} \leq \frac{\epsilon}{2\|X\|}, \text{for all } n > N_1$$
$$\text{and} \quad \|X_n - X\|_{L^2} \leq \frac{\epsilon}{2}, \text{for all } n > N_2$$

where $\|\varphi_j(T_{\mathcal{W}_n})\| \leq 1$ for all n and $j \in C$ because the graphon spectral basis is orthonormal. Since the sequence $\{X_n\}$ is convergent, it is bounded and $\|X\| < \infty$. Let $M = \max(N_1, N_2)$. Then, it holds that

$$|[\hat{X}_n]_j - [\hat{X}]_j| \leq \frac{\epsilon}{2}\|\varphi_j(T_{\mathcal{W}_n})\| + \|X\|\frac{\epsilon}{2\|X\|} \leq \epsilon$$

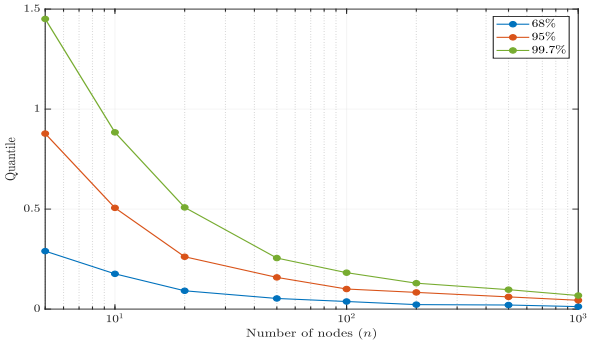

Fig. 2: Quantiles $(68\%, 95\%, 99.7\%)$ of the minimum normalized difference between GFTs of air pollution signals on G_1 and G_2 over 50 iterations for increasing n.

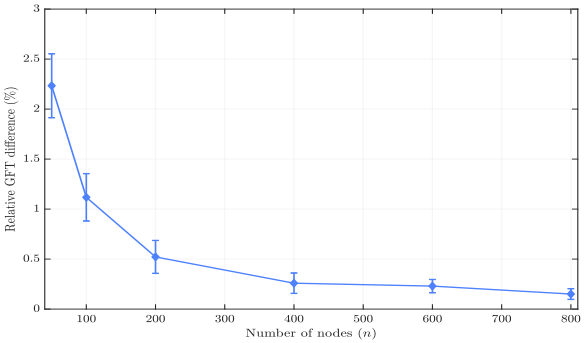

Fig. 3: Relative norm difference of the GFTs computed using the spectral basis of the n-node network and of the full 943-node network. Only the GFT coefficients associated with the first n eigenvalues were considered.

for all $n > M$, which we obtain by expressing $[\hat{X}_n]_j$ and $[\hat{X}]_j$ as inner products and applying the Cauchy-Schwarz inequality once. The coefficients with indices in C of the GFT of (G_n, X_n) therefore converge to those of the WFT of $(\mathcal{W}, X)$ as $n \to \infty$.

For $j \notin C$, the eigenfunctions $\varphi_j(T_{\mathcal{W}_n})$ may not converge to $\varphi_j(\mathcal{W})$, but they do converge to some function $\phi \in \mathcal{E}$. Since the graphon signal $(\mathcal{W}, X)$ is c-bandlimited, we have $\langle X, \varphi_j(T_{\mathcal{W}})\rangle = 0$ for all $j \notin C$. Using the same argument as above yields that the remaining GFT coefficients also converge to the WFT. □

5. NUMERICAL EXPERIMENTS

5.1. Convergence of air pollution signal spectra

We compare the spectral representations of air pollution signals measured on two distinct sensor networks of same size and drawn from the same random graph model, and show that the GFTs of the pollution signals on both networks converge as the number of nodes n increases. This problem can be seen as comparing the spectra of air pollution signals in two cities. Two air pollution sensor networks are drawn from a soft random geometric graph model [25] where, given nodes i, j and their (x, y) coordinates, the probability of the edge (i, j) is

$$p(i,j) \propto \exp\big(-\beta\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}\big)\,. \qquad (11)$$

Assuming that we only care about pollutant dispersion in the cross-wind direction $y \geq 0$, the x coordinate is fixed so that eq. (11) only depends on y_i and y_j. We also normalize y_i and y_j by setting $u_i = y_i/y_{\max}$ and $u_j = y_j/y_{\max}$ so that eq. (11) fits the expression of a graphon.

The pollution dispersion model is given by $S(y) \propto \exp(-y^2/2\sigma_y^2)$, where we have fixed the source of pollution at $y = 0$ and σ_y^2 is a cross-wind mixing term [26]. By normalizing y like before as $u = y/y_{\max}$, $S(u)$ can be interpreted as a signal on the graphon of eq. 11.

For a range of values of n, we sample two n-node graphs G_1 and G_2 at locations $\{u_i^1\}$ and $\{u_i^2\}$ drawn i.i.d from $U[0, 1]$. Then, we generate graph signals $(G_1, \mathbf{s}_1)$ and $(G_2, \mathbf{s}_2)$ by evaluating $[\mathbf{s}_1]_i = S(u_i^1)$ and $[\mathbf{s}_2]_i = S(u_i^2)$ for $1 \leq i \leq n$. Because of the way in which G_1 and G_2 and $\mathbf{s}_1$ and $\mathbf{s}_2$ are sampled, following Theorem 1 the GFTs of $(G_1, \mathbf{s}_1)$ and $(G_2, \mathbf{s}_2)$ should both converge to the WFT of S with probability 1. We verify this empirically by computing the GFTs of $(G_1, \mathbf{s}_1)$ and $(G_2, \mathbf{s}_2)$ and comparing their minimum norm difference.

The 68%, 95% and 99.7% quantile curves of the GFT norm difference for 50 realizations of this experiment are plotted in Figure 2. The minimum norm difference was computed by sorting and subtracting $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$, and the experiment was run for graphs of increasing size n. Figure 2 shows that the GFT converges as expected: across realizations, the minimum norm difference of the GFTs gets more and more concentrated around 0 as n increases.

5.2. GFT transferability in recommender systems

In this experiment, we use the MovieLens dataset [27] with 100,000 ratings from 943 users to 1,682 movies to analyze the behavior of the GFT on user similarity networks of increasing size. The networks are built by computing Pearson correlations between ratings that users have given to movies, in the fashion of [3]. The ratings vary between 1 and 5 and are seen as signals on the user network. We consider the movie "Toy Story", and use as the reference graph signal $\mathbf{x}$ the user ratings predicted for this movie on the full 943-user network and using the method in [3].

The experiment is as follows: we (i) generate a graph from a pool of users of size n, (ii) compute its spectral basis, (iii) sample the graph signal $\mathbf{x}_n$ from $\mathbf{x}$ and (iv) compare the GFT of $\mathbf{x}_n$ in the n-node user network, $[\hat{\mathbf{x}}_n]_{i=1}^n$, with the GFT coefficients $[\hat{\mathbf{x}}]_{i=1}^n$ of the complete signal $\mathbf{x}$ associated with the n largest eigenvalues of the full 943-user network. Although these are networks built from real data to which we cannot attribute a generating graphon, our goal here is to illustrate how our results can be implicitly observed and used to leverage transferability even in graphs that are not related by a common statistical model, but that are "similar" in some empirical sense.

The average relative norm difference between the GFTs and its standard deviation are shown in Figure 3 for 10 realizations of networks with n users each. We see that the GFT difference consistently decreases with n. Perhaps more importantly, notice that this difference is as low as 0.5% on average for networks as small as $n = 200$. In other words, the spectral decomposition of the signal corresponding to "Toy Story" ratings in a network with less than a quarter of the users of the full network is already an accurate enough representation of the complete spectra of this signal.

6. CONCLUSIONS

In this work, we have introduced a Fourier transform for graphon signals and shown that, for convergent sequences of graph signals, the GFT converges to the WFT. This result is empirically demonstrated in the numerical experiments of Section 5. The definition of graphon signals and of the WFT, together with Theorem 1, lay the ground for a graphon signal processing framework on which signals can be analyzed and systems designed in a *centralized* and *transferable* way, to then be applied to sequences of graphs or multiple instances of the same type of graph.

7. REFERENCES

[1] M. Eisen and A. Ribeiro, "Optimal wireless resource allocation with random edge graph neural networks," *arXiv:1909.01865 [eess.SP]*, 2019.

[2] R. Levie, E. Isufi, and G. Leus, Kutyniok, "On the transferability of spectral graph filters," *arXiv:1901.10524 [cs.LG]*, 29 Jan. 2019.

[3] W. Huang, A. G. Marques, and A. Ribeiro, "Rating prediction via graph signal processing," *IEEE Trans. Signal Process.*, vol. 66, no. 19, pp. 5066 – 5081, Oct. 2018.

[4] L. Ruiz, F. Gama, G. Marques, and A. Ribeiro, "Invariance-preserving localized activation functions for graph neural networks," *arXiv:1903.12575 [eess.SP]*, 29 March 2019.

[5] D. I Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, May 2013.

[6] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr. 2013.

[7] A. Sandyhaila and J. M. F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3042–3054, June 2014.

[8] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[9] S. Segarra, A. G. Marques, and A. Ribeiro, "Optimal graph-filter design and applications to distributed linear network operators," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4117–4131, Aug. 2017.

[10] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Graph filters," in *IEEE Int. Conf. Acoust., Speech and Signal Process.* IEEE, 2013, pp. 6163–6166.

[11] G. J. Wills, "NicheWorks – interactive visualization of very large graphs," *Journal of computational and Graphical Statistics*, vol. 8, no. 2, pp. 190–212, 1999.

[12] L. Rui, H. Nejati, and N. Cheung, "Dimensionality reduction of brain imaging data using graph signal processing," in *2016 IEEE Int. Conf. on Image Process.*, Phoenix, AZ, 25-28 Sep. 2016, pp. 1329–1333, IEEE.

[13] R. B. Morgan, "Computing interior eigenvalues of large matrices," *Linear Algebra and its Applications*, vol. 154, pp. 289–309, 1991.

[14] C. C. Paige, *The computation of eigenvalues and eigenvectors of very large sparse matrices*, Ph.D. thesis, University of London, 1971.

[15] L. Lovász, *Large networks and graph limits*, vol. 60, American Mathematical Society, 2012.

[16] P. J. Wolfe and S. C. Olhede, "Nonparametric graphon estimation," *arXiv:1309.5936 [math.ST]*, 2013.

[17] E. M. Airoldi, T. B. Costa, and S. H. Chan, "Stochastic blockmodel approximation of a graphon: Theory and consistent estimation," in *Advances in Neural Information Processing Systems*, 2013, pp. 692–700.

[18] M. Avella-Medina, F. Parise, M. Schaub, and S. Segarra, "Centrality measures for graphons: Accounting for uncertainty in networks," *IEEE Transactions on Network Science and Engineering*, 2018.

[19] P. Diao, D. Guillot, A. Khare, and B. Rajaratnam, "Model-free consistency of graph partitioning," *arXiv:1608.03860 [math.CO]*, 2016.

[20] F. Parise and A. Ozdaglar, "Graphon games," in *Proceedings of the 2019 ACM Conference on Economics and Computation*. ACM, 2019, pp. 457–458.

[21] M. W. Morency and G. Leus, "Signal processing on kernel-based random graphs," in *Eur. Signal Process. Conf.* IEEE, 2017, pp. 365–369.

[22] P. D. Lax, *Functional Analysis*, Wiley, 2002.

[23] M. T. Schaub, S. Segarra, and H. Wai, "Spectral partitioning of time-varying networks with unobserved edges," in *2019 IEEE Int. Conf. Acoust., Speech and Signal Process.*, Brighton, UK, 12-17 May 2019, IEEE.

[24] L. Ruiz, L. F. O. Chamon, and A. Ribeiro, "The Graphon Fourier Transform (extended)," `https://www.seas.upenn.edu/~rubruiz/graphons-icassp19.pdf`, 2019.

[25] M. D. Penrose, "Connectivity of soft random geometric graphs," *The Annals of Applied Probability*, vol. 6, no. 2, pp. 986–1028, 2016.

[26] S. P. et al Arya, *Air pollution meteorology and dispersion*, vol. 6, Oxford University Press New York, 1999.

[27] F. M. Harper and J. A. Konstan, "The movielens datasets: History and context," *ACM Transactions on Interactive Intelligent Systems (TiiS) - Regular Articles and Special issue on New Directions in Eye Gaze for Interactive Intelligent Systems*, vol. 5, pp. 1–19, Jan. 2016.