RANDOM WALKS ON DENSE GRAPHS AND GRAPHONS

JULIEN PETIT[1,2], RENAUD LAMBIOTTE[3], AND TIMOTEO CARLETTI[2]

May 20, 2020

ABSTRACT. Graph-limit theory focuses on the convergence of sequences of increasingly large graphs, providing a framework for the study of dynamical systems on massive graphs, where classical methods would become computationally intractable. Through an approximation procedure, the standard ordinary differential equations are replaced by nonlocal evolution equations on the unit interval. In this work, we adopt this methodology to prove the validity of the continuum limit of random walks, a largely studied model for diffusion on graphs. We focus on two classes of processes on dense weighted graphs, in discrete and in continuous time, whose dynamics are encoded in the transition matrix of the associated Markov chain, or in the random-walk Laplacian. We further show that previous works on the discrete heat equation, associated to the combinatorial Laplacian, fall within the scope of our approach. Finally, we characterize the relaxation time of the process in the continuum limit.

Key words. random walk, dense graph, graphon, continuum limit
AMS subject classification. 05C81, 34G10, 45K05, 47D06

1. INTRODUCTION

Graphs are everywhere, they appear for example in physics, engineering, biology, climat science, expectedly every time discrete entities interact through links of some nature [1]. They offer a conceptually simple but general enough modeling approach to real-life networks of various degrees of complexity. But as very large graphs have become commonplace in scientific research and real-world applications – online social networks or the brain are relevant examples – a range of graph-theoretical methods, algorithms and computational problems on graphs face scalability issues. In a recent series of works on the continuum limit of graphs, graphons have emerged as an appropriate limit object, defined when the number of nodes goes to infinity [25, 6, 3, 24]. They offer a both elegant and efficient workaround allowing for the analysis of massive graphs, while simultaneously providing a non-parametric network generation method that reaches far beyond the classical stochastic block model [18]. The versatility of graphons and their practical interest is revealed in their present-day use in many diverse domains, e.g. network identification [14], power network dynamics [21], epidemics spreading [35], reaction-diffusion [20], synchronization of oscillators [27, 29] and game theory [31]. However, the majority of these examples focus on a specific type of diffusion, namely

[1] DEPARTMENT OF MATHEMATICS, ROYAL MILITARY ACADEMY, BRUSSELS, BELGIUM
[2] DEPARTMENT OF MATHEMATICS AND NAMUR INSTITUTE FOR COMPLEX SYSTEMS (NAXYS), UNIVERSITY OF NAMUR, BELGIUM
[3] MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD, UK

E-mail addresses: `julien.petit@unamur.be`, `renaud.lambiotte@maths.ox.ac.uk`, `timoteo.carletti@unamur.be`.

Fickean diffusion, while alternative random-walk processes are usually preferred to build network algorithms for node ranking [7] or community detection [10]. This constitutes an important limitation that we address in this paper.

Diffusion on finite graphs is an extensive topic of research, which is relevant both from a theoretical and an applied perspective, and is often modeled as some variant of random walk process. Random walks on graphs are useful in many ways. They can for example identify clusters of well-connected nodes, also known as communities [34, 10], or measure the relative importance or centrality of the nodes in a networked system [7, 23, 22], and as mentioned are a paradigm for various diffusive and spreading processes on graphs [26]. There are overall three dominant classes of random walks. The first one is the discrete-time walk, in which case the walker performs a new jump at every discrete time step. The destination node of the jump is uniformly chosen among the neighbors in the graph structure. Secondly there is the continuous-time, node-centric variant. The difference resides in that the jumps take place at any point in time, as dictated by a continuous random variable governing the resting time of the walker on a node. Finally, the third class corresponds to the continuous-time edge-centric walk, which can interestingly be viewed as the discrete version of the heat equation. These are two different forms of normal diffusion equations that can be derived from Fick's law, that essentially implies a flux from regions of high to low concentration, across a concentration gradient. The difference between the node-centric and edge-centric processes is clear when observing the matrices controlling their dynamics, the random-walk and the combinatorial Laplacian respectively [32].

In this work, we first revisit existing results for the continuum limit of the discrete heat equation and some nonlinear variants. This limit was the subject of a series of papers recently [27, 28, 29], but a random walk interpretation is useful and was still lacking. We then concentrate on the continuum limit of the node-centric case, hence considering the limit of the random-walk Laplacian operator. In general, for non-regular graphs, this operator differs from the combinatorial Laplacian, which is often preferred in algorithmic implementations such as spectral clustering [36], because it properly accounts for the heterogeneous degree distributions observed in real-life networks. The random-walk operator in this work shouldn't be confused with a another operator common in the machine learning community, also called random-walk Laplacian, which has an established convergence to the Laplace-Beltrami operator [17, 2].

Our approach is based on graph-limit theory [24], which does not rely on the assumption that the data generating the graphs is sampled from a distribution on a manifold [16, 33]. Our main contribution is to prove the convergence of the space-discrete problem to a continuous problem in some appropriate setting. The problem on the continuum then falls in the realm of nonlocal evolution equations, as it is a volume-constrained diffusion problem [9]. Its analysis is limited to some consequences of spectral theory applied to our operators.

Importantly, graph-limit theory defines a framework for the convergence of graphs of increasing size, but it may as well be seen as a possibly random graph-generating method. From that perspective, our work demonstrates that one may analyze the continuum model, to draw valid conclusions regarding the dynamics on the graphs generated by that model. As such, we answer positively the question of transferability, showing that we can avoid the cost of repeatedly storing each graph, designing,

visualizing or analyzing it, or computing the associated spectral properties, a computationally demanding task for large graphs [30].

The paper is organized as follows. Section 2 contains the basics about graph-limit theory, and introduces graphons as the limit objects of dense graph sequences. A short presentation of the main random walk models opens section 3. Then follows a random walk interpretation of the continuum limit of the heat equation on graphs, before we focus on our main concern, the continuous-time node-centric walk. Well-posedness of the continuum problem is the subject of section 4. The main convergence results are presented in section 5. These results apply to dense graphs, and follow from a semigroup approach. We distinguish between different scenarios : first the discrete problem on graphs is sampled from the continuum version, and then the other way around. We then proceed with an analysis of the relaxation-time of the process based on spectral theory in section 6. In section 7 we revisit the discrete-time problem, before the conclusion of section 8.

2. Preliminaries

We first set the notation and recall definitions for various graph-related concepts, following closely [37]. For the sake of self-consistency we then introduce key notions about graph-limit theory [24].

2.1. **Graphs.** Let a $G = (V, E)$ be a graph where V is a finite set of vertices (or nodes), and $E \subset V \times V$ is the set of edges. We consider a symmetric adjacency relation $\sim$ such that the graph is undirected. Each edge may be attributed a weight, making the otherwise unweighted graph into a weighted one. Let $|V|$ and $|E|$ respectively denote the number of vertices and edges. The density ρ of the graph is the fraction of edges that are actually present, compared with the maximum possible number of connections : $\rho = \frac{|E|}{\binom{|V|}{2}}$. When it makes sense to take the limit $|V| \to \infty$, one says the graph is dense if $|E| = O(|V|^2)$, and sparse otherwise.

The number of neighbors of node v is denoted by $\deg(v)$. In weighted graphs, $\mathrm{str}(v)$ stands for the weighted degree, or strength, of node v, namely the sum of the weights of all edges attached to v. For the sake of simplicity, the notation $\mathrm{str}(v)$ will henceforth also apply to unweighted graphs and will refer to the degree, thereby identifying unweighted graphs with graphs with binary weights, either 0 or 1.

A path between two nodes v, w is an ordered sequence of nodes $[v_1, \ldots, v_n]$ such that $v = v_1$, $w = v_n$ and $v_i \sim v_{i+1}$, $i = 1, \ldots, n-1$. A graph is connected if every pair of nodes is linked by a path. Let $\mathbb{M}_n$ be the space $n \times n$ matrices. The adjacency matrix $A \in \mathbb{M}_n$ of a graph G with n vertices is the square matrix where A_{ij} is the weight of the edge between nodes with labels i and j and zero if no such edge exists. Unweighted graphs have binary adjacency matrix.

2.2. **Graphons.** Recent research [25, 6, 3, 24] provides a theoretical framework to study convergence of dense graphs sequences. As a starting point, the so-called cut (or rectangular) metric allows to define the notion of Cauchy sequence of graphs of increasing number of nodes. Their limit object, called graphon, is a symmetric Lebesgue-measurable function $W : [0, 1]^2 \to [0, 1]$[1]. Therefore, the space of

---

[1]Note that this choice of domain and range is somehow restrictive by comparison with other works where graphons may be unbounded and only feature some integrability property. However, we will work with the standard definition because it achieves the desired degree of generality.

graphons is essentially the completion of the set of finite graphs seen as step functions (see Section 2.3), endowed with the so-called cut metric[2] which we introduce hereafter in its graphon version [24, 27, 15]. The cut norm for graphons is given by

$$\|W\|_\square = \sup_{S,T\subset\mathfrak{M}[0,1]} \iint_{S\times T} W(x,y)dxdy,$$

where the supremum is over measurable subsets of $[0,1]$. The notation $\|W\|_p$ refers to the usual L^p norm of function defined on $[0,1]^2$, for $1\leq p\leq\infty$. The following inequalities are immediate consequences of this definition, and of the inclusion theorem of L^p spaces on finite measure spaces :

(1) $$\|W\|_\square \leq \|W\|_1 \leq \|W\|_2 \leq \|W\|_\infty \leq 1.$$

Graphons are unique up to a composition with an invertible measure preserving mapping $\phi : [0,1] \to [0,1]$, which amounts to invariance of the limit graphon with respect to a relabeling of the nodes of the graphs. The graphons W^ϕ defined by $W^\phi(x,y) = W(\phi(x),\phi(y))$ and W are in the same equivalence class. The cut metric $\delta_\square$ between two graphons U and W is therefore defined by

(2) $$\delta_\square(U,W) = \inf_{\phi\in\mathfrak{L}} \|U^\phi - W\|_\square$$

where $\mathfrak{L}$ is the space of the Lebesgue measurable bijections on the unit interval. The definition is similar for the $\delta_p(\cdot,\cdot)$ metrics based on the L^p norms, $1\leq p\leq\infty$. Since two different graphons U,W can satisfy $\delta_\square(U,W) = 0$, strictly speaking $\delta_\square$ is a metric only when we identify such graphons U and W [6]. Let us denote by $\mathbb{W}$ the space of graphons after this identification.

It holds that the metric space $(\mathbb{W}, \delta_\square)$ is compact, namely sequences of graphons posses at least one convergent subsequence in the cut metric. Unless explicitly mentioned, in this work we assume convergence of graphons in the L^2 norm topology. Hence by completeness, Cauchy sequences in $(\mathbb{W}, \|\cdot,\cdot\|_2)$ converge in the L^2 metric, and thus also in the δ_2 and $\delta_\square$ metrics, the limit being the same.

Many attributes of graphs have natural counterparts in the realm of graphons. A prominent example is the notion of strength, which plays a key role in this paper. For a given graphon we let $k(x) := \int_0^1 W(x,y)dy$ denote the (generalized) degree function. Since in this work graphons are bounded function $W : [0,1]^2 \to [0,1]$, the degree function is bounded, $0 \leq k(x) \leq 1$ for all $x \in [0,1]$.

2.3. **Graphs as step graphons and graphs from graphon models.** The connection between graphs and graphons is a two-way street. First, graphs can be mapped to the graphon space through a step function representation of their adjacency matrix. Let $\mathcal{P} = \{P_1,\dots,P_n\}$ be a uniform partition of $[0,1]$, where $P_i = \left[\frac{i-1}{n},\frac{i}{n}\right]$ for $i = 1,\dots,n-1$, and $P_n = \left[\frac{n-1}{n},1\right]$. Then let $\eta : \mathbb{M}_n \to \mathbb{W}$ be a mapping such that

(3) $$\eta(G)(x,y) = \sum_{i=1}^n \sum_{j=1}^n A_{ij}\chi_{P_i}(x)\chi_{P_j}(y),$$

where χ_S is the indicator function of set S and A the adjacency matrix of graph G. The mapping thus defines the step (or empirical) graphon $\eta(G)$ associated to

---

[2]There is a different though equivalent notion of convergence for dense graph sequences. It is called subgraph convergence, and is defined via associated sequences of induced subgraph densities [6].

G. Similarly, η maps vectors $\mathbf{u} = (u_1, \ldots, u_n)$ to piecewise constant functions on $[0, 1]$, so that

$$(4) \qquad \eta(\mathbf{u})(x) = \sum_{i=1}^{n} u_i \chi_{P_i}(x).$$

On the other hand, graphons can be considered as deterministic or (exchangeable [12]) random graph models, but in this work we adopt and present the deterministic setting. Let $W \in \mathbb{W}$ be a graphon and let the integer n denote the desired number of nodes in the graph. Then W generates a dense graph by assigning weights to the edges, which can be done in two ways. In a first approach, the weight A_{ij} of the edge between two nodes i and j equals the mean value of W on the corresponding cell of the partition of the unit square:

$$(5) \qquad A_{ij} = n^2 \int_{P_i} \int_{P_j} W(x, y) dx dy, \quad i, j = 1, \ldots, n.$$

This results in the so-called quotient graph $W/\mathcal{P}$. One can prove that there is almost everywhere point-wise convergence of the associated step graphon $\eta(W/\mathcal{P})$ to W ([6], lemma 3.2).

A second approach to generate a graph from a given graphon $W \in \mathbb{W}$, is to define

$$(6) \qquad A_{ij} = W\left(\frac{i}{n}, \frac{j}{n}\right), \quad i, j = 1, \ldots, n,$$

in a way that is reminiscent of W-random graphs [25]. Let us denote $W_{[n]}$ the corresponding graph. Observe that $\eta\left(W_{[n]}\right) \to W$ point-wise at every point of continuity of W [27].

2.4. Graphons as kernels of operators.
Every graphon $W \in \mathbb{W}$ can be considered as a kernel, allowing to formally define an integral operator $\mathcal{W}$ on functional spaces on $[0, 1]$ through

$$(7) \qquad \mathcal{W}f(x) = \int_0^1 W(x, y) f(y) dy.$$

The composition (product) of two such operators is given by

$$(8) \qquad \mathcal{U}\mathcal{W}f(x) = \int_0^1 (U \circ W)(x, y) f(y) dy,$$

where $\circ$ is the operator product between the graphon kernels, defined by

$$(9) \qquad (U \circ W)(x, y) = \int_0^1 U(x, z) W(z, y) dz, \quad \forall x, y \in [0, 1].$$

Observe that in general, $U \circ W$ is not a symmetric function. We denote $W^{\circ n}$ the operator product of the kernel, as opposed to the point-wise product $W^n(x, y) = (W(x, y))^n$, which is associated to the operator $\mathcal{W}^n$. It follows from eq. (9) that $W^{\circ n}(x, y) = \int W(x, z_1) W(z_1, z_2) \ldots W(z_{n-1}, y) dz_1 dz_2 \ldots dz_{n-1}$.

3. Random walks and their continuum limit

The aim of this section is twofold. Firstly, we introduce the three main random walk models. Secondly, we give a random walk perspective on the continuum limit of the discrete heat equation, namely the edge-centric walk. We then formally derive the continuum limit of the so-called node-centric walk. Our approach implies edges are directed, because they are associated with possible moves of the walker, with an origin and a destination. Therefore, the symmetry of the adjacency matrix indicates there exists a reciprocal to each edge and that both have the same weight. Further, we may assume that the graph is connected, otherwise the random walk is considered independently on each connected component, that is, each connected subgraph that is connected to no other additional node of the original graph.

3.1. **Random walks in discrete and continuous time.** In discrete-time, we introduce a random walk on a connected graph as a Markov chain where V is the state-space and the transition probability from node v_i to v_j is encoded in the matrix

$$(10) \qquad T_{ij} = \begin{cases} 1/\operatorname{str}(v_i) & \text{if } v_i \sim v_j, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathbf{p}(\ell) = (p_1(\ell), \dots, p_n(\ell))$ be the row vector of residence probabilities on the nodes, that is, $p_i(\ell)$ is the probability that the walker is located on node number i after ℓ steps. Then

$$(11) \qquad \mathbf{p}(\ell + 1) = \mathbf{p}(\ell)T,$$

where $T = D^{-1}A$. Here D denotes the diagonal matrix of the strengths, or degrees in unweighted graphs. It follows from eq. (11) that for any $\ell \in \mathbb{N}$, $\mathbf{p}(\ell) = \mathbf{p}(0)T^\ell$.

In the continuous-time node-centric variant, when the walker arrives on node v, a probability density function $\psi_v(t)$ determines the waiting-time until the next jump, in which case a destination node is selected uniformly among the neighboring ones. We limit ourselves to Poissonian walks, for which the waiting-time follows a memoryless exponential distribution $\psi_v(t) = \mu_v \exp(-\mu_v t)$ with rate μ_v ($t \geq 0$). The master equation for $u_i(t)$, the probability to find the walker on node i at time t, reads

$$(12) \qquad \dot{u}_i(t) = \sum_{j=1}^{N} \mu_j u_j \frac{1}{\operatorname{str}(v_j)} A_{ji} - \mu_i u_i, \qquad i = 1, \dots, n.$$

Assume that in eq. (12) the rate μ_j on the nodes is the same for all nodes, $\mu_j = \kappa > 0$ for all j. Then κ sets the timescale, and after a scaling of time, $t \mapsto \kappa t$, under matrix form the master equation eq. (12) rewrites

$$(13) \qquad \dot{\mathbf{u}} = \mathbf{u}(D^{-1}A - I),$$

where $\mathbf{u}(t) = (u_1(t), \dots, u_n(t))$ is a row vector. The matrix $L^{rw} = D^{-1}A - I$ is the random walk Laplacian. Moreover, it is easy to show that the discrete-time walk and the continuous-time version share the same asymptotic state, and that it is proportional to $(\operatorname{str}(v_1), \dots, \operatorname{str}(v_n))$.

In the edge-centric variant[3], the rate of the exponential distribution is proportional to the degree of the node, $\mu_j = \kappa \operatorname{str}(v_j)$, allowing a constant rate of jump across all edges of the graph. Hence in matrix form, Equation (12) rewrites

$$(14) \qquad \dot{\mathbf{u}} = \kappa \mathbf{u}(A - D)$$

Here, $L = A - D$ is called the combinatorial Laplacian of the graph. This model exhibits a homogeneous asymptotic state. Observe that the number of jumps is not trajectory-independent, as is the case in both the discrete- and continuous-time node-centric walks.

3.2. **Formal derivation of the continuum limit.** Let us first take a closer look at the edge-centric walk, and assume for simplicity an unweighted graph. If $\kappa > 0$, then $\kappa \deg(v_j) \to \infty$ if $\deg(v_j) \to \infty$, which will happen for some if not all nodes of a dense graph. The walker would perform jumps at an infinite rate, which is physically unrealistic. Normalizing the rate of the process according to the number of vertices avoids this situation. If κ becomes dependent on n, say $\kappa_n = \frac{1}{n}$, the resulting rate in each node remains bounded, $\kappa_n \deg(v_j) \leq 1$ for all j independently of the number of nodes. This explains the normalization that was required to justify the continuum limit of eq. (14) in [27].

In contrast with the edge-centric model, no normalization of the rate parameter κ of the node-centric walk is needed when the number of nodes grows to infinity, since the rate does not depend on the structure of graph. The continuum limit therefore directly applies to the unmodified discrete model. For a formal derivation in this case, consider again the vector $\mathbf{u}(t)$ satisfying eq. (13) and the uniform partition $\mathcal{P} = \{P_1, \ldots, P_n\}$ of $[0, 1]$, with $u(\cdot, t) := \eta(\mathbf{u}(t))$ an associated step function on the interval. Let k_η denote the generalized degree function of the step graphon $\eta(G)$. Observe that this degree function is actually the normalized strength (or also degree, when the graph is unweighted) of the nodes in G:

$$(15) \qquad \operatorname{str}(v_i) = n \sum_{j=1}^{n} \int_{P_j} A_{ij} dy = n \sum_{j=1}^{n} \int_{P_j} \eta(G)(x, y) dy = n k_\eta(x)$$

for all $x \in P_i$. It follows that

$$\sum_{j=1}^{n} \frac{A_{ij}}{\operatorname{str}(v_j)} u_j(t) = \sum_{j=1}^{n} n \int_{P_j} \frac{A_{ij}}{\operatorname{str}(v_j)} u(y, t) dy$$

$$(16) \qquad = \sum_{j=1}^{n} n \int_{P_j} \frac{A_{ij}}{n k_\eta(y)} u(y, t) dy = \int_0^1 \frac{\eta(G)(x, y)}{k_\eta(y)} u(y, t) dy,$$

for every $x \in P_i$. Hence, the node-centric walk on the graph has an equivalent continuum domain formulation

$$(17) \qquad \frac{\partial}{\partial t} u(x, t) = \int_0^1 \frac{\eta(G)(x, y)}{k_\eta(y)} u(x, t) dy - u(x, t).$$

---

[3] The fact that this walk can be formulated in terms of edges dynamics, where the walker passively follows the activations of the edges explains the alternative designation of "fluid model" [26]. It is the graph version of the heat equation on a continuum.

The goal of this work is to prove convergence in the appropriate norm of the solution of eq. (17) to the solution of the evolution equation on the continuum

$$(18) \qquad \frac{\partial}{\partial t} w(x,t) = \int_0^1 \frac{W(x,y)}{k(y)} w(y,t) dy - w(x,t),$$

where W is the limit graphon of $\eta(G)$ in the L^2 metric.

Observe that similarly, a discrete equation of the form eq. (13) is obtained starting from eq. (18), when the graph is $W/\mathcal{P}$ or $W_{[n]}$.

4. Well-posedness of the continuum initial value problem

Before we prove the above-mentioned convergence, let us determine whether the up-to-now formal eq. (18), together with initial condition $w(x,0) = g(x)$, defines a well-posed initial-value problem (IVP).

4.1. **Connectedness of the graphon.**

Care will be taken first regarding how connectedness in the graph translates to graphons, and how it affects the integrability of W/k. The following definition follows from [19, 24].

Definition 4.1. A graphon W is connected if $\int_{S\times([0,1]\setminus S)} W(x,y) dx dy > 0$ for every $S \in \mathfrak{M}[0,1]$ with lebesgue measure $\mu(S) \in (0,1)$.

Notice at this point that the connectedness (or lack thereof) of the graphs G_n of the sequence does not imply that of their limit [19]. Indeed, one could always make all the (otherwise disconnected) graphs of the sequence connected by a adding each time a node connected to all other nodes. This would leave the limit unchanged. And conversely, disconnecting one node in each connected graph of the sequence would not change the limit either. Also note that if a graphon W is (dis)connected, then so are all the kernels in the same equivalence class ([19], theorem 1.16). Let us now look into the implications of connectedness of the graphon on the positiveness of the degree function and hence on the definition of the random walk Laplacian operator.

Proposition 4.2. *Let W be a connected graphon, then $k > 0$ μ-almost everywhere (a.e.).*

Proof. Let $N_x = \{y \in [0,1] : W(x,y) > 0\}$ denote the neighborhood of $x \in [0,1]$ in W. Since W is connected, $\mu(N_x) > 0$ for μ-almost every x ([19], lemma 5.1) and therefore,

$$(19) \qquad k(x) = \int_{N_x} W(x,y) dy > 0 \ \text{ for } \ \mu\text{-a.e. } x.$$

$\square$

Remark 4.3. The connectedness of the graphon does not imply however that the degree function is bounded away from zero, namely that there exists a constant c such that $0 < c \le k$ on $[0,1]$. Take for instance $W(x,y) = x^m y^m$ with $m > 0$, for which $k(x) = x^m/(m+1)$.

That k can be arbitrarily small influences the integrability of the kernel $K(x,y) := \frac{W(x,y)}{k(y)}$ in eq. (18), as discussed in the following remark.

Remark 4.4. The connectedness of the graphon does not imply that the integral kernel $K(x, y)$ is in $L^p[0, 1]^2$ for $p > 1$. Consider for example the binary graphon $W = \chi_{x^\alpha + y^\alpha \leq 1}$ for $\alpha > 0$, where the subscript $x^\alpha + y^\alpha \leq 1$ is short for the set of couples $(x, y) \in [0, 1]^2$ such that the inequality is satisfied. By a direct integration, $k(x) = (1 - x^\alpha)^{\frac{1}{\alpha}}$. The integral

$$(20) \qquad \|K\|_p^p = \iint_{x^\alpha + y^\alpha \leq 1} (1 - y^\alpha)^{-\frac{p}{\alpha}} dx dy = \int_0^1 (1 - y^\alpha)^{\frac{1-p}{\alpha}} dy$$

is finite if and only if $p < 1 + \alpha$. Hence, K is in $L^2[0, 1]$ only if $\alpha > 1$, and in particular, the kernel K of the threshold graphon [11] obtained with $\alpha = 1$ is not square-integrable. However, using Fubini-Tonelli it is easy to show that $\|K\|_1 = 1$ for all connected graphons, such that K is always in $L^1[0, 1]^2$.

Based on the preceding remark, in order to ensure that the kernel is square integrable, we will make the following assumption :

Assumption 4.5. There exists a constant c such that $0 < c \leq k$ on $[0, 1]$.

If W is bounded away from zero, so is k, but graphons with localized support may still fulfill the assumption, as shown by Figure 1.

FIGURE 1. The stripe graphon $W(x, y) = \chi_{|x-y| \leq 1/4}(x, y)$ (left panel) and its degree function (right panel). Observe that the support of W is localized on a subset of the square, but k is bounded away from zero.

4.2. The IVP with functions in $L^2[0, 1]$.
Resting on the operator in the right-hand side of eq. (18), we come to the following definition.

Definition 4.6. Let $W \in \mathbb{W}$ be a connected graphon that verifies assumption 4.5. The random-walk Laplacian operator $\mathcal{L}^{rw} : L^2[0, 1] \to L^2[0, 1]$ is defined by

$$(21) \qquad \mathcal{L}^{rw} f(x) = \int_0^1 \frac{W(x, y)}{k(y)} f(y) dy - f(x).$$

By definition of W and assumption 4.5, $K(x, y) = W(x, y)/k(y)$ is a Hilbert-Schmidt kernel and $\mathcal{K} : L^2[0, 1] \to L^2[0, 1]$ defined by

$$(22) \qquad \mathcal{K}f(x) = \int_0^1 \frac{W(x, y)}{k(y)} f(y) dy, \quad \forall x \in [0, 1] \text{ and } f \in L^2[0, 1]$$

is a compact Hilbert-Schmidt operator. Following definition 4.6, the continuum IVP has the form

$$(23\text{a}) \qquad \frac{\partial}{\partial t} w(x,t) = \mathcal{L}^{rw} w(x,t),$$

$$(23\text{b}) \qquad w(x,0) = g(x) \in L^2[0,1].$$

Theorem 4.7. *Let $W \in \mathbb{W}$ be connected and satisfying assumption 4.5. Then there exists a unique classical solution to the initial-value problem eq. (23).*

Proof. The operator $\mathcal{K}$ is linear, and continuous hence bounded. It follows that $\mathcal{L}^{rw}$ is linear and bounded. Hence it is closed. Therefore, $\mathcal{L}^{rw}$ is the infinitesimal generator of the (uniformly and thus) strongly continuous semigroup

$$(24) \qquad \mathcal{T}^{rw}(t) = e^{\mathcal{L}^{rw} t} := \sum_{\ell=0}^{\infty} \frac{t^{\ell} (\mathcal{L}^{rw})^{\ell}}{\ell!}.$$

Proposition 6.2 in [13] allows to conclude. $\qquad \square$

Remark 4.8. (Classical solution) By definition of classical solution of the abstract Cauchy problem eq. (23), the orbit maps $t \in \mathbb{R}^+ \mapsto w(x,t) \in L^2[0,1]$ are continuously differentiable.

Remark 4.9. The asymptotic steady state w_{∞} of eq. (23) follows from $\mathcal{L}^{rw} w_{\infty} = 0$ and is proportional to the degree, $w_{\infty} \propto k$.

4.3. **Positivity.** The continuum IVP eq. (23) would loose physical relevance if its solution were to loose the positivity of the initial condition, $w(\cdot,0) \geq 0$. Before we proceed to a proof of positivity, let us first introduce a notation. For $g \in L^{\infty}[0,1]$, and $1 \leq p \leq \infty$, let $\mathcal{M}_g : L^p[0,1] \to L^p[0,1]$ denote the multiplication operator defined by $\mathcal{M}_g f(x) = g(x) f(x)$.

Proposition 4.10. *Let W be a connected graphon satisfying assumption 4.5 and let $w(\cdot,0) = g \geq 0$ be the initial condition of IVP eq. (23). Then the classical solution $w(x,t)$ of the IVP satisfies $w(\cdot,t) \geq 0$ for all $t \geq 0$.*

Proof. Let us define $\widetilde{\mathcal{L}} = \mathcal{M}_{1/k} \mathcal{L}^{rw} \mathcal{M}_k$, yielding by a direct calculation

$$(25) \qquad \widetilde{\mathcal{L}} f(x,t) = \frac{1}{k(x)} \int_0^1 W(x,y) f(y,t) dy - f(x,t), \quad \forall f \in L^2[0,1].$$

Further let $u = \mathcal{M}_{1/k} w$ with $w(x,t)$ the solution of eq. (23) such that

$$\frac{\partial}{\partial t} u = \mathcal{M}_{1/k} \frac{\partial}{\partial t} w = \mathcal{M}_{1/k} \mathcal{L}^{rw} w = \mathcal{M}_{1/k} \mathcal{L}^{rw} \mathcal{M}_k u = \widetilde{\mathcal{L}} u.$$

Since $w(\cdot,t) \geq 0 \iff u(\cdot,t) \geq 0$, it remains to prove the positivity of $u(\cdot,t)$. Choose $\epsilon > 0$ arbitrarily and let $v(x,t) = u(x,t) + \epsilon t$. Observe that $\widetilde{\mathcal{L}} v = \widetilde{\mathcal{L}} u$, and hence

$$\frac{\partial}{\partial t} v - \widetilde{\mathcal{L}} v = \frac{\partial}{\partial t} u + \epsilon - \widetilde{\mathcal{L}} u = \epsilon.$$

Let us show $v(x,t)$ reaches its minimum at some $(a,0)$, $a \in [0,1]$. Assume by contradiction that there exists $(a,\tau) \in [0,1] \times (0,T)$ for some $T > 0$ such that $v(x,t) \geq v(a,\tau)$ for all x and t. It follows that

$$\widetilde{\mathcal{L}} v(a,\tau) = \frac{1}{k(a)} \int_0^1 W(a,y) v(y,\tau) dy - v(a,\tau) \geq \frac{v(a,\tau)}{k(a)} \int_0^1 W(a,y) dy - v(a,\tau) = 0.$$

Hence, $\frac{\partial}{\partial t}v(a,\tau) = \widetilde{\mathcal{L}}v(a,\tau) + \epsilon = \epsilon > 0$ which is in contradiction with the assumption of v attaining its minimum in (a,τ) with $\tau > 0$, so $\tau = 0$. We have thus proved $v(x,t) \geq v(a,0)$, so that

$$u(x,t) + \epsilon t = v(x,t) \geq v(a,0) = u(a,0) = \frac{g(a)}{k(a)} \geq 0.$$

Since ϵ is arbitrary, this allows to conclude. $\qquad\square$

4.4. **The IVP with probability density functions.** Let us observe that when $w(\cdot,t)$ in eq. (23) is a probability density function, it is natural to consider $w(\cdot,t) \in L^1[0,1]$, and one may define $\mathcal{L}^{rw}$ as a mapping $L^1[0,1] \to L^1[0,1]$. Indeed, as in eq. (22) let us still write $\mathcal{K}$ the integral part of $\mathcal{L}^{rw}$ defined on $L^1[0,1]$. By Fubini-Tonelli, the operator norm $\|\mathcal{K}\|_{1,1} := \|\mathcal{K}\|_{L^1[0,1]\to L^1[0,1]}$ satisfies

$$(26) \qquad \|\mathcal{K}\|_{1,1} \leq \sup_{\|f\|_1=1} \int_{[0,1]^2} |K(x,y)f(y)|dxdy = \sup_{\|f\|_1=1} \int_{[0,1]} |f(y)|dy = 1.$$

This, combined with the fact that $\|\mathcal{K}f\|_1 = 1$ if $f = 1$, shows that $\|\mathcal{K}\|_{1,1} = 1$, and so even without assumption 4.5, $\mathcal{L}^{rw}$ is a bounded mapping of $L^1[0,1]$ into itself. Additionally, theorem 4.7 about the existence and unicity of a solution to the IVP has a similar formulation and proof in the present case. Further, the positivity established in section 4.3 also applies here, and this would still not require assumption 4.5. The only significant change in the proof of proposition 4.10 would be to use the auxiliary operator $\mathcal{L}^{rw}\mathcal{M}_k$ instead of $\widetilde{\mathcal{L}} = \mathcal{M}_{1/k}\mathcal{L}^{rw}\mathcal{M}_k$. When $w(\cdot,0) \geq 0$ we further have conservation of the L^1 norm :

$$(27) \qquad \frac{\partial}{\partial t}\|w(\cdot,t)\|_1 = \frac{\partial}{\partial t}\int_0^1 |w(x,t)|dx = \frac{\partial}{\partial t}\int_0^1 w(x,t)dx = 0.$$

In the remainder of the paper, for the sake of simplicity and in order to benefit from the Hilbert space framework at a later stage, we will however assume that W satisfies assumption 4.5. This allows to define $\mathcal{L}^{rw}$ as an operator acting on $L^2[0,1]$ and we do not use L^1 but rather the stronger L^2 norms also present in other works about dynamics on graphons [27, 29].

5. Convergence on dense graphs

This section is divided in three parts. The first two parts show that the solution of the discretized problem on $W/\mathcal{P}$ or $W_{[n]}$ converges to that of the continuum IVP in norm $\|\cdot\|_{C([0,T],L^2[0,1])}$ for any $T > 0$. The goal of the third part is to prove that the discrete problem can be approximated by its continuum version.

5.1. **Convergence on the quotient graph $W/\mathcal{P}$.** Let us start with two simple lemmas.

Lemma 5.1. *Let $\mathcal{A}_\eta : L^2[0,1] \to L^2[0,1]$ be an integral operator with bounded kernel A_η. Assume that A_η is a.e.-constant on every cell $P_i \times P_j$ of the uniform partition of $[0,1]^2$. Further let $f \in L^2[0,1]$ and define f_η by*

$$f_\eta(x) = n \sum_{i=1}^n \int_{P_i} f(y)dy \chi_{P_i}(x), \quad \forall x \in [0,1].$$

Then for all $\ell \in \mathbb{N}_0$, it holds that $\mathcal{A}_\eta^\ell f = \mathcal{A}_\eta^\ell f_\eta$.

Proof. The proof in the case $\ell = 1$ follows from a direct calculation, see for instance [15], lemma 3. The claim for $\ell > 1$ is a direct consequence since then

$$\mathcal{A}_\eta^\ell f = \mathcal{A}_\eta^{\ell-1} \mathcal{A}_\eta f = \mathcal{A}_\eta^{\ell-1} \mathcal{A}_\eta f_\eta = \mathcal{A}_\eta^\ell f_\eta.$$

□

Lemma 5.2. *Let $\mathcal{A}, \mathcal{B} : L^2[0,1] \to L^2[0,1]$ be two Hilbert-Schmidt integral operators with respective kernels A and B defined on the unit square, with $A \leq \beta$ for some constant $\beta > 0$. Then, for all $f \in L^2[0,1]$ and $\ell \in \mathbb{N}_0$*

$$\|\mathcal{A}^\ell f - \mathcal{B}^\ell f\|_2 \leq \beta^{\ell-1} \|A - B\|_2 \|f\|_2 + \|(\mathcal{A}^{\ell-1} - \mathcal{B}^{\ell-1})\mathcal{B}f\|_2.$$

Proof. Using the Minkowski inequality, we have

$$\|\mathcal{A}^\ell f - \mathcal{B}^\ell f\|_2 = \|\mathcal{A}^{\ell-1}\mathcal{A}f - \mathcal{B}^{\ell-1}\mathcal{B}f\|_2$$
$$= \|\mathcal{A}^{\ell-1}\mathcal{A}f - \mathcal{A}^{\ell-1}\mathcal{B}f + \mathcal{A}^{\ell-1}\mathcal{B}f - \mathcal{B}^{\ell-1}\mathcal{B}f\|_2$$
$$= \|\mathcal{A}^{\ell-1}(\mathcal{A}f - \mathcal{B}f) + (\mathcal{A}^{\ell-1} - \mathcal{B}^{\ell-1})\mathcal{B}f\|_2$$
$$(28) \qquad \leq \|\mathcal{A}^{\ell-1}(\mathcal{A}f - \mathcal{B}f)\|_2 + \|(\mathcal{A}^{\ell-1} - \mathcal{B}^{\ell-1})\mathcal{B}f\|_2.$$

Now $\mathcal{A}^{\ell-1}$ is a Hilbert-Schmidt integral operator with kernel $A^{\circ(\ell-1)}$. For such operator, as a product of the Cauchy-Schwarz inequality it is known about the operator norm $\|\cdot\|$ that $\|\mathcal{A}^{\ell-1}\| \leq \|A^{\circ(\ell-1)}\|_2$, or equivalently

$$(29) \qquad \|\mathcal{A}^{\ell-1} f\|_2 \leq \|A^{\circ(\ell-1)}\|_2 \|f\|_2.$$

The first term in the right hand side of eq. (28) therefore satisfies

$$(30) \qquad \|\mathcal{A}^{\ell-1}(\mathcal{A}f - \mathcal{B}f)\|_2 \leq \|A^{\circ(\ell-1)}\|_2 \|\mathcal{A}f - \mathcal{B}f\|_2 \leq \beta^{\ell-1} \|\mathcal{A}f - \mathcal{B}f\|_2$$

where we use $\|A^{\circ(\ell-1)}\|_2 \leq \|A\|_2^{\ell-1}$ ([15], lemma 6) and $A(x,y) \leq \beta$ for all $0 \leq x, y \leq 1$ to obtain the last inequality. Using again eq. (29) with $\ell = 2$, we also have $\|\mathcal{A}f - \mathcal{B}f\|_2 \leq \|A - B\|_2 \|f\|_2$ which, together with eq. (28) and eq. (30) leads to the conclusion. □

Now we are in a place to formulate the convergence results. The continuous formulation of the discrete problem associated to eq. (23) on the quotient graph reads[4]

$$(31a) \qquad\qquad \frac{\partial}{\partial t} u(x,t) = \mathcal{L}_\square^{rw} u(x,t)$$

$$(31b) \qquad\qquad u(x,0) = g_\square(x)$$

where the random walk Laplacian operator on $W/\mathcal{P}$ satisfies

$$(32) \qquad \mathcal{L}_\square^{rw} f(x) = \int_0^1 \frac{\eta\,(W/\mathcal{P})\,(x,y)}{k_\square(y)} f(y)dy - f(x), \quad \forall f \in L^2[0,1],$$

and the initial condition is averaged on each cell of the partition as

$$(33) \qquad g_\square(x) = n \sum_{i=1}^n \int_{P_i} g(y)dy \chi_{P_i}(x), \quad \forall x \in [0,1].$$

Based on the following proposition, operator $\mathcal{L}_\square^{rw}$ is well-defined.

---

[4]The subscript $\square$ refers to fact that the averaging is performed on square cells of $[0,1]^2$. To lighten the notations, we do not refer explicitly to the number of nodes of the graph, so we write $u(x,t)$ instead of, for instance, $u^{(n)}(x,t)$.

Proposition 5.3. *Let W be a connected graphon satisfying assumption 4.5, then the strength of every node of the quotient graph determined by the partition $\mathcal{P} = \{P_1, \ldots, P_n\}$ of $[0, 1]$ is positive.*

Proof. The strength of the i-th node v_i, $i = 1, \ldots, n$, is given by $\mathrm{str}(v_i) = nk_\square(x)$, for every $x \in P_i$. We have

$$(34) \qquad k_\square(x) = \int_0^1 \sum_{j=1}^n A_{ij} \chi_{P_j}(y) dy = \frac{1}{n} \sum_{j=1}^n A_{ij}, \quad \forall x \in P_i,$$

where A_{ij} was defined by eq. (5). Hence,

$$k_\square(x) = n \int_{P_i} \sum_{j=1}^n \int_{P_j} W(x', y') dy' dx' = n \int_{P_i} \int_0^1 W(x', y') dy' dx' = n \int_{P_i} k(x') dx',$$

showing $k_\square(x) \geq c$ where $c > 0$ is the constant from assumption 4.5. $\qquad\square$

Remark 5.4. It follows that the finite-dimensional IVP eq. (31) on the quotient graph has a unique solution given by $e^{t\mathcal{L}_\square} g_\square$.

Theorem 5.5 (Convergence with $W/\mathcal{P}$)**.** *Let W be a connected graphon satisfying assumption 4.5, and let $w(x, t)$ be the solution of IVP eq. (23). Further let $u(x, t)$ be the solution of the associated discrete problem eq. (31). Then for all $t \in \mathbb{R}^+$ it holds that*

$$\|u(\cdot, t) - w(\cdot, t)\|_2 \to 0 \quad as \quad n \to \infty.$$

Proof. Using remark 5.4, by the Minkowski inequality we have

$$\|u(\cdot, t) - w(\cdot, t)\|_2 = \left\| e^{t\mathcal{L}_\square} g_\square - e^{t\mathcal{L}} g \right\|_2$$

$$= \left\| \sum_{k=0}^\infty \frac{t^k}{k!} \mathcal{L}_\square^k g_\square - \sum_{k=0}^\infty \frac{t^k}{k!} \mathcal{L}^k g \right\|_2$$

$$(35) \qquad\qquad \leq \|g_\square - g\|_2 + \sum_{k=1}^\infty \frac{t^k}{k!} \underbrace{\left\| \mathcal{L}_\square^k g_\square - \mathcal{L}^k g \right\|_2}_{(*)}.$$

Let us write $\mathcal{L}^{rw} = \mathcal{K} - \mathcal{I}$ where $\mathcal{K}$ is the operator previously defined in eq. (22) and $\mathcal{I}$ is the identity operator. We have a similar decomposition $\mathcal{L}_\square^{rw} = \mathcal{K}_\square - \mathcal{I}$ for the Laplacian of the step graphon. For $k \geq 1$ and $0 \leq m \leq k$ let us write $\alpha_{mk} = (-1)^m \binom{k}{m}$, and consider $(*)$ in the right-hand side of eq. (35). Using Newton's binomial theorem we have

$$\left\| \mathcal{L}_\square^k g_\square - \mathcal{L}^k g \right\|_2 = \left\| (\mathcal{K}_\square - \mathcal{I})^k g_\square - (\mathcal{K}^k - \mathcal{I}) g \right\|_2$$

$$= \left\| \sum_{m=0}^k \alpha_{mk} \mathcal{K}_\square^{k-m} g_\square - \sum_{m=0}^k \mathcal{K}^{k-m} g \right\|_2$$

$$\leq \left\| \sum_{m=0}^{k-1} \alpha_{mk} \left(\mathcal{K}_\square^{k-m} g_\square - \mathcal{K}^{k-m} g \right) \right\|_2 + \left\| \alpha_{kk} (g_\square - g) \right\|_2$$

with $|\alpha_{mk}| = \binom{k}{m}$ and using lemma 5.1,

$$(36) \qquad\qquad \leq \sum_{m=0}^{k-1} \binom{k}{m} \underbrace{\left\| \left(\mathcal{K}_\square^{k-m} - \mathcal{K}^{k-m} \right) g \right\|_2}_{(**)} + \|(g_\square - g)\|_2.$$

By assumption 4.5 and proposition 5.3 there exists some constant $c > 0$ such that $k_\square(y) \geq c$ for all $y \in [0,1]$. Further, $0 \leq W \leq 1$ on $[0,1]^2$, and so

$$(37) \qquad \|K_\square\|_2 = \left\|\frac{\eta(W/\mathcal{P})}{k}\right\|_2 = \|\eta(W/\mathcal{P})\|_2 \left\|\frac{1}{k_\square}\right\|_2 \leq \frac{1}{c} =: \beta_\square,$$

where $K_\square$ denotes the integral kernel of $\mathcal{K}_\square$. For $\ell \in \mathbb{N}_0$, let us define $\mathcal{E}_\ell := \mathcal{K}_\square^\ell - \mathcal{K}^\ell$ and $E_\ell := K_\square^\ell - K^\ell$. Then, applying lemma 5.2 successively $\ell - 1$ times to $(\ast\ast)$ in eq. (36) with $\ell = k - m$, we obtain

$$\|\mathcal{E}_\ell g\|_2 \leq \beta_\square^{\ell-1} \|E_1\|_2 \|g\|_2 + \|\mathcal{E}_{\ell-1}\mathcal{K}g\|_2$$
$$\leq \left(\beta_\square^{\ell-1} + \beta_\square^{\ell-2}\right) \|E_1\|_2 \|g\|_2 + \|\mathcal{E}_{\ell-2}\mathcal{K}^2 g\|_2$$
$$\vdots$$
$$\leq \left(\sum_{j=1}^{\ell-1} \beta_\square^{\ell-j}\right) \|E_1\|_2 \|g\|_2 + \|\mathcal{E}_1 \mathcal{K}^{\ell-1} g\|_2,$$

and since E_ℓ is the kernel of $\mathcal{E}_\ell$ if $\ell = 1$,

$$\leq \left(\sum_{j=1}^{\ell-1} \beta_\square^{\ell-j}\right) \|E_1\|_2 \|g\|_2 + \|E_1\|_2 \|\mathcal{K}^{\ell-1} g\|_2$$

and with $\|\mathcal{K}^{\ell-1} g\|_2 \leq \|K^{\circ(\ell-1)}\|_2 \|g\|_2 \leq \|K\|_2^{\ell-1} \|g\|_2$,

$$\leq \left(\|K\|_2^{\ell-1} + \sum_{j=1}^{\ell-1} \beta_\square^{\ell-j}\right) \|E_1\|_2 \|g\|_2$$
$$(38) \qquad \leq \ell \beta^{\ell-1} \|E_1\|_2 \|g\|_2,$$

where the last inequality stems from $\beta := \max\{\|K\|_2, \beta_\square\} \geq 1$. Combining eq. (36) and eq. (38) yields

$$\left\|\mathcal{L}_\square^k g_\square - \mathcal{L}^k g\right\|_2 \leq \sum_{m=0}^{k-1} \binom{k}{m} (k-m) \beta^{k-m-1} \|K_\square - K\|_2 \|g\|_2 + \|(g_\square - g)\|_2$$
$$(39) \qquad \leq \beta^{k-1} \sum_{m=0}^{k-1} \binom{k}{m} (k-m) \|K_\square - K\|_2 \|g\|_2 + \|(g_\square - g)\|_2.$$

From eqs. (35) and (39) we obtain

$$\|u(\cdot,t) - w(\cdot,t)\|_2 \leq \|g_\square - g\|_2 + \sum_{k=1}^\infty \frac{t^k}{k!} \|(g_\square - g)\|_2$$
$$+ \|K_\square - K\|_2 \|g\|_2 \sum_{k=1}^\infty \frac{t^k}{k!} \beta^{k-1} \sum_{m=0}^{k-1} \binom{k}{m} (k-m)$$
$$= \|g_\square - g\|_2 e^t + \|K_\square - K\|_2 \|g\|_2 \sum_{k=1}^\infty \frac{t^k}{k!} \beta^{k-1} \sum_{m=0}^{k-1} \binom{k}{m} (k-m)$$

and with $\sum_{m=0}^{k-1} \binom{k}{m}(k-m) = k2^{k-1}$,

$$= \|g_\square - g\|_2 e^t + \|K_\square - K\|_2 \|g\|_2 \sum_{k=1}^{\infty} \frac{t^k}{k!} k(2\beta)^{k-1}$$

$$\leq \|g_\square - g\|_2 e^t + \|K_\square - K\|_2 \|g\|_2 t \sum_{k=1}^{\infty} \frac{(2\beta t)^{k-1}}{(k-1)!}$$

(40)
$$\leq \|g_\square - g\|_2 e^t + \underbrace{\|K_\square - K\|_2}_{(***)} \|g\|_2 t e^{2\beta t}.$$

By the Lebesgue differentiation theorem, $g_\square \to g$ pointwise for almost every $x \in [0,1]$ as $n \to \infty$, so that

(41)
$$\|g_\square - g\|_2 \xrightarrow[n \to 0]{} 0$$

by dominated convergence [27]. Let us consider $(***)$ in eq. (40) :

$$\|K_\square - K\|_2^2 = \int_{[0,1]^2} \left(\frac{\eta(W/\mathcal{P})(x,y)}{k_\square(x,y)} - \frac{W(x,y)}{k(y)} \right)^2 dxdy$$

$$\leq \operatorname*{ess\,sup}_{y \in [0,1]} \frac{1}{k_\square^2(y) k^2(y)} \int_{[0,1]^2} \left(\eta(W/\mathcal{P})(x,y)k(y) - W(x,y)k_\square(y) \right)^2 dxdy$$

$$\leq \beta^2 \int_{[0,1]^2} \left(\eta(W/\mathcal{P})(x,y)(k(y) - k_\square(y)) \right)^2 dxdy$$

$$+ \beta^2 \int_{[0,1]^2} \left((W(x,y) - \eta(W/\mathcal{P})(x,y))k_\square(y) \right)^2 dxdy$$

and because $\|\eta(W/\mathcal{P})\|_2 \leq 1$ and $\|k_\square\|_2 \leq 1$,

(42)
$$\leq \beta^2 \left(\|k - k_\square\|_2^2 + \|W - \eta(W/\mathcal{P})\|_2^2 \right).$$

By the Cauchy-Schwarz inequality,

$$\|k - k_\square\|_2^2 = \int_0^1 \left(\int_0^1 (W(y,z) - \eta(W/\mathcal{P})(y,z)) \, dz \right)^2 dy$$

$$\leq \int_0^1 \int_0^1 (W(y,z) - \eta(W/\mathcal{P})(y,z))^2 \, dzdy$$

$$= \|W - \eta(W/\mathcal{P})\|_2^2,$$

which together with eq. (42) yields

(43)
$$\|K_\square - K\|_2^2 \leq 2\beta^2 \|W - \eta(W/\mathcal{P})\|_2^2.$$

By the same argument leading to eq. (41), we have $\|W - \eta(W/\mathcal{P})\|_2 \to 0$ as $n \to \infty$ which with eq. (43) implies

(44)
$$\|K_\square - K\|_2 \xrightarrow[n \to 0]{} 0.$$

Combining eqs. (40), (41), and (44) allows to conclude. $\qquad \square$

5.2. **Convergence on the sampled graph $W_{[n]}$.** The case of the discrete problem on $W_{[n]}$ can be handled similarly as the discrete problem on $W/\mathcal{P}$, and the convergence theorem follows mainly from the observation in section 2.3 that $W_{[n]} \to W$ at every point of continuity of W. The necessary convergence in L^2 will follow from the supplemental assumption that the graphon is almost everywhere continuous. The discrete problem (in its step function form) associated to eq. (23) on the sampled graph $W_{[n]}$ reads

$$(45a) \qquad\qquad\qquad \frac{\partial}{\partial t}u(x,t) = \mathcal{L}_{[n]}^{rw}u(x,t)$$

$$(45b) \qquad\qquad\qquad u(x,0) = g_\square(x)$$

where the random walk Laplacian operator on $W_{[n]}$ satisfies

$$(46) \qquad \mathcal{L}_{[n]}^{rw}f(x) = \int_0^1 \frac{\eta\left(W_{[n]}\right)(x,y)}{k_{[n]}(y)}f(y)dy - f(x), \quad \forall f \in L^2[0,1],$$

and the initial condition is again averaged on each cell of the partition as in eq. (33). One needs to assume sufficiently large n to guarantee $k_{[n]}$ to be bounded away from 0 and so the Laplacian to be well-defined.

Theorem 5.6 (Convergence with $W_{[n]}$). *Let W be a connected, almost everywhere continuous graphon satisfying assumption 4.5 and let $w(x,t)$ be the solution of IVP eq. (23). Further let $u(x,t)$ be the solution of the associated discrete problem eq. (45). Then for all $t \in \mathbb{R}^+$ it holds that*

$$\|u(\cdot,t) - w(\cdot,t)\|_2 \to 0 \quad as \quad n \to \infty.$$

The proof is similar as for theorem 5.5.

Remark 5.7. The initial condition could have been sampled in a similar fashion as the graphon, to yield the step function $g_{[n]} = \sum_{i=1}^n g\left(\frac{i}{n}\right)\chi_{P_i}$. Almost everywhere continuity of g would ensure that $\|g - g_{[n]}\|_2 \to 0$ when $n \to \infty$, and would be part of the hypothesis of a convergence theorem. The proof of theorem 5.6 would only require minor changes, which are similar to those discussed next in the new context of section 5.3.

5.3. **Convergence for a sequence of discrete problems.** This time we consider a sequence of problems defined on graphs with increasing number of nodes. We assume the sequence of dense connected graphs, say (G_n), converges to a limit graphon W in the L^2 metric, in the sense that $\|\eta(G_n) - W\|_2 \to 0$ as $n \to \infty$. Let k_n denote the degree function of the empirical graphon $\eta(G_n)$. Consider the family of discrete problems under the mapping η

$$(47a) \qquad\qquad\qquad \frac{\partial}{\partial t}u(x,t) = \mathcal{L}_n^{rw}u(x,t)$$

$$(47b) \qquad\qquad\qquad u(x,0) = g_n(x) \in L^2[0,1],$$

where the random walk Laplacian operator $\mathcal{L}_n^{rw}$ satisfies

$$(48) \qquad \mathcal{L}_n^{rw}f(x) = \int_0^1 \frac{\eta\left(G_n\right)(x,y)}{k_n(y)}f(y)dy - f(x), \quad \forall f \in L^2[0,1].$$

Similarly as before, we write $\mathcal{L}_n^{rw} = \mathcal{K}_n - \mathcal{I}$.

Theorem 5.8 (Convergence with (G_n))**.** *Let (G_n) be a sequence of connected graphs that converges to a connected graphon W satisfying assumption 4.5. Let $w(x,t)$ be the solution of the IVP eq. (23) associated to W with initial condition $w(\cdot,0) = g \in L^2[0,1]$. Further let $u(x,t)$ be the solution of the corresponding discrete problem eq. (47), and assume that $\|g_n - g\|_2 \to 0$ as $n \to \infty$. Then for all $t \in \mathbb{R}^+$ it holds that*

$$\|u(\cdot,t) - w(\cdot,t)\|_2 \to 0 \quad as \quad n \to \infty.$$

Proof. The proof follows the same steps as for theorem 5.5. However, using lemma 5.1 to obtain eq. (36) is now prohibited due to the initial condition of a discrete problem no longer resulting from an averaging of the continuous IVP. Consider a sufficiently large n such that the degree function of the empirical graphon satisfies $k_n \geq c$ for some constant $c > 0$. Not relying this time on lemma 5.1, we write

$$\|\mathcal{K}_n^{\ell-m} g_n - \mathcal{K}^{k-m} g\|_2 = \|\mathcal{K}_n^{\ell-m} g_n - \mathcal{K}_n^{\ell-m} g + \mathcal{K}_n^{\ell-m} g - \mathcal{K}^{k-m} g\|_2$$
$$\leq \|\mathcal{K}_n^{\ell-m}(g_n - g)\|_2 + \|(\mathcal{K}_n^{\ell-m} - \mathcal{K}^{\ell-m})g\|_2,$$

with the first term in the right-hand side newly present. Following the same steps leading to eq. (36), we obtain

$$\left\|\mathcal{L}_n^k g_n - \mathcal{L}^k g\right\|_2 \leq \sum_{m=0}^{k-1} \binom{k}{m} \|\mathcal{K}_n^{\ell-m}(g_n - g)\|_2$$
$$+ \sum_{m=0}^{k-1} \binom{k}{m} \left\|\left(\mathcal{K}_n^{k-m} - \mathcal{K}^{k-m}\right) g\right\|_2 + \|(g_n - g)\|_2,$$

where again the first term right of the inequality is new. In fashion similar to the proof of theorem 5.5, with $\beta := \max\left\{\|K\|_2, \left\|\frac{1}{k_n}\right\|_\infty\right\}$ we have

$$\|u(\cdot,t) - w(\cdot,t)\|_2 \leq \sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_{m=0}^{k-1} \binom{k}{m} \|\mathcal{K}_n^{\ell-m}(g_n - g)\|_2$$
$$+ \|g_n - g\|_2 e^t + \|K_n - K\|_2 \|g\|_2 t e^{2\beta t}.$$

Using $\|\mathcal{K}^{k-m}\| \leq \beta^{k-m} \leq \beta^k$ and $\sum_{m=0}^{k-1} \binom{k}{m} \leq 2^k$, we have

$$\sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_{m=0}^{k-1} \binom{k}{m} \|\mathcal{K}_n^{\ell-m}(g_n - g)\|_2 \leq \sum_{k=1}^{\infty} \frac{t^k}{k!} 2^k \beta^k \|(g_n - g)\|_2 = \|(g_n - g)\|_2 e^{2\beta t},$$

leading to

$$\|u(\cdot,t) - w(\cdot,t)\|_2 \leq \|g_n - g\|_2 \left(e^t + e^{2\beta t}\right) + \|K_n - K\|_2 \|g\|_2 t e^{2\beta t}.$$

$\qquad\square$

6. Relaxation

The evolution of a system towards its asymptotic state w_∞ starting from any initial condition is know as relaxation. The so-called relaxation time characterizes the rate of this evolution. In the continuum limit of the node-centric walk, it is determined by the spectral properties of $\mathcal{K}$, in a way reminiscent of random walks on finite graphs. For the node-centric continuous-time walk, we will show now that

this rate can be exponential. Let us define a normalized adjacency operator, which is then used in the definition of a normalized Laplacian.

Definition 6.1 (Normalized adjacency operator). Under assumption 4.5, let the normalized adjacency operator be the integral operator $\mathcal{A}^{norm} : L^2[0,1] \to L^2[0,1]$ defined by $\mathcal{A}^{norm} = \mathcal{M}_{1/\sqrt{k}} \mathcal{K} \mathcal{M}_{\sqrt{k}}$.

Observe that under assumption 4.5 the kernel $W(x,y)/\sqrt{k(x)k(y)}$ of $\mathcal{A}^{norm}$ is square-integrable and symmetric. Hence $\mathcal{A}^{norm}$ is a compact, self-adjoint Hilbert-Schmidt integral operator and the Hilbert-Schmidt theorem applies. Therefore, there exists an orthonormal basis of eigenfunctions $\{\phi_m\}$ with associated eigenvalues θ_m, so that operator $\mathcal{A}^{norm}$ has the canonical form

$$(49) \qquad \mathcal{A}^{norm} = \sum_{m=1}^{\infty} \theta_m \left(\phi_m, \cdot\right) \phi_m.$$

The operator $\mathcal{L}^{norm} := \mathcal{A}^{norm} - \mathcal{I}$ is the associated normalized (or sometimes also called symmetric) Laplacian. Note that for $\ell \in \mathbb{N}$, $(\mathcal{A}^{norm})^{\ell}$ has eigenfunctions ϕ_m and eigenvalues θ_m^{ℓ}, and that $(\mathcal{L}^{rw})^{\ell} = \mathcal{M}_{\sqrt{k}} (\mathcal{L}^{norm})^{\ell} \mathcal{M}_{1/\sqrt{k}}$. Combined with eq. (49) this yields the singular value decomposition

$$e^{\mathcal{L}^{rw} t} = \mathcal{M}_{\sqrt{k}} \left(\sum_{\ell=0}^{\infty} \frac{t^{\ell}}{\ell!} \left(\sum_{m=1}^{\infty} \theta_m \left(\phi_m, \cdot\right) \phi_m - \mathcal{I} \right)^{\ell} \right) \mathcal{M}_{1/\sqrt{k}}$$

$$= \mathcal{M}_{\sqrt{k}} \left(\sum_{\ell=0}^{\infty} \frac{t^{\ell}}{\ell!} \left(\sum_{m=1}^{\infty} (\theta_m - 1) \left(\phi_m, \cdot\right) \phi_m \right)^{\ell} \right) \mathcal{M}_{1/\sqrt{k}}$$

$$= \mathcal{M}_{\sqrt{k}} \left(\sum_{m=1}^{\infty} \sum_{\ell=0}^{\infty} \frac{t^{\ell}}{\ell!} \lambda_m^{\ell} \left(\phi_m, \cdot\right) \phi_m \right) \mathcal{M}_{1/\sqrt{k}}$$

$$(50) \qquad = \sum_{m=1}^{\infty} e^{\lambda_m t} \left(\frac{\phi_m}{\sqrt{k}}, \cdot \right) \sqrt{k} \phi_m$$

with $\lambda_m = \theta_m - 1$ the eigenvalues of $\mathcal{L}^{norm}$. By letting $\psi_m = \frac{\phi_m}{\sqrt{k}}$ and $\zeta_m = \sqrt{k} \phi_m$, the solution of IVP eq. (23) reads

$$(51) \qquad w(x,t) = \sum_{m=1}^{\infty} e^{\lambda_m t} \left(\psi_m, g\right) \zeta_m(x).$$

The following proposition allows for a characterization of the rate of the relaxation towards w_{∞}.

Proposition 6.2. *Let W be a graphon satisfying assumption 4.5, then the eigenvalues λ_m of $\mathcal{L}^{norm}$ are non-positive reals, and the largest eigenvalue is zero. If moreover W is connected, then the eigenvalue zero has multiplicity one.*

Proof. That the eigenvalues are reals results from $\mathcal{L}^{norm}$ being a self-adjoint operator on $L^2[0,1]$. Let λ be an eigenvalue of $\mathcal{L}^{norm}$ associated to ϕ. Then λ is given by the Rayleigh quotient

$$(52) \qquad \lambda = \frac{(\lambda \phi, \phi)}{(\phi, \phi)} = \frac{(\mathcal{L}^{norm} \phi, \phi)}{(\phi, \phi)}.$$

Consider the numerator in the right-hand side of eq. (52). For all $f \in L^2[0,1]$ we can write

$$(\mathcal{L}^{norm} f, f) = \int_0^1 \int_0^1 \frac{W(x,y)}{\sqrt{k(x)}\sqrt{k(y)}} f(x)f(y)dxdy - \int_0^1 f^2(x)dx$$

$$= \frac{1}{2}\left(2 \int_0^1 \int_0^1 \frac{\sqrt{W(x,y)}}{\sqrt{k(x)}} \frac{\sqrt{W(x,y)}}{\sqrt{k(y)}} f(x)f(y)dxdy \right.$$

$$\left. - \int_0^1 \int_0^1 \frac{W(x,y)}{k(x)} f^2(x)dxdy - \int_0^1 \int_0^1 \frac{W(x,y)}{k(y)} f^2(y)dxdy \right)$$

$$(53) \qquad = -\frac{1}{2}\int_0^1 \int_0^1 \left(\frac{\sqrt{W(x,y)}}{\sqrt{k(x)}} f(x) - \frac{\sqrt{W(x,y)}}{\sqrt{k(y)}} f(y)\right)^2 dxdy,$$

which is non-positive. The claim that zero is an eigenvalue follows from the fact that $\mathcal{L}^{norm}\sqrt{k(x)} = 0$ on $[0,1]$. Finally, let us show that the nullspace of $\mathcal{L}^{norm}$ has dimension one if W is connected. By defining $g = \mathcal{M}_{1/\sqrt{k}}f$ on $[0,1]$, we have to show that if the right-hand side of eq. (53) is zero, namely

$$(54) \qquad -\frac{1}{2}\iint_{[0,1]^2} W(x,y)\left(g(x) - g(y)\right)^2 dxdy = 0,$$

then g has to be a constant function on $[0,1]$. By contradiction, assume that there exists some non-constant function g that verifies eq. (54). For simplicity, consider the case that $g = c_1$ on some $S \in \mathfrak{M}[0,1]$ with $\mu(S) \in (0,1)$, and $g = c_2$ on $S^c := [0,1] \setminus S$, with $c_1, c_2 \in \mathbb{R}$, $c_1 \neq c_2$. The reasoning would be similar if g is a piecewise constant function on any other partition of $[0,1]$, and can be extended by density to any not piecewise constant g. Based on eq. (54), we can write
(55)
$$0 = \iint_{[0,1]^2} W(x,y)\left(g(x) - g(y)\right)^2 dxdy \geq \iint_{S \times S^c} W(x,y)\left(g(x) - g(y)\right)^2 dxdy,$$

and the integral in the right-hand side is zero. Since W is connected, we have $\iint_{S \times S^c} W(x,y)dxdy > 0$, and hence there exists a positive-measured subset $E \times F$ of $S \times S^c$ such that $W > 0$ on $E \times F$. Therefore, $g(x) - g(y) = 0$ for almost every $(x,y) \in E \times F$. But then, since $g = c_1$ on $E \subset S$, $g = c_1$ on $F \subset S^c$, a contradiction. $\qquad\square$

Remark 6.3 (Spectral gap). The last claim of proposition 6.2 means that the spectral gap of $\mathcal{L}^{norm}$, namely the positive difference between the largest and the second largest eigenvalue, is nonzero when the graphon is connected with degree function bounded away from zero. Observe that if k is not bounded away from zero, we may no longer write $\mathcal{A}^{norm} = \mathcal{M}_{1/\sqrt{k}}\mathcal{K}\mathcal{M}_{\sqrt{k}}$ because $1/\sqrt{k}$ is not bounded. This implies that the spectrum of $\mathcal{L}^{rw}$ can no longer be deduced directly from the spectrum of the compact self-adjoint operator $\mathcal{L}^{norm}$. However, the eigenvalues of $\mathcal{L}^{rw}$ may in some cases be computed directly, see example 6.4 hereafter. If the graphon is not connected, one can analyze the dynamics on each connected component independently, as follows from the decomposition introduced in [19]. Therefore, it only remains to consider relaxation in the case of a connected graphon where assumption 4.5 is not satisfied, meaning $k(x)$ becomes arbitrarily small on positive measured subsets of $[0,1]$, and such that $\mathcal{K}$ is still well-defined. In this situation

the eigenvalues of the generally non-self-adjoint operator $\mathcal{L}^{rw}$ may be embedded in the continuous spectrum whilst in the discrete or discretized version, the spectrum is composed only of eigenvalues and there can be a positive spectral gap as a result of connectedness in a finite graph.

Example 6.4 (Eigenvalues of $\mathcal{L}^{rw}$ on a separable graphon). Consider the separable[5] graphon $W(x,y) = xy$. The degree function is $k(x) = x/2$, in which case $\mathcal{L}^{rw} = \mathcal{K} - \mathcal{I}$ with $\mathcal{K}f(x) = 2x \int_0^1 f(y)dy$. Any eigenvalue $\lambda_{\mathcal{K}}$ of $\mathcal{K}$ satisfies

$$(56) \qquad 2x \int_0^1 \phi_{\mathcal{K}}(y)dy = \lambda_{\mathcal{K}}\phi_{\mathcal{K}}(x), \quad x \in [0,1],$$

where $\phi_{\mathcal{K}}$ is an eigenfunction. It suffices to subtract one to the eigenvalues of $\mathcal{K}$ and to hold the same eigenfunctions to obtain the eigenpairs of $\mathcal{L}^{rw}$. From (56), one finds $\lambda_{\mathcal{K}} = 1$ with the one-dimensional eigenspace span$\{x\}$, or $\lambda_{\mathcal{K}} = 0$ with the infinite-dimensional eigenspace $\{1\}^{\perp}$. Observe these spaces are not orthogonal, but their sum is $L^2[0,1]$.

7. Extension to the discrete-time walk

The analysis of the node-centric continuous-time walk carries over to the discrete-time version eq. (11). The corresponding IVP on the continuum reads

$$(57a) \qquad w(\cdot, \ell+1) = \mathcal{K}w(\cdot, \ell), \quad \ell \in \mathbb{N}_0$$

$$(57b) \qquad w(\cdot, 0) = w_0 \in L^2[0,1],$$

with solution given by $w(\cdot, \ell) = \mathcal{K}^{\ell}w_0$ for every $\ell \in \mathbb{N}$.

Following the same steps as in sections 3.2, 4, and 5, we obtain a similar convergence result on the quotient graph $W/\mathcal{P}$. Indeed, the proof in continuous time already contains the necessary bound on the operator norm of the difference $\mathcal{K}^{\ell} - \mathcal{K}_{\square}^{\ell}$, see eq. (38). The same holds true on the sampled graph $W_{[n]}$ and for a sequence of discrete problems. Analogously as for eq. (51), the spectral expansion of the solution of the discrete-time IVP eq. (57) reads

$$(58) \qquad w(\cdot, \ell) = \sum_{m=1}^{\infty} \lambda_m^{\ell} \left(\psi_m, w_0\right) \zeta_m, \quad \ell \geq 0.$$

8. Conclusion

There are two main arguments motivating this work. On the one hand, random walks and Laplacians play a central role in the study of graphs, and a better understanding of their behavior on graphons has a clear mathematical interest, with theoretical and algorithmic objectives. On the other hand, even though large networks become ever more common in numerous fields of research, a rigorous study of the continuum limit of the different types of random walks on graphs was still lacking.

This paper is intended as a first step towards a systematic study of classes of random walks on discrete domains, relying on the adequate framework provided by graph-limit theory. We have first shown that the continuum-limit of the discrete heat equation [27] could be interpreted as the limit of a rescaled edge-centric

---

[5]The graphon $W(x,y)$ is separable if it can be written as $W(x,y) = \zeta(x)\zeta(y)$ for some function ζ.

continuous-time Poisson random walk. We have then studied the continuum limit of the remaining two fundamental classes of random walks on graphs, which complement the discrete heat equation : the discrete-time walk, and its continuous-time generalization. A final part of the document was devoted to spectral aspects of the introduced random walk Laplacian operator, thereby allowing for a characterization of the relaxation time of the process.

The world of random walks is a very broad one, and this initial work calls for extensions. A promising research direction would consist in generalizing the semigroup approach developed here, or the one in [27, 28, 29], to the diverse classes of random walk processes omitted here, for instance walks on temporal or directed networks. A second line of research could focus on the case of sparse graphs, and the way they affect the approximation procedure we have applied. Sparsity is indeed known to be the norm rather than the exception in real-life networks. Such extension was already provided for the graph-limit version of the heat equation, using L^p graphons [5, 28, 20]. Another possible venue of investigation could follow from recent work on sparse exchangeable graphs generated via graphon processes or graphexes [3, 4, 8].

References

[1] A.-L. BARABÁSI, *Network science*, Cambridge university press, 2016.

[2] M. BELKIN AND P. NIYOGI, *Towards a theoretical foundation for laplacian-based manifold methods*, Journal of Computer and System Sciences, 74 (2008), pp. 1289 – 1308. Learning Theory 2005.

[3] C. BORGS, J. CHAYES, L. LOVÁSZ, V. SÓS, AND K. VESZTERGOMBI, *Limits of randomly grown graph sequences*, European Journal of Combinatorics, 32 (2011), pp. 985 – 999.

[4] C. BORGS, J. T. CHAYES, H. COHN, AND N. HOLDEN, *Sparse exchangeable graphs and their limits via graphon processes*, arXiv preprint arXiv:1601.07134, (2016).

[5] C. BORGS, J. T. CHAYES, H. COHN, AND Y. ZHAO, *An l^p theory of sparse graph convergence i: limits, sparse random graph models, and power law distributions*.

[6] C. BORGS, J. T. CHAYES, L. LOVÁSZ, V. T. SÓS, AND K. VESZTERGOMBI, *Convergent sequences of dense graphs i: Subgraph frequencies, metric properties and testing*, Advances in Mathematics, 219 (2008), pp. 1801–1851.

[7] S. BRIN AND L. PAGE, *Anatomy of a large-scale hypertextual web search engine*, Proceedings of the Seventh International World Wide Web Conference, (1998), pp. 107–117.

[8] F. CARON AND E. B. FOX, *Sparse graphs using exchangeable random measures*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79 (2017), pp. 1295–1366.

[9] M. D'ELIA, Q. DU, M. GUNZBURGER, AND R. LEHOUCQ, *Nonlocal convection-diffusion problems on bounded domains and finite-range jump processes*, Computational Methods in Applied Mathematics, 17 (2017), pp. 707–722.

[10] J. C. DELVENNE, S. N. YALIRAKI, AND M. BARAHONA, *Stability of graph communities across time scales*, Proc. Natl. Acad. Sci. USA, 107 (2010), pp. 12755–12760.

[11] P. DIACONIS, S. HOLMES, AND S. JANSON, *Threshold graph limits and random threshold graphs*, Internet Mathematics, 5 (2008), pp. 267–320.

[12] P. DIACONIS AND S. JANSON, *Graph limits and exchangeable random graphs*, arXiv preprint arXiv:0712.2749, (2007).

[13] K.-J. ENGEL AND R. NAGEL, *One-parameter semigroups for linear evolution equations*, in Semigroup Forum, vol. 63, Springer, 2001, pp. 278–280.

[14] C. GAO, Y. LU, H. H. ZHOU, ET AL., *Rate-optimal graphon estimation*, The Annals of Statistics, 43 (2015), pp. 2624–2652.

[15] S. GAO AND P. E. CAINES, *The control of arbitrary size networks of linear systems via graphon limits: An initial investigation*, in 2017 IEEE 56th Annual Conference on Decision and Control (CDC), IEEE, 2017, pp. 1052–1057.

[16] E. Giné, V. Koltchinskii, et al., *Empirical graph laplacian approximation of laplace–beltrami operators: Large sample results*, in High dimensional probability, Institute of Mathematical Statistics, 2006, pp. 238–259.

[17] M. Hein, J.-Y. Audibert, and U. v. Luxburg, *Graph laplacians and their convergence on random neighborhood graphs*, Journal of Machine Learning Research, 8 (2007), pp. 1325–1368.

[18] P. W. Holland, K. B. Laskey, and S. Leinhardt, *Stochastic blockmodels: First steps*, Social networks, 5 (1983), pp. 109–137.

[19] S. Janson, *Connectedness in graph limits*, arXiv preprint arXiv:0802.3795, (2008).

[20] D. Kaliuzhnyi-Verbovetskyi and G. S. Medvedev, *The semilinear heat equation on sparse random graphs*, SIAM Journal on Mathematical Analysis, 49 (2017), pp. 1333–1355.

[21] C. Kuehn and S. Throm, *Power network dynamics on graphons*, SIAM Journal on Applied Mathematics, 79 (2019), pp. 1271–1292.

[22] R. Lambiotte and M. Rosvall, *Ranking and clustering of nodes in networks with smart teleportation*, Physical Review E, 85 (2012), p. 056107.

[23] A. N. Langville and C. D. Meyer, *Deeper inside pagerank*, Internet Mathematics, 1 (2004), pp. 335–380.

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

[25] L. Lovász and B. Szegedy, *Limits of dense graph sequences*, Journal of Combinatorial Theory, Series B, 96 (2006), pp. 933–957.

[26] N. Masuda, M. A. Porter, and R. Lambiotte, *Random walks and diffusion on networks*, Physics Reports, (2017).

[27] G. S. Medvedev, *The nonlinear heat equation on dense graphs and graph limits*, SIAM Journal on Mathematical Analysis, 46 (2014), pp. 2743–2766.

[28] G. S. Medvedev, *Small-world networks of kuramoto oscillators*, Physica D: Nonlinear Phenomena, 266 (2014), pp. 13–22.

[29] G. S. Medvedev and X. Tang, *The kuramoto model on power law graphs: Synchronization and contrast states*, Journal of Nonlinear Science, (2018), pp. 1–23.

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

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

[32] J. Petit, R. Lambiotte, and T. Carletti, *Classes of random walks on temporal networks with competing timescales*, Applied Network Science, 4 (2019), pp. 72–91.

[33] L. Rosasco, M. Belkin, and E. D. Vito, *On learning with integral operators*, Journal of Machine Learning Research, 11 (2010), pp. 905–934.

[34] M. Rosvall, D. Axelsson, and C. T. Bergstrom, *The map equation*, The European Physical Journal Special Topics, 178 (2009), pp. 13–23.

[35] R. Vizute, P. Frasca, and F. Garin, *Graphon-based sensitivity analysis of sis epidemics*, IEEE Control Systems Letters, 4 (2020), pp. 542–547.

[36] U. Von Luxburg, M. Belkin, and O. Bousquet, *Consistency of spectral clustering*, The Annals of Statistics, (2008), pp. 555–586.

[37] W. Woess, *Random walks on infinite graphs and groups*, vol. 138, Cambridge university press, 2000.