Random walks on dense graphs and graphons

Julien Petit
julien.petit@unamur.be
Department of Mathematics, Royal Military Academy, Brussels (Belgium)
aXys, Namur Institute for Complex Systems, Namur (Belgium)

Renaud Lambiotte
renaud.lambiotte@maths.ox.ac.uk
Mathematical Institute, University of Oxford, Oxford (UK)

Timoteo Carletti
timoteo.carletti@unamur.be
naXys, Namur Institute for Complex Systems, Namur (Belgium)

September 27, 2019

Abstract

Graph-limit theory focuses on the convergence of sequences of graphs when the number of nodes becomes arbitrarily large. This framework defines a continuous version of graphs allowing for the study of dynamical systems on very large graphs, where classical methods would become computationally intractable. Through an approximation procedure, the standard system of coupled ordinary differential equations is replaced by a nonlocal evolution equation on the unit interval. In this work, we adopt this methodology to explore the continuum limit of random walks, a popular model for diffusion on graphs. We focus on two classes of processes on dense weighted graph, in discrete and in continuous time, whose dynamics are encoded in the transition matrix and the random-walk Laplacian. We also show that previous works on the discrete heat equation, associated to the combinatorial Laplacian, fall within the scope of our approach. Finally, we apply the spectral theory of operators to 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

As very large graphs become increasingly common in scientific research and real-world applications, a range of algorithms and computational problems face scalability issues. An elegant solution can be found by considering the continuum limit of graphs, defined when the number of nodes goes to infinity. This approach has been used, for instance, for network identification techniques [14] and spectral clustering methods [17]. Another application focuses on different classes of diffusion-based problems in networked systems [25, 27]. In that case, as we will see, the dynamics on a large graph can be approximated by the dynamics on its continuous limit, which translates into integro-differential equations mathematically.

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 a random walk. Different types of random walk models can be defined, depending on the stochastic processes determining the timings at which events
take place and the sequence of visited nodes. Random walks on graphs are useful in many ways. They can for example identify clusters of well-connected nodes, also known as communities [31, 10], or measure the relative importance or centrality of the nodes in a networked system [7, 21, 20]. They are also a paradigm for various diffusive and spreading processes on graphs [24, 1].

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 randomly 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 term 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. The difference between the last two processes is clear when observing the matrices controlling their dynamics, the random-walk and the combinatorial Laplacian respectively [28].

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 [25, 26, 27]. 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 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 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 [22], which does not rely on the assumption that the data generating the graphs is sampled from a distribution on a manifold [16, 30]. Most of our effort is directed toward the convergence of the discrete (in space) problem to a continuous problem in some appropriate setting. The problem on the continuum then falls in the realm of nonlocal evolution equations. More precisely, it is a volume-constrained diffusion problem [9], and its analysis is voluntarily limited to some immediate consequences of spectral theory applied to our operators.

The paper is organized as follows. Section 2 contains the basic notions about graphs and 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 notations for the various concepts from graph theory and recall the main definitions, following closely the standard definitions found in [32]. For the sake of self-consistency we then introduce the necessary notions about graph limits and graphons [22].
2.1 Graphs

Let a graph $G = (V, E)$ be a graph where $V$ is a finite set of vertices, and $E \subseteq V \times V$ is the set of edges. The vertices, also called nodes, are labeled by $\{1, 2, \ldots\}$. Two vertices $v$ and $w$ form an edge $[v, w]$ whenever $v \sim w$ where $\sim$ is an adjacency relation. We assume this relation to be symmetric, $v \sim w \iff w \sim v$, such that the resulting graph is undirected. Let $|V|$ (resp. $|E|$) denote the number of vertices (resp. edges). The density $\rho$ 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 $|E| = O\left(\frac{|V|^2}{2}\right)$, and sparse otherwise.

Each edge may be attributed a weight, making the otherwise unweighted graph into a weighted one. The number of neighbors of node $v$ is denoted by $\deg(v)$. In weighted graphs, one also defines $\str(v)$, the weighted degree or strength of node $v$, as the sum of the weights of all edges $[v, w]$ for which $w \sim v$.

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 of $n \times n$ matrices. The adjacency matrix $A \in \mathbb{M}_n$ of a finite 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, where the ones indicate existing edges. We denote $D = \text{diag}(\str(v_1), \ldots, \str(v_n))$ the diagonal matrix of the strengths (or degrees in unweighted graphs).

2.2 Graphons

Recent research [23, 6, 3, 22] 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 $W : [0, 1]^2 \to [0, 1]$. Therefore, the space of 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 which we introduce hereafter in its graphon version. Let us review the main concepts, as exposed also in [22, 25, 15]. The cut norm for graphons is given by

$$\|W\|_\Box = \sup_{S,T \subseteq [0,1]} \int \int_{S \times T} W(x,y) \, dx \, dy,$$

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:

$$\|W\|_\Box \leq \|W\|_1 \leq \|W\|_2 \leq \|W\|_\infty \leq 1.$$

Note that this choice of domain and range is somehow restrictive by comparison with other works where for instance $W : [0,1]^2 \to \mathbb{R}$. However, we will work with the standard definition because it achieves the desired degree of generality.

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].

3
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 \( W^\phi \) 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

\[
\delta_\square(U, W) = \inf_{\phi \in \mathcal{L}} \| U^\phi - W \|_\square
\]

where \( \mathcal{L} \) 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 \( \mathcal{W} \) the space of graphons after this identification.

It holds that the metric space \((\mathcal{W}, \delta_\square)\) 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 \((\mathcal{W}, \| \cdot \|_2)\) converge in the \( L^2 \) metric, and thus also in the \( \delta_2 \) and \( \delta_\square \) metrics, the limit being the same.

### 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 \( P = \{ P_1, \ldots, P_n \} \) be a uniform partition of \([0, 1]\), where \( P_i = \left( \frac{i-1}{n}, \frac{i}{n} \right) \) for \( i = 1, \ldots, n-1 \), and \( P_n = \left[ \frac{n-1}{n}, 1 \right] \). Then let \( \eta : \mathcal{M}_n \to \mathcal{W} \) be a mapping such that

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

where \( \chi_S \) 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 \( G \). Similarly, \( \eta \) maps vectors \( u = (u_1, \ldots, u_n) \) to piecewise constant functions on \([0, 1]\), so that

\[
\eta(u)(x) = \sum_{i=1}^{n} u_i \chi_{P_i}(x).
\]

On the other hand, graphons can be considered as deterministic or (exchangeable [12]) random graph models. In this work we present the deterministic version. Let \( W \in \mathcal{W} \) 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:

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

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

A second approach to generate a graph from a given graphon \( W \in \mathcal{W} \), is to define

\[
A_{ij} = W \left( \frac{i}{n}, \frac{j}{n} \right), \quad i, j = 1, \ldots, n,
\]
in a way that is reminiscent of \( W \)-random graphs \[23\]. Let us denote \( W_{[n]} \) the corresponding graph. Observe that \( \eta \left( W_{[n]} \right) \rightarrow W \) point-wise at every point of continuity of \( W \) \[23\].

### 2.4 Graphons as kernels of operators

Every graphon \( W \in W \) can be considered as a kernel, allowing to formally define an integral operator \( \mathcal{W} \) on functional spaces on \([0, 1]\) through

\[
\mathcal{W}f(x) = \int_0^1 W(x, y)f(y)dy
\]

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

\[
\mathcal{UW}f(x) = \int_0^1 U \circ W(x, y)f(y)dy,
\]

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

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

Observe that in general, \( U \circ W \) is not a symmetric function. We denote \( W^{on} \) 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 \( W^n \). It follows from \[9\] that

\[
W^{on}(x, y) = \int W(x, z_1)W(z_1, z_2)\ldots W(z_{n-1}, y)dz_1dz_2\ldots dz_{n-1}.
\]

### 3 Random walks and their continuum limit

The aim of this section is twofold. First, we shortly introduce the discrete- and continuous-time random walks on a connected\(^3\) graph \( G = (V, E) \) with \( n \) nodes. Secondly, we formulate their continuum version in the limit of an infinite number of nodes.

### 3.1 Random walks in discrete and continuous time

Arguably the simplest case, the discrete-time random walk on a connected graph is 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

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

Let \( p(\ell) = (p_1(\ell), \ldots, p_n(\ell)) \) 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 \( \ell \) steps. Then

\[
p(\ell + 1) = p(\ell)T,
\]

\[\text{If the graph is not connected, the random walk is considered independently on each connected component, that is, on each connected subgraph that is connected to no other additional node of the original graph.}\]
where \( T = D^{-1}A \). It follows that for any \( \ell \in \mathbb{N} \), \( p(\ell) = p(0)T^\ell \).

In the continuous-time (node-centric) version, 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 concentrate on Poissonian walks, for which the waiting-time follows a memoryless exponential distribution \( \psi_v(t) = \mu_v \exp(-\mu_v t) \) with rate \( \mu_v \) \((t \geq 0)\). The master equation for \( u_i(t) \), the probability to find the walker on node \( i \) at time \( t \), reads

\[
\dot{u}_i(t) = \sum_{j=1}^{N} \mu_j u_j \frac{1}{\text{str}(v_j)} A_{ji} - \mu_i u_i, \quad i = 1, \ldots, n. \tag{13}
\]

Assume that in (13) the rate \( \mu_j \) on the nodes is the same for all nodes, \( \mu_j = \kappa > 0 \) for all \( j \). Then \( \kappa \) sets the timescale, and after a rescaling of time, \( t \mapsto \kappa t \), the master equation (13) rewrites

\[
\dot{u}_i(t) = \sum_{j=1}^{N} u_j \frac{1}{\text{str}(v_j)} A_{ji} - u_i, \tag{14}
\]

or under matrix form, \( \dot{\mathbf{u}} = \mathbf{u}(D^{-1}A - I) \), where \( \mathbf{u}(t) = (u_1(t), \ldots, u_n(t)) \) is a row vector. The matrix \( L^w = D^{-1}A - I \) 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 \((\text{str}(v_1), \ldots, \text{str}(v_n))\).

In the edge-centric model\(^4\), the rate of the exponential distribution is proportional to the degree of the node, \( \mu_j = \kappa \text{str}(v_j) \), allowing a constant rate of jump across all edges of the graph. Hence in matrix form, equation (13) rewrites

\[
\dot{\mathbf{u}} = \kappa \mathbf{u}(A - D) \tag{15}
\]

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 \( \kappa \deg(v_j) \rightarrow \infty \) if \( \deg(v_j) \rightarrow \infty \), 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 \( \kappa \) becomes dependent on \( n \), say \( \kappa_n = \frac{\kappa}{n} \), 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 equation (15) in \[25\].

In contrast with the edge-centric model, no normalization of the rate parameter \( \kappa \) 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 (14) and the uniform

\[\text{The fact that this walk can be formulated in terms of edges dynamics, where the walker passively follows the activations of the edges explains the alternative designation of “fluid model” [24]. It is the graph version of the heat equation on a continuum.}\]
partition $\mathcal{P} = \{P_1, \ldots, P_n\}$ of $[0, 1]$, with $u(\cdot, t) := \eta(u(t))$ an associated step function on the interval. Let the degree function $k_\eta$ of the step graphon $\eta(G)$ be defined by

$$k_\eta(x) = \int_0^1 \eta(G)(x,y)dy \text{ for } x \in [0, 1]. \quad (16)$$

This degree function is actually the normalized strength (or also degree, when the graph is unweighted) of the nodes in $G$:

$$\text{str}(v_i) = n \sum_{j=1}^n A_{ij} dy = n \sum_{j=1}^n \int_{P_j} \eta(G)(x,y)dy = nk_\eta(x) \quad (17)$$

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

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

$$= \sum_{j=1}^n n \int_{P_j} \frac{A_{ij}}{nk_\eta(y)} u(y,t)dy = \int_0^1 \frac{\eta(G)(x,y)}{k_\eta(y)} u(y,t)dy,$ \quad (18)$$

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

$$\frac{\partial}{\partial t}u(x,t) = \int_0^1 \frac{\eta(G)(x,y)}{k_\eta(y)} u(x,t)dy - u(x,t). \quad (19)$$

The goal of this work is to prove convergence in the appropriate norm of the solution of (19) to the solution of the evolution equation on the continuum

$$\frac{\partial}{\partial t}w(x,t) = \int_0^1 \frac{W(x,y)}{k(y)} w(y,t)dy - w(x,t), \quad (20)$$

where $W$ is the limit graphon of $\eta(G)$ in the $L^2$ metric, and $k$ its degree function defined by $k(x) = \int_0^1 W(x,y)dy$ for every $x \in [0, 1]$.

Observe that similarly, a discrete equation of the form (14) is obtained starting from (20), when the graph is $W/\mathcal{P}$ or $W_{\lfloor n \rfloor}$.

### 4 Well-posedness of the continuum initial value problem

Before we prove the above-mentioned convergence, let us determine whether (20), together with initial condition $w(x,0) = g(x)$, defines a well-posed initial-value problem (IVP).

#### 4.1 Connectedness of the graphon and integrability of $\frac{W}{k}$

Care will be taken first regarding how connectedness in the graph translates to graphons. The following definition follows from [18][22].
Definition 1. A graphon $W$ is connected if

$$\int_S W(x,y) \, dx dy > 0$$

for every $S \in \mathcal{M}[0,1]$ 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 [18]. Indeed, one could always make all the (otherwise disconnected) graphs of the sequence connected by 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 ([18], 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 1. Let $W$ be a connected graphon, then $k \geq 0$.

Proof. For every $x \in [0,1]$, define the neighborhood of $x$ in $W$ as

$$N_x = \{y \in [0,1] : W(x,y) > 0\}.$$ 

Since $W$ is connected, $\mu(N_x) > 0$ for $\mu$-almost every $x$ ([18], lemma 5.1) and therefore,

$$k(x) = \int_{N_x} W(x,y) \, dy > 0 \text{ for } \mu\text{-a.e. } x.$$  \hfill (21)

Remark 1. 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 $0 < c \leq k$ on $[0,1]$. Take for instance $W(x,y) = x^m y^m$ with $m > 0$, for which $k(x) = \frac{x^m}{m+1}$.

That $k$ can be arbitrarily small influences the integrability of the kernel $K(x,y) := \frac{W(x,y)}{k(y)}$ in [20], as the following remark explains.

Remark 2. 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

$$\|K\|_p^p = \int_{x^\alpha + y^\alpha \leq 1} (1 - y^\alpha)^{-\frac{p}{\alpha}} \, dx dy = \int_0^1 (1 - y^\alpha)^{\frac{1-\alpha}{\alpha}} \, dy$$  \hfill (22)

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 1. 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 [11].
4.2 The IVP with functions in $L^2[0,1]$

Resting on the operator in the right-hand side of (20), we come to the following definition.

**Definition 2.** Let $W \in \mathcal{W}$ be a connected graphon that verifies assumption 1. Let the random-walk Laplacian operator $L_{rw} : L^2[0,1] \rightarrow L^2[0,1]$ be defined by

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

(23)

By definition, $W$ is bounded on $[0,1]^2$ and following our hypothesis, $\frac{1}{k(x)}$ is bounded on $[0,1]$. Therefore, $K(x,y) = \frac{W(x,y)}{k(y)}$ is a Hilbert-Schmidt kernel and $K : L^2[0,1] \rightarrow L^2[0,1]$ defined by

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

(24)

is a compact Hilbert-Schmidt operator. Following definition 3, the continuum IVP has the form

$$\frac{\partial}{\partial t} w(x,t) = L_{rw} w(x,t)$$

(25a)

$$w(x,0) = g(x) \in L^2[0,1]$$

(25b)

**Theorem 1.** Let $W \in \mathcal{W}$ be connected and satisfying assumption 1. Then there exists a unique classical solution to the initial-value problem (25).

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

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

(26)

Proposition 6.2 in [13] allows to conclude. \qed
Remark 3. (Classical solution) By definition of classical solution of the abstract Cauchy problem (25), the orbit maps 
\[ t \in \mathbb{R}^+ \mapsto w(t,x) \in L^2[0,1] \]
are continuously differentiable. The forthcoming convergence results of section 3 are established in norm
\[ \|w\|_{C([0,T],L^2[0,1])} = \sup_{t \in [0,T]} \|w(t,\cdot)\|_{L^2[0,1]} \] (27)
defined for any positive real \( T \).

Remark 4. The asymptotic steady state \( w_\infty \) of (25) follows from \( L^* w_\infty = 0 \) and is given by \( w_\infty = k \).

4.3 Positivity

The continuum IVP (25) would lose physical relevance if its solution were to lose the possible 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 \( M_g : L^p[0,1] \to L^p[0,1] \) denote the multiplication operator defined by
\[ M_g f(x) = g(x) f(x). \] (28)

Proposition 2. Let \( W \) be a connected graphon satisfying assumption 2 and let \( w(\cdot,0) = g \geq 0 \) be the initial condition of IVP (25). 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 \( \tilde{L} = M_1 \mathcal{L} \mathcal{L}^* \mathcal{M}_k \) by a direct calculation
\[ \tilde{L} f(x) = \frac{1}{k(x)} \int_0^1 W(x,y) f(y,t) dy - f(x,t), \quad \forall f \in L^2[0,1]. \] (29)

Further let \( u = M_1^\frac{1}{k} w \) with \( w(x,t) \) the solution of (25) such that
\[ \frac{\partial}{\partial t} u = M_1^\frac{1}{k} \frac{\partial}{\partial t} w = M_1^\frac{1}{k} \mathcal{L}^* \mathcal{L}^w w = M_1^\frac{1}{k} \mathcal{L}^* \mathcal{L}^w \mathcal{M}_k u = \tilde{L} u. \]

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 \( \tilde{L} v = \tilde{L} u \), and hence
\[ \frac{\partial}{\partial t} v - \tilde{L} v = \frac{\partial}{\partial t} u + \epsilon - \tilde{L} u = \epsilon. \]

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
\[ \tilde{L} v(a,\tau) = \frac{1}{k(a)} \int_0^1 W(a,y) v(y,\tau) dy - v(a,\tau) \geq \frac{1}{k(a)} \int_0^1 W(a,y) v(a,\tau) dy - v(a,\tau) = 0. \]
Hence, \( \frac{\partial}{\partial t} v(a, \tau) = \tilde{L} v(a, \tau) + \epsilon = \epsilon > 0 \) which is in contradiction with the assumption of \( v \) attaining its minimum in \((a, \tau)\) with \( \tau > 0 \), so \( \tau = 0 \). We have thus proved \( v(x, t) \geq v(a, 0) \), so that
\[
    u(x, t) + \epsilon t = v(x, t) \geq v(a, 0) = \frac{g(a)}{k(a)} \geq 0.
\]

Since \( \epsilon \) is arbitrary, this allows to conclude.

4.4 The IVP with probability density functions

Let us observe that when \( w(\cdot, t) \) in (25) is a probability density function, it is natural to consider \( w(\cdot, t) \in L^1[0, 1] \), and one may define \( L^{rw} \) as a mapping \( L^1[0, 1] \to L^1[0, 1] \). Indeed, as in (24) let us still write \( K \) the integral part of \( L^{rw} \) defined on \( L^1[0, 1] \). By Fubini-Tonelli, the operator norm \( \|K\|_{1,1} := \|K\|_{L^1[0,1] \to L^1[0,1]} \) satisfies
\[
    \|K\|_{1,1} \leq \sup_{\|f\|_1=1} \int_{[0,1]^2} |K(x, y)f(y)|\,dxdy = \sup_{\|f\|_1=1} \int_{[0,1]} |f(y)|\,dy = 1. \tag{30}
\]

This, combined with the fact that \( \|Kf\|_1 = 1 \) if \( f = 1 \), shows that \( \|K\|_{1,1} = 1 \), and so even without assumption \( 1 \), \( L^{rw} \) is a bounded mapping of \( L^1[0, 1] \) into itself. Additionally, theorem \( 4 \) 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 \( 1 \).

The only significant change in the proof of proposition \( 2 \) would be to use the auxiliary operator \( L^{rw} M_k \) instead of \( L = M_k L^{rw} M_k \). When \( w(\cdot, 0) \geq 0 \) we further have conservation of the \( L^1 \) norm:
\[
    \frac{\partial}{\partial t} \|w(\cdot, t)\|_1 = \frac{\partial}{\partial t} \int_0^1 |w(x, t)|\,dx = \frac{\partial}{\partial t} \int_0^1 w(x, t)\,dx = 0. \tag{31}
\]

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 \( 1 \). This allows to define \( L^{rw} \) 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 \([25, 27]\).

5 Convergence on dense graphs

This section is divided in three parts. In loose terms, the first two show that the solution of the discretized problem on \( W/P \) or \( W_{[n]} \) converges to that of the continuum IVP, in the norm of \( (27) \). 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 \( W/P \)

Let us start with two simple lemmas.
Lemma 1. Let $A_\eta : L^2[0,1] \to L^2[0,1]$ be an integral operator with bounded kernel $A_\eta$. Assume that $A_\eta$ 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 $f_\eta$ by

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

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

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

$$A_\eta^\ell f = A_\eta^{\ell-1} A_\eta f = A_\eta^{\ell-1} A_\eta f_\eta = A_\eta^\ell f_\eta.$$

Lemma 2. Let $A, B : L^2[0,1] \to L^2[0,1]$ 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 $\ell \in \mathbb{N}_0$

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

Proof. Using the Minkowski inequality, we have

$$\|A^\ell f - B^\ell f\|_2 = \|A^{\ell-1} Af - B^{\ell-1} Bf\|_2$$

$$= \|A^{\ell-1} Af - A^{\ell-1} B f + A^{\ell-1} B f - B^{\ell-1} Bf\|_2$$

$$\leq \|A^{\ell-1} (Af - Bf)\|_2 + \|(A^{\ell-1} - B^{\ell-1})Bf\|_2.$$

(32)

Now $A^{\ell-1}$ is a Hilbert-Schmidt integral operator with kernel $A^{\ell-1}$, as defined by [10]. For such operator, as a product of the Cauchy-Schwarz inequality it is known about the operator norm $\|\cdot\|$ that $\|A^{\ell-1}\| \leq \|A^{\ell-1}\|_2$, or equivalently

$$\|A^{\ell-1} f\|_2 \leq \|A^{\ell-1}\|_2 \|f\|_2.$$  

(33)

The first term in the right hand side of (32) therefore satisfies

$$\|A^{\ell-1} (Af - Bf)\|_2 \leq \|A^{\ell-1}\|_2 \|Af - Bf\|_2 \leq \beta^{\ell-1} \|Af - Bf\|_2$$

(34)

where we use $\|A^{\ell-1}\|_2 \leq \|A\|_2^{\ell-1}$ ([15], lemma 6) and $A(x,y) \leq \beta$ for all $0 \leq x, y \leq 1$ to obtain the last inequality. Using again (33) with $\ell = 2$, we also have $\|Af - Bf\|_2 \leq \|A - B\|_2 \|f\|_2$ which, together with (32) and (34) leads to the conclusion. 

Now we are in a place to formulate the convergence results. The continuous formulation of the discrete problem associated to (25) on the quotient graph reads\footnote{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, $u^{(n)}(x,t)$.}

$$\frac{\partial}{\partial t} u(x,t) = \mathcal{L}^w u(x,t)$$

(35a)

$$u(x,0) = g_{\square}(x)$$

(35b)
where the random walk Laplacian operator on $W/P$ satisfies
\[
L_{rw}f(x) = \int_{0}^{1} \eta(W/P)(x,y) f(y) dy - f(x), \quad \forall f \in L^2[0,1],
\]
and the initial condition is averaged on each cell of the partition as
\[
g_{\square}(x) = \frac{1}{n} \sum_{i=1}^{n} \int_{P_i} g(y) dy \chi_{P_i}(x), \quad \forall x \in [0,1].
\]
Based on the following proposition, operator $L_{rw}$ is well-defined.

**Proposition 3.** Let $W$ be a connected graphon satisfying assumption then the strength of every node of the quotient graph determined by the partition $P = \{P_1, \ldots, P_n\}$ of $[0,1]$ is positive.

**Proof.** The strength of the $i$-th node $v_i$, $i = 1, \ldots, n$, is given by $\text{str}(v_i) = \frac{1}{n} \sum_{j=1}^{n} A_{ij} \chi_{P_j}(x)$, for every $x \in P_i$.

We have
\[
k_{\square}(x) = \int_{0}^{1} \sum_{j=1}^{n} A_{ij} \chi_{P_j}(y) dy = \frac{1}{n} \sum_{j=1}^{n} A_{ij}, \quad \forall x \in P_i,
\]
where $A_{ij}$ was defined by (4). Hence,
\[
k_{\square}(x) = \frac{1}{n} \sum_{j=1}^{n} \int_{P_j} W(x',y') dy' dx' \\
= \frac{1}{n} \int_{P_i} \int_{0}^{1} W(x',y') dy' dx' \\
= \frac{1}{n} \int_{P_i} k(x') dx' \geq c,
\]
where $c > 0$ is the constant from assumption \[1\].

**Remark 5.** It follows that the finite-dimensional IVP (35) on the quotient graph has a unique solution given by $e^{tL_{\square}}g_{\square}$.

**Theorem 2** (Convergence with $W/P$). Let $W$ be a connected graphon satisfying assumption then and let $w(x,t)$ be the solution of IVP (25). Further let $u(x,t)$ be the solution of the associated discrete problem (35). Then for all $t \in \mathbb{R}^+$ it holds that
\[
\|u(\cdot,t) - w(\cdot,t)\|_2 \to 0 \quad \text{as} \quad n \to \infty.
\]

**Proof.** Using remark [5] by the Minkowski inequality we have
\[
\|u(\cdot,t) - w(\cdot,t)\|_2 = \|e^{tL_{\square}}g_{\square} - e^{tL}g\|_2 \\
= \left\| \sum_{k=0}^{\infty} \frac{t^k}{k!} L_{\square}^k g_{\square} - \sum_{k=0}^{\infty} \frac{t^k}{k!} L^k g \right\|_2 \\
\leq \|g_{\square} - g\|_2 + \sum_{k=1}^{\infty} \frac{t^k}{k!} \|L_{\square}^k g_{\square} - L^k g\|_2. \quad (39)
\]
Let us write $L^w = K - I$ where $K$ is the operator previously defined in [24] and $I$ is the identity operator. We have a similar decomposition $L^w = K - I$ for the Laplacian of the step graphon. For $k \geq 1$ and $0 \leq m \leq k$ let us write $\alpha_m = (-1)^m \binom{k}{m}$, and consider $(\star)$ in the right-hand side of (39). Using Newton’s binomial theorem we have

$$
\|L^k g - L^k q\|_2 = \|(K - I)^k g - (K^k - I)q\|_2
$$

$$
= \left\| \sum_{m=0}^{k} \alpha_m k^m g - \sum_{m=0}^{k} k^{k-m} g \right\|_2
$$

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

with $|\alpha_m| = \binom{k}{m}$ and using lemma 1

$$
\leq \sum_{m=0}^{k-1} \binom{k}{m} \left\| (K^k - K^{k-m}) g \right\|_2 + \|g - g\|_2.
$$

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

$$
\|K\|_2 = \left\| \frac{\eta(W/P)}{k} \right\|_2 = \|\eta(W/P)\|_2 \frac{1}{k} \leq \frac{1}{c} =: \beta_k.
$$

where $K$ denotes the integral kernel of $K$. For $\ell \in \mathbb{N}_0$, let us define $E_\ell := K^\ell - K^\ell$ and $E_\ell := K^\ell - K^\ell$. Then, applying lemma 2 successively $\ell - 1$ times to $(\star)$ in (40) with $\ell = k - m$, we obtain

$$
\|E_\ell g\|_2 \leq \beta^{\ell-1}_k \|\ell_1\|_2 \|g\|_2 + \|E_{\ell-1} K g\|_2
$$

$$
\leq (\beta^{\ell-1}_k + \beta^{\ell-2}_k) \|\ell_1\|_2 \|g\|_2 + \|E_{\ell-2} K^2 g\|_2
$$

$$
\vdots
$$

$$
\leq \left( \sum_{j=1}^{\ell-1} \beta^{\ell-j}_k \right) \|\ell_1\|_2 \|g\|_2 + \|E_1 K^{\ell-1} g\|_2.
$$

and since $E_\ell$ is the kernel of $E_\ell$ if $\ell = 1$,

$$
\leq \left( \sum_{j=1}^{\ell-1} \beta^{\ell-j}_k \right) \|\ell_1\|_2 \|g\|_2 + \|E_1\|_2 \|K^{\ell-1} g\|_2
$$

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

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

$$
\leq \ell \beta^{\ell-1}_k \|E_1\|_2 \|g\|_2.
$$

(42)
where the last inequality stems from \( \beta := \max \{ \| K \|_2, \beta_\square \} \geq 1 \). Combining (40) and (42) yields
\[
\| L^k_{\square} g - L^k g \|_2 \leq \sum_{m=0}^{k-1} \left( \frac{k}{m} \right)(k-m)\beta^{k-m-1}\| K_{\square} - K \|_2 g_2 + \| (g_{\square} - g) \|_2 \\
\leq \beta^{k-1} \sum_{m=0}^{k-1} \left( \frac{k}{m} \right)(k-m)\| K_{\square} - K \|_2 g_2 + \| (g_{\square} - g) \|_2.
\]
From (39) and (43) we obtain
\[
\| u(\cdot, t) - w(\cdot, t) \|_2 \leq \| g_{\square} - g \|_2 + \sum_{k=1}^{\infty} \frac{t^k}{k!} \| (g_{\square} - g) \|_2 \\
+ \| K_{\square} - K \|_2 g_2 \sum_{k=1}^{\infty} \frac{t^k}{k!} \beta^{k-1} \sum_{m=0}^{k-1} \left( \frac{k}{m} \right)(k-m) \\
= \| g_{\square} - g \|_2 e^t + \| K_{\square} - K \|_2 g_2 \sum_{k=1}^{\infty} \frac{t^k}{k!} \beta^{k-1} \sum_{m=0}^{k-1} \left( \frac{k}{m} \right)(k-m) \\
\qquad + \sum_{m=0}^{k-1} \left( \frac{k}{m} \right)(k-m) = k2^{k-1},
\]
and with \( \sum_{m=0}^{k-1} \left( \frac{k}{m} \right)(k-m) = k2^{k-1} \),
\[
= \| g_{\square} - g \|_2 e^t + \| K_{\square} - K \|_2 g_2 \sum_{k=1}^{\infty} \frac{t^k}{k!} k(2\beta)^{k-1} \\
\leq \| g_{\square} - g \|_2 e^t + \| K_{\square} - K \|_2 g_2 e^t \sum_{k=1}^{\infty} \frac{(2\beta t)^{k-1}}{(k-1)!} \\
\leq \| g_{\square} - g \|_2 e^t + \| K_{\square} - K \|_2 g_2 e^{2\beta t}. \tag{44}
\]
By the Lebesgue differentiation theorem, \( g_{\square} \to g \) pointwise for almost every \( x \in [0, 1] \) as \( n \to \infty \), so that
\[
\| g_{\square} - g \|_2 \underset{n \to \infty}{\to} 0 \tag{45}
\]
by dominated convergence \cite{25}. Let us consider \((**)\) in (44):
\[
\| K_{\square} - K \|_2^2 = \int_{[0,1]^2} \left( \frac{\eta(W/P)(x, y)}{k_{\square}(x, y)} - \frac{W(x, y)}{k(y)} \right)^2 \, dx \, dy \\
\leq \text{ess sup}_{y \in [0, 1]} \frac{1}{k_{\square}(y)k^2(y)} \int_{[0,1]^2} \left( \eta(W/P)(x, y)k(y) - W(x, y)k_{\square}(y) \right)^2 \, dx \, dy \\
\leq \beta^2 \int_{[0,1]^2} \left( \eta(W/P)(x, y)k(y) - k_{\square}(y) \right)^2 \, dx \, dy \\
+ \beta^2 \int_{[0,1]^2} \left( W(x, y) - \eta(W/P)(x, y)k_{\square}(y) \right)^2 \, dx \, dy.
\]
and because $\|\eta(W/P)\|_2 \leq 1$ and $\|k\|_2 \leq 1$,

$$\leq \beta^2 \left( \|k - k\|_2^2 + \|W - \eta(W/P)\|_2^2 \right),$$

(46)

By the Cauchy-Schwarz inequality,

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

$$\leq \int_0^1 \int_0^1 (W(y,z) - \eta(W/P)(y,z))^2 \, dz \, dy$$

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

which together with (46) yields

$$\|K - K\|_2^2 \leq 2\beta^2 \|W - \eta(W/P)\|_2^2.$$  

(47)

By the same argument leading to (45), we have $\|W - \eta(W/P)\|_2 \to 0$ as $n \to \infty$ which with (47) implies

$$\|K - K\|_2 \to 0 \quad n \to 0.$$  

(48)

Combining (44), (45) and (48) allows to conclude.

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 $W/P$, 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 (25) on the sampled graph $W[n]$ reads

$$\frac{\partial}{\partial t} u(x,t) = \mathcal{L}^w_{[n]} u(x,t)$$

$$u(x,0) = g(x)$$

(49a)

(49b)

where the random walk Laplacian operator on $W[n]$ satisfies

$$\mathcal{L}^w_{[n]} f(x) = \int_0^1 \eta(W[n]) \left( \frac{x}{k[n]} \right) f(y) dy - f(x),$$

$$\forall f \in L^2[0,1],$$

(50)

and the initial condition is again averaged on each cell of the partition as in (37). 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 3** (Convergence with $W[n]$). Let $W$ be a connected, almost everywhere continuous graphon satisfying assumption 7 and let $w(x,t)$ be the solution of IVP (25). Further let $u(x,t)$ be the solution of the associated discrete problem (49). Then for all $t \in \mathbb{R}^+$ it holds that

$$\|u(\cdot,t) - w(\cdot,t)\|_2 \to 0 \quad as \quad n \to \infty.$$
Remark 6. The initial condition could have been sampled in a similar fashion as the graphon, to yield the step function
\[ g[n] = \sum_{i=1}^{n} g\left( \frac{i}{n} \right) \chi_{P_i}. \] (51)
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 \( \text{3} \) would only require minor changes, which are similar to those discussed next in the new context of section \( \text{5.3} \).

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 \( \eta \)

\[ \frac{\partial}{\partial t} u(x,t) = L_n^{rw} u(x,t) \] (52a)
\[ u(x,0) = g_n(x) \in L^2[0,1], \] (52b)

where the random walk Laplacian operator \( L_n^{rw} \) satisfies

\[ L_n^{rw} f(x) = \int_0^1 \frac{\eta(G_n)(x,y)}{k_n(y)} f(y)dy - f(x), \quad \forall f \in L^2[0,1]. \] (53)

Similarly as before, we write \( L_n^{rw} = K_n - I \).

Theorem 4 (Convergence with \( (G_n) \)). Let \( (G_n) \) be a sequence of connected graphs that converges to a connected graphon \( W \) satisfying assumption \( 4 \). Let \( w(x,t) \) be the solution of the IVP \( (25) \) 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 \( (52) \), and assume that \( \| g_n - g \|_2 \to 0 \) as \( n \to \infty \). Then for all \( t \in \mathbb{R}^+ \) it holds that

\[ \| u(\cdot,t) - w(\cdot,t) \|_2 \to 0 \quad \text{as} \quad n \to \infty. \]

Proof. The proof follows the same steps as for theorem \( \text{2} \). However, using lemma \( \text{4} \) to obtain \( (40) \) 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 \( \text{4} \) we write

\[ \| K_n^{\ell-m} g_n - K_n^{k-m} g \|_2 = \| K_n^{\ell-m} g_n - K_n^{k-m} g + K_n^{\ell-m} g - K_n^{k-m} g \|_2 \]
\[ \leq \| K_n^{\ell-m} (g_n - g) \|_2 + \| (K_n^\ell - K_n^k) g \|_2, \]

with the first term in the right-hand side newly present. Following the same steps leading to \( (40) \), we
obtain
\[
\|\mathcal{L}^k_{n} g_n - \mathcal{L}^k g\|_2 \leq \sum_{m=0}^{k-1} \binom{k}{m} \|\mathcal{K}^{\ell-m}_n (g_n - g)\|_2 \\
+ \sum_{m=0}^{k-1} \binom{k}{m} \|\mathcal{K}^{k-m}_n - \mathcal{K}^{k-m}_m\|_2 + \|(g_n - g)\|_2, 
\]
where again the first term right of the inequality is new. In fashion similar to the proof of theorem 2 with \(\beta := \max\{\|K\|_2, \frac{1}{k_n}\}\) we have
\[
\|u(\cdot, t) - w(\cdot, t)\|_2 \leq \sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_{m=0}^{k-1} \binom{k}{m} \|K^{\ell-m}_n (g_n - g)\|_2 \\
+ \|g_n - g\|_2 e^t + \|K_n - K\|_2 \|g\|_2 e^{2\beta t}. 
\]
Using \(\|K^{k-m}\| \leq \beta^{k-m} \leq \beta^k\) and \(\sum_{m=0}^{k-1} \binom{k}{m} \leq 2^k\), we have
\[
\sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_{m=0}^{k-1} \binom{k}{m} \|K^{\ell-m}_n (g_n - g)\|_2 \leq \sum_{k=1}^{\infty} \frac{t^k}{k!} 2^k \beta^k \|g_n - g\|_2 = \|g_n - g\|_2 e^{2\beta t}, 
\]
leading to
\[
\|u(\cdot, t) - w(\cdot, t)\|_2 \leq \|g_n - g\|_2 (e^t + e^{2\beta t}) + \|K_n - K\|_2 \|g\|_2 e^{2\beta t}. 
\]

6 Relaxation time through spectral analysis

The rate at which the system evolves towards the asymptotic state \(w_\infty\) starting from any initial condition, is known as relaxation time. In the continuum limit of the node-centric walk, it is determined by the spectral properties of \(K\), in a way reminiscent of random walks on finite graphs. For the node-centric continuous-time walk, we will show now that this rate is exponential. Under assumption \([1]\) let us define integral operator \(S: L^2[0, 1] \to L^2[0, 1]\) by \(S = \mathcal{M}_W K \mathcal{M}_\sqrt{\mathcal{K}}\), namely
\[
Sf(x) = \int_0^1 \frac{W(x, y)}{\sqrt{k(x)} \sqrt{k(y)}} f(y) dy, \quad \forall f \in L^2[0, 1]. 
\]
Under assumption \([1]\) the kernel is square-integrable and symmetric. Hence, \(S\) is a compact, self-adjoint Hilbert-Schmidt integral operator and the Hilbert-Schmidt theorem applies (\([29]\), theorem VI.6). Therefore, there exists an orthonormal basis of eigenfunctions \(\{\phi_m\}\) with associated eigenvalues \(\lambda_m\), so that operator \(S\) has the canonical form
\[
S = \sum_{m=1}^{\infty} \lambda_m (\phi_m, \cdot) \phi_m. 
\]
Also note that for \( \ell \in \mathbb{N} \), \( S^\ell \) has eigenfunctions \( \phi_m^\ell \) and eigenvalues \( \lambda_m^\ell \), and that \((L_{rw})^\ell = M \sqrt{k} (S - I)^\ell \frac{1}{\sqrt{k}}\). Combined with (55) this yields

\[
e^{L_{rw} t} = M \sqrt{k} \sum_{\ell=0}^{\infty} \frac{t^\ell}{\ell!} \left( \sum_{m=1}^{\infty} \lambda_m (\phi_m, \cdot) \phi_m - I \right)^\ell \frac{1}{\sqrt{k}}
\]

\[
= M \sqrt{k} \sum_{\ell=0}^{\infty} \frac{t^\ell}{\ell!} \left( \sum_{m=1}^{\infty} (\lambda_m^\ell - 1) (\phi_m, \cdot) \phi_m \right) \frac{1}{\sqrt{k}}
\]

\[
= M \sqrt{k} \sum_{m=1}^{\infty} \sum_{\ell=0}^{\infty} \frac{t^\ell}{\ell!} \theta_m^\ell (\phi_m, \cdot) \phi_m \frac{1}{\sqrt{k}}
\]

\[
= \sum_{m=1}^{\infty} e^{\theta_m t} \left( \phi_m \sqrt{k} \right) \phi_m
\]

(56)

with \( \theta_m = \lambda_m - 1 \) the eigenvalues of \( S - I \). By letting \( \psi_m = \frac{\phi_m}{\sqrt{k}} \) and \( \zeta_m = \sqrt{k} \phi_m \), the solution of IVP (25) reads

\[
w(x,t) = \sum_{m=1}^{\infty} e^{\theta_m t} (\psi_m, g) \zeta_m(x).
\]

(57)

The following proposition allows for a characterization of the rate of the relaxation towards \( w_\infty \).

**Proposition 4.** Let \( W \) be a graphon satisfying assumption [4] then the eigenvalues \( \theta_m \) of \( S - I \) 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 \( S - I \) being a self-adjoint operator on \( L^2[0,1] \). Let \( \lambda \) be an eigenvalue associated to \( \phi \). Then \( \lambda \) is given by the Rayleigh quotient

\[
\lambda = \frac{(\lambda \phi, \phi)}{(\phi, \phi)} = \frac{(S - I) \phi, \phi}{(\phi, \phi)}.
\]

(58)

Consider the numerator of (58). For all \( f \in L^2[0,1] \) we can write

\[
((S - I) f, f) = \int_0^1 \int_0^1 \frac{W(x,y)}{\sqrt{k(x)} \sqrt{k(y)}} f(x) f(y) dx dy - \int_0^1 f^2(x) dx
\]

\[
= \frac{1}{2} \left( \int_0^1 \int_0^1 \sqrt{W(x,y)} \frac{\sqrt{W(x,y)}}{\sqrt{k(x)} \sqrt{k(y)}} f(x) f(y) dx dy - \int_0^1 \int_0^1 \frac{W(x,y)}{k(x)} f^2(x) dx dy \right)
\]

\[- \int_0^1 \int_0^1 \frac{W(x,y)}{k(y)} f^2(x) dx dy - \int_0^1 \int_0^1 \frac{W(x,y)}{k(y)} f^2(y) dx dy \right)
\]

\[
= - \frac{1}{2} \int_0^1 \int_0^1 \left( \frac{\sqrt{W(x,y)}}{\sqrt{k(x)}} f(x) - \frac{\sqrt{W(x,y)}}{\sqrt{k(y)}} f(y) \right)^2 dx dy.
\]

(59)
which is non-positive. The claim that zero is an eigenvalue follows from the fact that $(S - I)\sqrt{k(x)} = 0$ on $[0, 1]$. Finally, let us show that the nullspace of $S - I$ has dimension one if $W$ is connected. By defining $g = M_{\frac{1}{\pi f}}$ on $[0, 1]$, we have to show that if the right-hand side of (59) is zero, namely

$$ -\frac{1}{2} \int_{[0,1]^2} W(x, y) (g(x) - g(y))^2 \, dx \, dy = 0, \quad (60) $$

then $g$ has to be a constant function on $[0, 1]$. By contradiction, assume that there exists some non-constant function $g$ that verifies (60). For simplicity, consider the case that $g = c_1$ on some $S \in \mathcal{W}[0, 1]$ 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 $[0, 1]$, or if $g$ is not a piecewise constant function. Based on (60), we can write

$$ 0 = \int_{[0,1]^2} W(x, y) (g(x) - g(y))^2 \, dx \, dy \geq \int_{S \times S^c} W(x, y) (g(x) - g(y))^2 \, dx \, dy, \quad (61) $$

and the integral in the right-hand side is zero. Since $W$ is connected, we have $\int_{S \times S^c} W(x, y) \, dx \, dy > 0$, 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.

Remark 7. The last claim of proposition 4 means that the spectral gap of $S - I$, namely the positive difference between zero and the second largest eigenvalue, is nonzero when the graphon is connected.

We leave for upcoming work a more detailed study of the spectral properties, and the resulting relaxation time of linear processes with the graphon-based combinatorial and random-walk Laplacian operators.

### 7 The case of the discrete-time walk

The analysis of the node-centric continuous-time walk carries over to the discrete-time version (12). The corresponding IVP on the continuum reads

$$ w(\cdot, \ell + 1) = K w(\cdot, \ell), \quad \ell \in \mathbb{N}_0 \quad (62a) \\
 w(\cdot, 0) = w_0 \in L^2(0, 1), \quad (62b) $$

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

Following the same steps as in sections 3.2.4 and (5), we obtain similar convergence results on the quotient graph $W/\mathcal{P}$, on the sampled graph $W_{[n]}$ and for a sequence of discrete problems. Analogously as for (57), the spectral expansion of the solution of the discrete-time IVP (62) reads

$$ w(\cdot, \ell) = \sum_{m=1}^\infty \lambda_m^\ell \psi_m(w_0) \zeta_m, \quad \ell \geq 0. \quad (63) $$

20
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, as large networks become more and 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 was 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 [25] could be interpreted as the limit of a rescaled edge-centric 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. A more comprehensive study of its spectral properties and applications will be the subject of future research.

The world of random walks is a very broad one, and in this respect the scope of this initial work had to be narrowed. Hence, a promising research direction would consist in generalizing the semigroup approach developed here, or the one in [25, 26, 27], 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, 26, 19].

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] M. Asllani, T. Carletti, F. Di Patti, D. Fanelli, and F. Piazza, Hopping in the crowd to unveil network topology, Physical review letters, 120 (2018), p. 158301.

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

[19] 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.

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

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

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

[23] L. Lovász and B. Szegedy, *Limits of dense graph sequences*, Journal of Combinatorial Theory, Series B, 96 (2006), pp. 933–957.
[24] N. Masuda, M. A. Porter, and R. Lambiotte, Random walks and diffusion on networks, Physics Reports, (2017).

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

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

[27] 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.

[28] 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.

[29] M. Reed and B. Simon, Vol 1: Functional Analysis, Methods of modern mathematical physics, Academic Press, San Diego, 1980.

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

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

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