Continuum limit of the nonlocal p-Laplacian evolution problem on random inhomogeneous graphs

Hafiene Yosra* Jalal M. Fadili* Christophe Chesneau† Abderrahim Elmoataz*

Abstract. In this paper we study numerical approximations of the evolution problem for the nonlocal p-Laplacian operator with homogeneous Neumann boundary conditions on inhomogeneous random convergent graph sequences. More precisely, for networks on convergent inhomogeneous random graph sequences (generated first by deterministic and then random node sequences), we establish their continuum limits and provide rate of convergence of solutions for the discrete models to their continuum counterparts as the number of vertices grows. Our bounds reveals the role of the different parameters, and in particular that of p and the geometry/regularity of the data.

Key words. Nonlocal diffusion; p-Laplacian; inhomogeneous random graphs; graph limits; numerical approximation.

AMS subject classifications. 35A35, 65N12, 65N15, 41A17, 05C80.

1 Introduction

1.1 Problem statement

Our main goal in this paper is to study numerical approximations on random inhomogeneous graphs to a nonlocal nonlinear diffusion problem, involving the nonlocal p-Laplacian operator with homogeneous Neumann boundary conditions. More precisely, the nonlocal p-Laplacian evolution problem with Neumann boundary conditions that we deal with is

$$\begin{cases} \frac{\partial}{\partial t} u(x,t) = -\mathbf{\Delta}_p^K(u(x,t)), & x \in \Omega, t > 0, \\ u(x,0) = g(x), & x \in \Omega, \end{cases} \qquad (\mathcal{P})$$

where

$$\mathbf{\Delta}_p^K(u(x,t)) = -\int_\Omega K(x,y)\big|u(y,t)-u(x,t)\big|^{p-2}(u(y,t)-u(x,t))dy,$$

where $\Omega \subset \mathbb{R}$ a bounded domain, and without loss of generality we take $\Omega = [0,1]$[1]. The kernel $K \in L^\infty(\Omega^2)$ is a symmetric measurable and nonnegative mapping. Throughout the paper, we will assume that $p \in]1,+\infty[$. Existence and uniqueness of a strong solution to ($\mathcal{P}$) in the space $L^p(\Omega)$ was shown in [17, Theorem 3.1] (relying on arguments from [2]).

*Normandie Univ, ENSICAEN, UNICAEN, CNRS, GREYC, France.
†Normandie Univ, ENSICAEN, UNICAEN, CNRS, LMNO, France.

[1] Only boundedness of Ω is actually needed but we take Ω as a closed set as well to conform to our setting of graphs. Moreover, though we here focus on the one-dimensional case $\Omega \subset \mathbb{R}$, several of our results can be extended to higher dimension.

The nonlocal p-Laplacian operator appears naturally in the study of many nonlocal diffusion phenomena. It arises for instance in a number of applications such as continuum mechanics, phase transition phenomena, population dynamics, image processing and game theory (see [1, 2, 16, 20] and the references therein). On the other hand, recently, there has been a high interest in adapting and applying discretized versions of PDEs such as ($\mathcal{P}$) on data defined on arbitrary graphs and networks. Given the discrete nature of data in practice, graphs constitute a natural structure suited to their representation. The demand for such methods is motivated by existing and potential future applications, such as in machine learning and mathematical image/signal/data processing (see among other references [10, 11, 15, 8]). Indeed, any kind of discrete data can be represented by a graph in an abstract form in which the vertices are associated to the data and the edges correspond to relationships within the data. These practical considerations naturally lead to a discrete time and space approximation of ($\mathcal{P}$).

To do this, fix $n \in \mathbb{N}^*$. Let $G_n = (V(G_n), E(G_n))$, where $V(G_n)$ stands for the set of nodes and $E(G_n) \subset V(G_n) \times V(G_n)$ denotes the edges set, be a sequence of simple graphs, i.e. undirected graphs without loops and parallel edges.

Next, we consider the fully discrete counterpart of ($\mathcal{P}$) on a graph G_n using the forward Euler scheme. For that, let us consider a partition (not necessarily uniform) of the time interval $[0, T]$ into intervals of sizes $\{\tau_h\}_{h=1}^N$, $N \in \mathbb{N}^*$, and denote $\tau = \max_{h\in[N]} \tau_h$, where $[N] \stackrel{\text{def}}{=} \{1, \cdots, N\}$ for any integer N. Denote $u_i^h \stackrel{\text{def}}{=} u(x_i, t_h)$ and $g_i \stackrel{\text{def}}{=} g(x_i)$. Then for $h \in [N]$, consider

$$\begin{cases} \dfrac{u_i^h - u_i^{h-1}}{\tau_h} = \dfrac{1}{n} \displaystyle\sum_{j:(i,j)\in E(G_n)} \left|u_j^{h-1} - u_i^{h-1}\right|^{p-2} (u_j^{h-1} - u_i^{h-1}), \\ u_i^0 = g_i, i \in \{1, \cdots, n\}. \end{cases} \tag{$\mathcal{P}^d_{n,\tau}$}$$

Thus, ($\mathcal{P}^d_{n,\tau}$) induces a discrete diffusion process parametrized by the structure of the graph whose adjacency matrix captures the (nonlocal) interactions. As such, it can be viewed as a discrete approximation of a continuum problem such as ($\mathcal{P}$).

Several questions then naturally arise:

- Does the discrete problem ($\mathcal{P}^d_{n,\tau}$), and in what sense, has a continuum limit (as $n \to +\infty$) ?
- What is the rate of convergence to this limit ? Is this limit consistent/related with the unique strong solution of ($\mathcal{P}$) ?
- What are the parameters involved in this rate and what is their influence on the convergence rate ?

This paper provides answers to these questions for graphs drawn from a random inhomogeneous model introduced by [4]. The 'classical' random graph models, in particular dense graphs, are 'homogeneous', in the sense that the nodes degrees tend to be concentrated around a typical value, so that all vertices are exactly equivalent in the definition of the model. Furthermore, in a typical realization, most vertices are in some sense similar to most others. In contrast, some graphs arising in real-world applications do not have this property and are inhomogeneous. One reason is that the vertices may have been 'born' at different times, with old and new vertices having very different properties. Thus, there has been a lot of recent interest in defining and studying networks on inhomogeneous random graph models (see Section 2 for further details). Our aim is to investigate this graph model and study the continuum limit of the p-Laplacian discrete approximation on these random graph models that can be dense or sparse.

1.2 Contributions

In [17], we provided a rigorous justification of the continuum limit ($\mathcal{P}$) for the discrete p-Laplacian on deterministic dense graphs (graphs with n vertices and $\Theta(n^2)$ edges[2]). The analysis of the continuum limit in [17] uses ideas from the theory of dense graph limits [22, 6, 21], which for every convergent family of dense graphs defines the limiting object, a measurable symmetric bounded and nonnegative function K called *graphon* (see Section 2 for a brief overview on graphons). This object characterizes the completion of the space of all graphs with respect to an appropriate metric. In [17], for convergent sequences of deterministic dense graphs $\{G_n\}_{n\in\mathbb{N}}$, it was shown that with the kernel in ($\mathcal{P}$) taken to be the graphon associated to $\{G_n\}_{n\in\mathbb{N}^*}$, the solution of ($\mathcal{P}$) is well-approximated by those of the totally discrete problems ($\mathcal{P}^d_{n,\tau}$). We gave precise convergence rates as a function of n the discretization time step τ.

However, the analysis in [17] does not deal with inhomogeneous graphs, see [4], which allow for sparse (but not too sparse) graphs with $o(n^2)$ but $\omega(n)$ edges. It does not either exhibit the typical error bounds achieved over a sequence of (random) graphs drawn from this inhomogeneous graph model. The main concern of this paper is to bridge this gap by studying continuum limits of ($\mathcal{P}^d_{n,\tau}$) on inhomogeneous random graphs.

Combining tools from evolution equations, random graph theory and deviation inequalities, we establish *nonasymptotic* rate of convergence of the discrete solution to its continuum limit with a controlled (high) probability. More precisely, we start by considering the case of random graph models generated by a deterministic sequence of nodes. We prove nonasymptotic error bounds that hold with a controlled probability (Theorem 3.1). These results serve as a basis to deal with the totally random graph model, i.e., where both the nodes and edges are random (Theorem 3.2). In turn, this shows convergence of solutions of the discrete models to the solution of the continuum problem as the number of vertices n grows. To get the corresponding convergence rate, we additionally assume that the kernel K and the initial data g belong to the versatile class Lipschitz spaces $\mathrm{Lip}(s, L^q(\Omega^2))$ and $\mathrm{Lip}(s', L^q(\Omega))$. Roughly speaking, $\mathrm{Lip}(s, L^q(\Omega^2))$ contains functions with s "derivatives" in $L^q(\Omega^2)$. They contain in particular functions of bounded variation and those of fractal structure for appropriate values of s, see (see Appendix A for a brief introduction to these functional spaces). Using in addition arguments from approximation theory on these spaces, we reveal the influence of the value of p, the density of the graph, the regularity of the graphon K and that of the initial data g both on the convergence rate and the probability of success. In particular, we isolate different regimes where the rate exhibits different scalings.

1.3 Relation to prior work

In [25] and earlier [24], the author studied convergence of discrete approximations of a nonlinear heat equation governed by a Lipschitz continuous potential, first on dense deterministic graphs and then on dense random ones, without discretization of time. However, though the work of [25] was important to us, it differs markedly from ours in many crucial aspects. Indeed, we use some standard arguments from numerical analysis of evolution problems but also specific and sophisticated ones tied to the p-Laplacian. Typically, well-posedness and Lipschitz continuity of the solutions w.r.t.

[2] We recall the usual Landau asymptotic notations: $O(.)$, $\Theta(.)$, $o(.)$ and $\omega(.)$.

to the kernel and initial data for the evolution problem with the p-Laplacian is much harder to establish than for the problem considered in [24, 25] (see [17]). Second, comparing [25] and our current work, we use completely different paths to prove consistency in the random case. Indeed, while the claim in [25] is asymptotic by nature as it completely relies on application of the central limit theorem (CLT), the latter argument cannot be applied to our evolution problem (except for the trivial case $p = 2$). Rather, we establish a nonasymptotic deviation inequality, both in the partly and completely random graph model, relying on a careful control of a random process using sharp inequalities from probability theory (Rosenthal and Bernstein, see Lemma A.1). Thus, we are able to provide the probability of success of our bound for fixed n and we exhibit the dependence of both the error bound and the probability on the problem parameters (p, T, graph model, kernel K, initial data g). This is in a stark contrast to the asymptotic claims in [25].

In [19], the authors extended the analysis of [25] to sparse random graphs corresponding to $L^2(\Omega^2)$ graphons and proved almost sure consistency. While a first version of this paper was under review, we also became aware of the recent preprint [26] which studied the Kuramoto model on a sequence of converging dense and sparse graph sequences. It proved almost sure convergence of the discrete problems on such graphs to continuum limit with time intervals of size $T = O(\log(n))$. In addition to the fact that our evolution problem is different and more intricate, our random graph model is different from that of [19, 26]. Both models allow for sparse graphs, but ours only for those with $o(n^2)$ but $\omega(n)$ edges with bounded graphons, while theirs covers graphs with $O(n)$ edges and $L^q(\Omega^2)$ graphons. Whether our results on the p-Laplacian can be extended to such sparse graphs is an open problem. In fact, even well-posedness (existence and uniqueness) of the p-Laplacian evolution problem ($\mathcal{P}$) with unbounded kernels K remains completely open in the literature. Our results can also cope with time intervals $T = O(\log(n))$ as discussed in Remark 3.3(v). Observe finally that the convergence claim of [19] is asymptotic (almost sure convergence), relying on the standard Markov inequality and Borel-Cantelli lemma, while ours are nonasymptotic with a precise rate and probability of success.

1.4 Paper organization

The rest of the paper is organized as follows. In Section 2, we provide some pre-requisites on graph limits and graphons and then define our inhomogeneous K-random graph model that we deal with throughout the paper. We also specify the assumptions needed to get our results and give a class of graphs for which our assumptions hold true. Section 3 is devoted to the main result of the paper. We begin our analysis by treating random graph sequences generated by deterministic nodes in Section 3.1. Then, in Section 3.2 we consider the general model defined previously in Section 2. After getting the convergence of the discrete model to its continuum limit and identifying the corresponding rate, in Section 3.3, we discuss the different regimes of the convergence rate as a function of the problem parameters. Some technical material is deferred to Appendix A.

1.5 Notations

For a given vector $u = (u_1, \cdots, u_n)^\top \in \mathbb{R}^n$, we define the norm $\|\cdot\|_{p,n}$

$$\|u\|_{p,n} = \left(\frac{1}{n}\sum_{i=1}^{n} |u_i|^p\right)^{\frac{1}{p}}.$$

For an integer $n \in \mathbb{N}^*$, we denote $[n] = \{1, \cdots, n\}$. For any set S, $\overline{S}$ is its closure and $|S|$ is its cardinality or its Lebesgue measure (to be understood from the context). χ_S is the characteristic function of the set S (takes 1 in it and 0 otherwise).

$C(0,T; L^p(\Omega))$ denotes the space of uniformly time continuous functions with values in $L^p(\Omega)$. We endow this space with the norm

$$\|u\|_{C(0,T;L^p(\Omega))} \stackrel{\text{def}}{=} \sup_{t\in[0,T]} \|u(t)\|_{L^p(\Omega)}.$$

For $d \in \{1, 2\}$, $\text{Lip}(s, L^q(\Omega^d))$ is the Lipschitz space which consists of functions with, roughly speaking, s "derivatives" in $L^q(\Omega^d)$ [9, Ch. 2, Section 9]. Only values $s \in]0, 1]$ are of interest to us. See Section A.2 for further details on these spaces and approximation theoretic results on them.

2 The random inhomogeneous graph model

2.1 Graph limits and graphons

We start with some important results from the theory of graph limits that will be crucial to our exposition. The theory of graph limits was introduced by Lovász and Szegedy in 2006 [22] and then further developed in a series of papers (see the book [21] for a comptrehensive bibliography). A key goal of Lovász and Szegedy was to understand large graph structures by characterizing convergence for sequences of graphs which grow unboundedly, thereby constructing a natural limit object.

Let $G_n = (V(G_n), E(G_n))$, $n \in \mathbb{N}^*$, be a sequence of finite and simple graphs. Every finite simple graph G_n such that $V(G_n) = [n]$ can be represented by a measurable function $K_{G_n} : \Omega^2 \to \Omega$ called a *pixel kernel*. Its construction is as follows: split the interval Ω into n equal intervals, and for every $(x, y) \in [\frac{i-1}{n}, \frac{i}{n}[\times[\frac{j-1}{n}, \frac{j}{n}[$, define

$$K_{G_n}(x,y) = \begin{cases} 1 & \text{if} \quad (i,j) \in E(G_n), \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

For *weighted* graphs with edge weights $\{\beta(i,j)\}_{(i,j)\in V(G)^2}$, the pixel kernel K_{G_n} becomes

$$K_{G_n}(x,y) = \begin{cases} \beta(i,j) & \text{if} \quad (i,j) \in E(G_n), \\ 0 & \quad\text{otherwise.} \end{cases} \tag{2}$$

This construction is not unique, however given a graph, the set of pixel kernels arising via (1) can be considered to be equivalent via the weakly isomorphic relation.

Convergent graph sequences $\{G_n\}_{n\in\mathbb{N}^*}$ have a limit object, which can be represented as a measurable and symmetric function $K : \Omega^2 \to [0, 1]$ called *graphon*, and the function K is uniquely determined up to measure-preserving transformation; see [7, Theorem 2.1]. Intuitively, a graphon can be thought of as a generalization of the adjacency matrix of a (weighted) graph which has a continuum number of vertices. Actually, the space of graphons is the the completion of the metric space of graphs, or equivalently pixel kernels, relative to the so-called cut distance; see [6, Theorem 2.6]. Conversely, [7, Theorem 2.1] also shows that every measurable symmetric and $[0, 1]$-valued function K arises as the limit of a convergent graph sequence.

The result of [7, Theorem 2.1] proves existence and uniqueness of the limit graphon but it is not a constructive result. In fact, there is a natural "limit object" in the form of a symmetric measurable function $K : \Omega^2 \to [0,1]$ which arises as a limit of an appropriate graph sequence but this limit is not explicitly known for every graph sequence. The natural question is then whether, given a graphon K, one can construct a sequence fo graphs $\{G_n\}_{n\in\mathbb{N}^*}$ whose limit is K. It turns out that there is such a random construction as we show in the next two sections.

2.2 Random graphs

The theory of random graphs was founded in the 50s-60s by Erdös and Rényi [12], who started the systematic study of the space of graphs with n labeled vertices and $M = M(n)$ edges, with all graphs equiprobable. Nearly the same time, Gilbert [14] introduced the closely related model of random graphs on n labeled vertices obtained as follows: join each pair $(i, j) \in [n]^2$ of vertices independently, with probability $p = p(n)$. These graphs are now known as Erdös-Rényi random graphs.

The aim is to turn the set of all graphs with n vertices into a probability space. Intuitively we should be able to generate G_n randomly as follows: for each edge $(i, j) \in [n]^2$, we decide by some random experiment wether or not (i, j) shall be an edge of G_n, these experiments are performed independently.

Lovász et al. [22, 7] defined a more general random graph model called K-random graph, as follows: given any symmetric measurable function $K : \Omega^2 \to [0,1]$ and an integer $n \in \mathbb{N}^*$, generate n independent numbers $\mathbf{X}_1, \ldots, \mathbf{X}_n$ from the uniform distribution on Ω, and then connect nodes i and j with probability $K(\mathbf{X}_i, \mathbf{X}_j)$. The Erdös-Rényi graph corresponds to K being the constant p-valued function. It was shown in [22, Corollary 2.6] that this random construction provides a sequence of graphs $\{G_n\}_{n\in\mathbb{N}^*}$ which converges almost surely to the graphon K.

2.3 K-random inhomogeneous graph model

2.3.1 Model generation

The random graph models defined above are "homogeneous" in the sense that all vertices are exactly equivalent in the definition of the model. Furthermore, in a typical realization, most vertices are in some sense similar to most others. For example, the vertex degrees in the Erdös-Rényi model do not vary very much: their distribution is close to a Poisson distribution. However, many large real-world graphs are inhomogeneous. One reason is that the vertices may have been 'born' at different times, with old and new vertices having very different properties. This has led to the introduction and analysis of many new random graph models designed to incorporate or explain inhomogeneous features. We here focus on a particular random graph model that will be used throughout. This random graph model is motivated by the construction of inhomogeneous random graphs proposed in [3, 4, 5].

Definition 2.1. *Fix $n \in \mathbb{N}^*$ and let $K : \Omega^2 \to [0,1]$ be a symmetric measurable function. Generate the undirected graph $G_n = (V(G_n), E(G_n)) \stackrel{\text{def}}{=} G_{q_n}(n, K)$ as follows:*

1) *Generate n independent and identically distributed (i.i.d.) random variables $(\mathbf{X}_1, \ldots, \mathbf{X}_n) \stackrel{\text{def}}{=} \mathbf{X}$ from the uniform distribution on Ω. Let $\left\{\mathbf{X}_{(i)}\right\}_{i=1}^n$ be the order statistics of the random vector $\mathbf{X}$, i.e. $\mathbf{X}_{(i)}$ is the i-th smallest value.*

2) Conditionally on $\mathbf{X}$, join each pair $(i,j) \in [n]^2$ of vertices independently, with probability $q_n \hat{K}{}^{\mathbf{X}}_{nij}$, i.e. for every $(i,j) \in [n]^2$, $i \neq j$,

$$\mathbb{P}\left((i,j) \in E(G_n)|\mathbf{X}\right) = q_n \hat{K}{}^{\mathbf{X}}_{nij}, \tag{3}$$

where

$$\hat{K}{}^{\mathbf{X}}_{nij} \overset{\text{def}}{=} \min\left(\frac{1}{\left|\Omega^{\mathbf{X}}_{nij}\right|}\int_{\Omega^{\mathbf{X}}_{nij}} K(x,y)dxdy, 1/q_n\right), \tag{4}$$

and

$$\Omega^{\mathbf{X}}_{nij} \overset{\text{def}}{=}]\mathbf{X}_{(i-1)}, \mathbf{X}_{(i)}] \times]\mathbf{X}_{(j-1)}, \mathbf{X}_{(j)}] \tag{5}$$

where q_n is non-negative and uniformly bounded in n.

A graph $G_{q_n}(n,K)$ generated according to this procedure is called a K-random inhomogeneous graph generated by a random sequence $\mathbf{X}$.

Following [4], we write the parameter q_n as a subscript to emphasize that it is part of the normalization. For appropriate choices of q_n, this model allows to sample both dense and sparse graphs from the kernel K. In the latter, we think of a sparse graph generated from K, rather than a "sparse kernel" $q_n K$.

At this stage, the following important remark is in order.

Remark 2.1. *In the context of numerical analysis, we are primarily interested not only in the error bounds of the discrete problem, but more importantly in the (nonasymptotic) rate of convergence. This is why our attention aims specifically at this graph model and not at the original inhomogeneous random model defined in [3, 4], i.e. the model constructed replacing (3) by*

$$\mathbb{P}\left((i,j) \in E(G_n)\right) = \min\left(q_n K(\mathbf{X}_i, \mathbf{X}_j), 1\right).$$

Our error bounds of the discrete problem ($\mathcal{P}^d_{n,\tau}$) cover also this graph model, and more specifically, the first statements of Theorem 3.1 and Theorem 3.2 hold. However, with this model, even our convergence claim (not to mention the rate) of the discrete scheme does not hold unless the kernel K and the initial data g are additionally supposed almost everywhere continuous.

We denote by $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ the realization of $\mathbf{X}$. To lighten the notation, we also denote

$$\Omega^{\mathbf{X}}_{ni} \overset{\text{def}}{=}]\mathbf{X}_{(i-1)}, \mathbf{X}_{(i)}], \quad \Omega^{\mathbf{x}}_{ni} \overset{\text{def}}{=}]\mathbf{x}_{(i-1)}, \mathbf{x}_{(i)}], \quad \text{and} \quad \Omega^{\mathbf{x}}_{nij} \overset{\text{def}}{=}]\mathbf{x}_{(i-1)}, \mathbf{x}_{(i)}] \times]\mathbf{x}_{(j-1)}, \mathbf{x}_{(j)}] \quad i,j \in [n]. \tag{6}$$

As the realization of the random vector $\mathbf{X}$ is fixed, we define

$$\hat{K}{}^{\mathbf{x}}_{nij} \overset{\text{def}}{=} \min\left(\frac{1}{\left|\Omega^{\mathbf{x}}_{nij}\right|}\int_{\Omega^{\mathbf{x}}_{nij}} K(x,y)dxdy, 1/q_n\right), \quad \forall (i,j) \in [n]^2, \quad i \neq j. \tag{7}$$

In the rest of the paper, the following random variables will be useful. Let λ_{ij}, $(i,j) \in [n]^2, i \neq j$, be independent random variables such that $q_n \lambda_{ij}$ follows a Bernoulli distribution with parameter $q_n \hat{K}{}^{\mathbf{x}}_{nij}$. We consider the independent random variables Υ_{ij} such that the distribution of $q_n \Upsilon_{ij}$ conditionally on $\mathbf{X} = \mathbf{x}$ is that of $q_n \lambda_{ij}$. Thus $q_n \Upsilon_{ij}$ follows a Bernoulli distribution with parameter $\mathbb{E}\big(q_n \hat{K}{}^{\mathbf{X}}_{nij}\big)$, where $\mathbb{E}(\cdot)$ is the expectation operator (here with respect to the distribution of $\mathbf{X}$).

2.3.2 Model assumptions

We are now ready to formulate our assumptions on the graph sequence $\{G_{q_n}(n,K)\}_{n\in\mathbb{N}}$.

Assumption 2.1. *We suppose that q_n and K are such that the following hold:*

(A.1) $G_{q_n}(n,K)$ *converges almost surely and its limit is the graphon K;*

(A.2) $\sup_{n\geq 1} q_n \leq 1$.

At this stage, it is legitimate to discuss the validity of Assumption 2.1. As far as the boundedness assumption (A.2) is concerned, it is the least we can expect from q_n; otherwise the generated graphs are trivially empty. Of course, there is no loss of generality in taking 1 in the bound of (A.2). For (A.1), it is also reasonable as it allows one to assert that the randomly generated graph sequence has an (almost sure) limit object, which is the graphon K. The forthcoming proposition provides a large class of K-random graphs, corresponding to a particular scaling of q_n, for which both (A.1) and (A.2) hold. It is is inspired by the so-called *non-uniform* random graphs studied in [4, Section 3.4]. This choice of q_n allows to cover both dense and reasonably sparse graphs (see the discussion after Proposition 2.1) which are ubiquitous in various applications such as data (e.g., signal/image/point clouds) processing.

Proposition 2.1. *Suppose $K : \Omega^2 \to [0,1]$ is a symmetric measurable function. Choose the parameter $q_n = n^{-g(n)}$ where $g(n) = o(1)$. Then, assumptions (A.1) and (A.2) are in force.*

Proof . Since the graphon $K \in L^\infty(\Omega^2)$ and $q_n = n^{-o(1)}$, the arguments to prove [4, Lemma 3.5 and Lemma 3.8], that were designed for the graph model described in Remark 2.1, can be adapted to cover our graph model with (3) to show that the sequence of random graphs $G_{q_n}(n,K)$ indeed converges almost surely to the graphon K in the metric d_{sub} (see [4, Section 2.1] for details about this metric). This shows (A.1). (A.2) is trivially verified. □

The graph model of Proposition 2.1 encompasses the dense random graph model (i.e., with $\Theta(n^2)$ edges) extensively studied in [22, 7], by taking the choice $g(n)\log(n) = C$, for $C > 0$, and thus $q_n = e^{-C}$. This graph model allows also to generate sparse (but not too sparse) graphs. That is graphs with $o(n^2)$ but $\omega(n)$ edges, i.e., that the average degree tends to infinity with n. For example, one can take $g(n) = C\log(n)^{-\delta}$, where $\delta \in]0,1[$, in which case one has $q_n = \exp(-C\log(n)^{1-\delta}) = o(1)$, where such a choice of q_n will control the level of density/sparsity.

3 Consistency of the nonlocal p-Laplacian on random inhomogeneous graphs

Having defined the graph model, we are now in position to state our main error bounds between the discrete dynamics and their continuum counterparts. In Section 3.1, we first deal with the case where $\mathbf{X}$ is deterministic. Capitalizing on this result, we will then extend the bounds to the totally random model (i.e., where the nodes are also random) in Section 3.2 by a marginalization argument.

3.1 Networks on graphs generated by deterministic nodes

We define the parameter $\delta(n)$ as the maximal size of the spacings between the ordered values $\mathbf{x}_{(i)}$

$$\delta(n) = \max_{i\in[n]} \left|\mathbf{x}_{(i)} - \mathbf{x}_{(i-1)}\right|. \tag{8}$$

Recall from ($\mathcal{P}^d_{n,\tau}$) the definition of the time steps τ_h and maximal size τ.

Next, we consider the following system of difference equations on $G_{q_n}(n, K)$:

$$\begin{cases} \dfrac{u_i^h - u_i^{h-1}}{\tau_h} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} \lambda_{ij}\left|u_j^{h-1} - u_i^{h-1}\right|^{p-2}(u_j^{h-1} - u_i^{h-1}), \quad (i,h) \in [n]\times[N], \\ u_i^0 = g_i, i \in [n], \end{cases} \tag{$\mathcal{P}^{d,d}_n$}$$

where u_i^h is the value at vertex $i \in [n]$ and time $t_h \in [0, T]$, and

$$g_i = \frac{1}{\left|\Omega^{\mathbf{x}}_{ni}\right|}\int_{\Omega^{\mathbf{x}}_{ni}} g(x)dx.$$

Recall from Section 2 that $\{q_n\lambda_{ij}\}_{i,j}$ are independent Bernoulli variables with parameters $\{q_n \hat{K}{}^{\mathbf{x}}_{nij}\}_{i,j}$.

Before turning to our convergence result, we pause here to make the following observations.

Remark 3.1. *Coming back to Definition 2.1, $G_{q_n}(n, K)$ is a random variable taking its values on the set of simple graphs. It is then important to keep in mind that the evolution equations we write involving random variables must be understood in this sense.*

Remark 3.2. *As the reader may have remarked, the sum in the right-hand side of ($\mathcal{P}^{d,d}_n$) is divided by n instead of a weighted sum with weights $\left|\mathbf{x}_{(i)} - \mathbf{x}_{(i-1)}\right|^{-1}$ which would be expected if we interpret this sum as a Riemann sum. The scaling by n reminds us of an equidistant design regarding the space-discretization, despite the fact that the nodes are not necessarily equispaced. However, given that the $\mathbf{x}_i$'s are realizations of i.i.d. uniform variables on Ω, the uniform spacing choice still makes sense. Indeed, using classical results on order statistics of uniform variables, see, e.g., [27, Section 1.7], it can be shown that each spacing $\mathbf{X}_{(i)} - \mathbf{X}_{(i-1)}$ concentrates around $1/n$ for any $i \in [n]$.*

We are now in position to tackle our main goal: comparing the solutions of the discrete and continuum problems and establish our rate of convergence. Since the two solutions do not live on the same spaces, it is reasonable to represent some intermediate model that is the continuum extension of the discrete problem, using the vector $U^h = (u_1^h, u_2^h, \cdots, u_n^h)^\top$ whose components solve the previous system ($\mathcal{P}^{d,d}_n$) to obtain the following interpolation on $\Omega \times [0, T]$ (linear in time and piecewise constant in space)

$$\check{u}_n(x,t) = \frac{t_h - t}{\tau_h}u_i^{h-1} + \frac{t - t_{h-1}}{\tau_h}u_i^h \quad \text{if} \quad x \in \Omega^{\mathbf{x}}_{ni}, \quad t \in]t_{h-1}, t_h], \tag{9}$$

and a space-time piecewise constant approximation

$$\bar{u}_n(x,t) = \sum_{i=1}^{n}\sum_{h=1}^{N} u_i^{h-1}\chi_{]t_{h-1},t_h]}(t)\chi_{\Omega^{\mathbf{x}}_{ni}}(x). \tag{10}$$

Then, $\breve{u}_n$ formally solves the following problem

$$\begin{cases} \frac{\partial}{\partial t}\breve{u}_n(x,t) = -\mathbf{\Delta}_p^{\Lambda_n}(\bar{u}_n(x,t)), \quad x \in \Omega, t > 0, \\ \breve{u}_n(x,0) = g_n(x), \quad x \in \Omega, \end{cases} \tag{$\mathcal{P}_n$}$$

where the random variable

$$\Lambda_n(x,y) = \lambda_{ij} \quad \text{for} \quad (x,y) \in \Omega^{\mathbf{x}}_{nij},$$

and

$$g_n(x) = g_i \quad \text{if} \quad x \in \Omega^{\mathbf{x}}_{ni}, i \in [n].$$

Toward our goal of establishing error bounds, we need an intermediate discrete problem for the p-Laplacian. This is defined as

$$\begin{cases} \dfrac{v_i^h - v_i^{h-1}}{\tau_h} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} \hat{K}{}^{\mathbf{x}}_{nij} \left|v_j^{h-1} - v_i^{h-1}\right|^{p-2}(v_j^{h-1} - v_i^{h-1}), \quad (i,h) \in [n] \times [N], \\ v_i^0 = g_i, \quad i \in [n]. \end{cases} \tag{$\overset{\wedge}{\mathcal{P}}{}^d_n$}$$

The discrete problem ($\overset{\wedge}{\mathcal{P}}{}^d_n$) can also be viewed as a discrete p-Laplacian evolution problem over a weighted graph on n vertices, where the weight of edge (i,j) is $\hat{K}{}^{\mathbf{x}}_{nij}$.

Using the vector $V^h = (v_1^h, v_2^h, \cdots, v_n^h)^\top$ whose components solve the system ($\overset{\wedge}{\mathcal{P}}{}^d_n$), similarly to before, we define the following interpolation on $\Omega \times [0,T]$

$$\breve{v}_n(x,t) = \frac{t_h - t}{\tau_h} v_i^{h-1} + \frac{t - t_{h-1}}{\tau_h} v_i^h \quad \text{if} \quad x \in \Omega^{\mathbf{x}}_{ni}, \quad t \in]t_{h-1}, t_h], \tag{11}$$

and a space-time piecewise constant interpolation

$$\bar{v}_n(x,t) = \sum_{i=1}^{n}\sum_{h=1}^{N} v_i^{h-1} \chi_{]t_{h-1},t_h]}(t)\chi_{\Omega^{\mathbf{x}}_{ni}}(x). \tag{12}$$

We also define the piecewise-constant extension $\hat{K}_n$ on Ω^2

$$\hat{K}_n(x,y) = \sum_{(i,j)\in[n]^2} \hat{K}{}^{\mathbf{x}}_{nij}\chi_{\Omega^{\mathbf{x}}_{nij}}(x,y). \tag{13}$$

Then, by construction, $\breve{v}_n(x,t)$ formally solves the following problem

$$\begin{cases} \frac{\partial}{\partial t}\breve{v}_n(x,t) = -\mathbf{\Delta}_p^{\hat{K}_n}(\bar{v}_n(x,t)), \quad x \in \Omega, t > 0, \\ \breve{v}_n(x,0) = g_n(x), \quad x \in \Omega, \end{cases} \tag{$\overset{\wedge}{\mathcal{P}}_n$}$$

where

$$g_n(x) = g_i \quad \text{for} \quad x \in \Omega^{\mathbf{x}}_{ni}, i \in [n].$$

The first main result of the paper is the following theorem.

Theorem 3.1. *Suppose that $p \in]1,+\infty[$, $K \in L^\infty(\Omega^2)$ is a symmetric and measurable mapping, and $g \in L^\infty(\Omega)$. Let u and U^h denote the solutions to $(\mathcal{P})$ and $(\mathcal{P}_n^{d,d})$, respectively. Let $\check{u}_n$ be the continuum extension of U^h given in (9). Then, the following hold:*

(i) for $T>0$, there exist positive constants C_1 and C_2, independent of n and T, such that for any $\beta > 0$

$$\left\|u-\check{u}_n\right\|_{C(0,T;L^p(\Omega))} \leq C_1 T \exp\left(O(T)\right)\left(\left(\beta\frac{\log(n)}{n}+\frac{\max\left(q_n^{-(p-1)},q_n^{-p/2}\right)}{n^{p/2}}\right)^{1/p} + \left\|K-\hat{K}_n\right\|_{L^p(\Omega^2)} + \left\|g-g_n\right\|_{L^p(\Omega)} + \tau\right), \quad (14)$$

with probability at least $1-n^{-C_2 q_n^{2p-1}\beta}$.

(ii) Suppose furthermore that $g \in \mathrm{Lip}(s, L^q(\Omega))$ and $K \in \mathrm{Lip}(s', L^q(\Omega^2))$, $q \in [1,+\infty]$, $s, s' \in]0,1]$, and $q_n \left\|K\right\|_{L^\infty(\Omega^2)} \leq 1$. Then, for $T > 0$, there exist positive constants C_1 and C_2, independent of n and T, such that for any $\beta > 0$

$$\left\|u-\check{u}_n\right\|_{C(0,T;L^p(\Omega))} \leq C_1 T \exp\left(O(T)\right)\left(\left(\beta\frac{\log(n)}{n}+\frac{\max\left(q_n^{-(p-1)},q_n^{-p/2}\right)}{n^{p/2}}\right)^{1/p} + \delta(n)^{\min(s,s')\min(1,q/p)} + \tau\right), \quad (15)$$

with probability at least $1-n^{-C_2 q_n^{2p-1}\beta}$, where $\delta(n)$ is the spacing parameter defined in (8).

Before proceeding to the proof, some remarks are in order.

Remark 3.3.

(i) The constant C_1 in (14) depends on p and the data via $\left\|g\right\|_{L^\infty(\Omega)}$ and $\left\|K\right\|_{L^\infty(\Omega^2)}$. For the bound (15), it also depends on (q, s, s'). Similarly, the constant in the $O(T)$ term in the exponential can be made explicit and depends on p, $\left\|g\right\|_{L^\infty(\Omega)}$ and $\left\|K\right\|_{L^\infty(\Omega^2)}$.

(ii) Thanks to the well-known inequality

$$(a+b)^\alpha \leq a^\alpha + b^\alpha, \quad \forall \alpha \in]0,1] \text{ and } a,b \geq 0, \quad (16)$$

it is clear that the first term in the bounds (14)-(15) can be replaced by

$$\beta^{1/p}\left(\frac{\log(n)}{n}\right)^{1/p} + \frac{\max\left(q_n^{-(1-1/p)},q_n^{-1/2}\right)}{n^{1/2}}.$$

(iii) The last term in the latter bound can be rewritten as

$$n^{-1/2}\max\left(q_n^{-(1-1/p)},q_n^{-1/2}\right) = \begin{cases} (q_n n)^{-1/2} & \text{if } p \in]1,2], \\ q_n^{1/p}(q_n^2 n)^{-1/2} & \text{if } p > 2. \end{cases} \quad (17)$$

Thus, if $\inf_{n\geq 1} q_n > 0$*, as is the case when the graph is dense (see discussion after Proposition 2.1), then the term* (17) *is in the order of* $n^{-1/2}$ *with probability at least* $1 - n^{-c\beta}$ *for some* $c > 0$*. If* q_n *is allowed to be* $o(1)$*, i.e., sparse graphs (see Proposition 2.1), then* (17) *is* $o(1)$ *if either* $q_n n \to +\infty$ *for* $p \in]1, 2]$*, or* $q_n^2 n \to +\infty$ *for* $p > 2$*. The probability of success is at least* $1 - e^{-C_2\beta \log(n)^{1-\delta}}$ *provided that* $q_n = \log(n)^{-\delta/(2p-1)}$*, with* $\delta \in [0, 1[$*. Observe that all these conditions on* q_n *are fulfilled by the graph model of Proposition 2.1 for* $g(n) = \delta/(2p-1)\log(\log(n))/\log(n)$*.*

(iv) In fact, if $\inf_{n\geq 1} q_n \geq c > 0$*, then we have* $\sum_{n\geq 1} n^{-C_2 q_n^{2p-1}\beta} \leq \sum_{n\geq 1} n^{-C_2 c^{2p-1}\beta} < +\infty$ *provided that* $\beta > (C_2 c^{2p-1})^{-1}$*. Thus, if this holds, invoking the (first) Borel-Cantelli lemma, it follows that the bounds of Theorem 3.1 hold almost surely. The same reasoning carries over for the bounds of Theorem 3.2.*

(v) For finite fixed T*, the term* $T\exp(c_1 T)$*, for* $c_1 > 0$*, in the bound becomes a constant. One can even allow for time intervals of size* $T = c_2 \log(n)$*,* $c_2 > 0$*, in which case this term scales as* $O(n^{c_1 c_2} \log(n))$*. Thus this term can be dominated by the other rates in* n *if* $c_1 c_2$ *is sufficiently small (see Remark 3.4(ii) for details).*

(vi) One may wonder if the functional space assumption made on g *and* K *in claim (ii) is reasonable or even makes sense. The answer is affirmative. Indeed, Lipschitz spaces are rich enough to include both functions with discontinuities and even fractal structure. For instance, from [21], one can show that the graphon corresponding to the nearest neighbour graphs, which are very popular in practice (e.g. in image processing [11, 10]), are typical examples satisfying Assumptions (A.1)-(A.2) with* $q_n = 1$ *and* K *is a* $\{0, 1\}$*-valued function living on the space of bounded variation functions, which in turn is* $\mathrm{Lip}(1, L^1(\Omega^2))$*.*

To prove Theorem 3.1, we first show the following key lemma.

Lemma 3.1. *Under the assumptions of Theorem 3.1, for* $T > 0$*, there exist positive constants* C_1 *and* C_2*, independent of* n *and* T*, such that for any* $\beta > 0$

$$\mathbb{P}\left(\left\|\check{v}_n - \check{u}_n\right\|_{C(0,T;L^p(\Omega))} \geq \varepsilon\right) \leq n^{-C_2 q_n^{2p-1}\beta},$$

where

$$\varepsilon = C_1 T \exp\left(O(T)\right)\left(\left(\beta\frac{\log(n)}{n} + \max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)\frac{1}{n^{p/2}}\right)^{1/p} + \tau\right).$$

Proof of Lemma 3.1. For $1 < p < +\infty$, we define the function

$$\begin{aligned}\Psi &: \mathbb{R} \to \mathbb{R}\\ x &\mapsto |x|^{p-2}\, x = \mathrm{sign}(x)|x|^{p-1}.\end{aligned}$$

First, for an appropriate choice of τ_h, using [17, Lemma 5.1], we have that both $(\mathcal{P}_n^{d,d})$ and $(\hat{\mathcal{P}}_n^{d})$ are well posed. In turn U^h and V^h are bounded and V^h uniquely solves $(\hat{\mathcal{P}}_n^{d})$, and similarly for $\check{u}_n$ and $\check{v}_n$ as solutions to $(\mathcal{P}_n)$ and $(\hat{\mathcal{P}}_n)$. Observe also that $\check{v}_n(\cdot, t)$ and $\check{u}_n(\cdot, t)$ are both constants over $\Omega^{\mathbf{x}}_{ni}$. Similarly, $\bar{v}_n(\cdot,t)$ and $\bar{u}_n(\cdot,t)$ are also constants over the cell $\Omega^{\mathbf{x}}_{ni}$. We therefore used the shorthand notations for the vector-valued functions $\bar{\mathbf{u}}_n(t) = (\bar{\mathbf{u}}_{ni}(t))_{i\in[n]} \stackrel{\text{def}}{=} (\bar{u}_n(\mathbf{x}_i,t))_{i\in[n]}$ and $\bar{\mathbf{v}}_n(t) = (\bar{\mathbf{v}}_n(t))_{i\in[n]} \stackrel{\text{def}}{=} (\bar{v}_n(\mathbf{x}_i,t))_{i\in[n]}$, and likewise for $\check{\mathbf{u}}_n(t)$ and $\check{\mathbf{v}}_n(t)$. Let us denote $\check{\xi}_n(t) = \check{\mathbf{u}}_n(t) - \check{\mathbf{v}}_n(t)$ and $\bar{\xi}_n(t) = \bar{\mathbf{u}}_n(t) - \bar{\mathbf{v}}_n(t)$. By subtracting both sides of ($\mathcal{P}_n$) from those of ($\overset{\wedge}{\mathcal{P}_n}$), evaluated at the cell $\Omega^{\mathbf{x}}_{ni}$, we obtain

$$
\begin{aligned}
\frac{d}{dt}\check{\xi}_{ni}(t) &= \frac{1}{n}\sum_{j=1}^{n}\Big(\lambda_{ij}\Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)) - \hat{K}{}^{\mathbf{x}}_{nij}\Psi(\bar{\mathbf{v}}_{nj}(t)-\bar{\mathbf{v}}_{ni}(t))\Big) \\
&= Z_{ni}(t) + \frac{1}{n}\sum_{j=1}^{n}\hat{K}{}^{\mathbf{x}}_{nij}\big(\Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)) - \Psi(\bar{\mathbf{v}}_{nj}(t)-\bar{\mathbf{v}}_{ni}(t))\big),
\end{aligned}
\tag{18}
$$

where

$$
Z_{ni}(t) = \frac{1}{n}\sum_{j=1}^{n}(\lambda_{ij} - \hat{K}{}^{\mathbf{x}}_{nij})\alpha_{ij}(t) \ \text{ and } \ \alpha_{ij}(t) = \Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)), \forall (i,j)\in[n]^2, t\in[0,T]. \tag{19}
$$

By our discussion above, we have $\sup_{(i,j)\in[n]^2, t\in[0,T]}|\alpha_{ij}(t)| < +\infty$. We multiply both sides of (18) by $\frac{1}{n}\Psi(\check{\xi}_{ni}(t))$ and sum over i to obtain

$$
\frac{1}{p}\frac{d}{dt}\|\check{\xi}_n(t)\|^p_{p,n} = \frac{1}{n}\sum_{i=1}^{n}Z_{ni}(t)\Psi(\check{\xi}_{ni}(t)) + \frac{1}{n^2}\sum_{i,j=1}^{n}\hat{K}{}^{\mathbf{x}}_{nij}\big(\Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)) - \Psi(\bar{\mathbf{v}}_{nj}(t)-\bar{\mathbf{v}}_{ni}(t))\big)\Psi(\check{\xi}_{ni}(t)). \tag{20}
$$

We estimate the first term on the right-hand side of (20) using the Hölder inequality, to get

$$
\frac{1}{n}\left|\sum_{i=1}^{n}Z_{ni}(t)\Psi(\check{\xi}_{ni}(t))\right| \le \frac{1}{n}\left(\sum_{i=1}^{n}|Z_{ni}(t)|^p\right)^{\frac{1}{p}} \times \left(\sum_{i=1}^{n}|\check{\xi}_{ni}(t)|^p\right)^{\frac{p-1}{p}} \le \|Z_n(t)\|_{p,n}\|\check{\xi}_n(t)\|^{p-1}_{p,n}. \tag{21}
$$

Now, using the fact that $\hat{K}{}^{\mathbf{x}}_{nij} \le \|K\|_{L^\infty(\Omega^2)}$ (see (4)), $\forall (i,j)\in[n]^2$, and applying a generalized mean value theorem ([17, Corollary B.1]) to the function Ψ, since $p>1$, between $a_{ij}(t) = \bar{\mathbf{v}}_{nj}(t) - \bar{\mathbf{v}}_{ni}(t)$ and $b_{ij}(t) = \bar{\mathbf{u}}_{nj}(t) - \bar{\mathbf{u}}_{ni}(t)$ (without loss of generality, we suppose that $b_{nij}(t) > a_{ij}(t)$), we get

$$
\begin{aligned}
&\left|\frac{1}{n^2}\sum_{i,j=1}^{n}\hat{K}{}^{\mathbf{x}}_{nij}\big(\Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)) - \Psi(\bar{\mathbf{v}}_{nj}(t)-\bar{\mathbf{v}}_{ni}(t)\big)\Psi(\xi_{ni}(t))\right| \\
&\le \frac{(p-1)\,\|K\|_{L^\infty(\Omega^2)}}{n^2}\sum_{i,j=1}^{n}|\bar{\xi}_{nj}-\bar{\xi}_{ni}||\eta_{ij}(t)|^{p-2}|\check{\xi}_{ni}|^{p-1},
\end{aligned}
\tag{22}
$$

where $\eta_{ij}(t)$ is an intermediate value between $a_{ij}(t)$ and $b_{ij}(t)$. Using that fact that $g\in L^\infty(\Omega)$ and the construction of $\bar{\mathbf{u}}_n(\cdot)$, we deduce from [17, Theorem 3.1(ii)] that for any $(i,j)\in[n]^2$ and $t\in[0,T]$, we have for $p\ge 2$

$$
|\eta_{ij}(t)|^{p-2} \le |\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)|^{p-2} \le \big(2\|u(\cdot,t)\|_{L^\infty(\Omega)}\big)^{p-2} \le \big(2\|g\|_{L^\infty(\Omega)}\big)^{p-2}.
$$

For $p \in]1,2[$, since $\inf_{(i,j,t)\in[n]^2\times[0,T]} |\eta_{ij}(t)| = C' > 0$, we have

$$|\eta_{ij}(t)|^{p-2} \leq C'^{p-2} < +\infty.$$

Altogether, we obtain

$$|\eta_{ij}(t)|^{p-2} \leq \max\left(\left(2\|g\|_{L^\infty(\Omega)}\right)^{p-2}, C'^{p-2}\right). \tag{23}$$

Let $C_2 = \max\left(\left(2\|g\|_{L^\infty(\Omega)}\right)^{p-2}, C'^{p-2}\right)\|K\|_{L^\infty(\Omega^2)}$. Inserting (23) into (22), and then using the Hölder and triangle inequalities, it follows that

$$\begin{aligned}
&\left|\frac{1}{n^2}\sum_{i,j=1}^{n} \hat{K}{}^{\mathbf{x}}_{nij}\left(\Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)) - \Psi(\bar{\mathbf{v}}_{nj}(t)-\bar{\mathbf{v}}_{ni}(t)\right)\Psi(\check{\xi}_{ni}(t))\right| \\
&\leq C_2\frac{p-1}{n^2}\sum_{i,j=1}^{n}\left|\bar{\xi}_{nj}(t)-\bar{\xi}_{ni}(t)\right|\left|\check{\xi}_{ni}\right|^{p-1} \\
&\leq C_2\frac{p-1}{n^2}\left(\left(\sum_{i,j=1}^{n}\left|\bar{\xi}_{nj}(t)-\bar{\xi}_{ni}(t)\right|^p\right)^{\frac{1}{p}}\left(\sum_{i,j}\left|\check{\xi}_{ni}(t)\right|^p\right)^{\frac{p-1}{p}}\right) \\
&\leq C_2\frac{p-1}{n^2}\left(\left(\sum_{i,j=1}^{n}\left|\bar{\xi}_{nj}(t)\right|^p\right)^{\frac{1}{p}} + \left(\sum_{i,j=1}^{n}\left|\bar{\xi}_{ni}(t)\right|^p\right)^{\frac{1}{p}}\right)\left(n^{\frac{2(p-1)}{p}}\left(\frac{1}{n}\sum_{i=1}^{n}\left|\check{\xi}_{ni}(t)\right|^p\right)^{\frac{p-1}{p}}\right) \\
&\leq C_2\frac{p-1}{n^2}\left(2n^{\frac{2}{p}}\|\bar{\xi}_n(t)\|_{p,n}\right)\left(n^{\frac{2(p-1)}{p}}\|\check{\xi}_n(t)\|_{p,n}^{p-1}\right) \\
&\leq 2C_2(p-1)\|\bar{\xi}_n(t)\|_{p,n}\|\check{\xi}_n(t)\|_{p,n}^{p-1}.
\end{aligned} \tag{24}$$

Using the triangle inequality combined with [17, Lemma 5.2], we have

$$\begin{aligned}
\|\bar{\xi}_n(t)\|_{p,n} &= \|\bar{\mathbf{v}}_n(t)-\bar{\mathbf{u}}_n(t)\|_{p,n} \\
&\leq \|\bar{\mathbf{v}}_n(t)-\check{\mathbf{v}}_n(t)\|_{p,n} + \|\check{\mathbf{v}}_n(t)-\check{\mathbf{u}}_n(t)\|_{p,n} + \|\check{\mathbf{u}}_n(t)-\bar{\mathbf{u}}_n(t)\|_{p,n} \\
&\leq C\tau + \|\check{\xi}_n(t)\|_{p,n} + C'\tau \\
&\leq C''\tau + \|\check{\xi}_n(t)\|_{p,n}.
\end{aligned} \tag{25}$$

Putting together (20), (21), (24) and (25), we have

$$\begin{aligned}
\frac{d}{dt}\|\check{\xi}_n(t)\|_{p,n}^p &\leq \|Z_n(t)\|_{p,n}\|\check{\xi}_n(t)\|_{p,n}^{p-1} + 2C_2(p-1)\left(C''\tau + \|\check{\xi}_n(t)\|_{p,n}\right)\|\check{\xi}_n(t)\|_{p,n}^{p-1} \\
&\leq \left(2C_3(p-1)\tau + \|Z_n(t)\|_{p,n}\right)\|\check{\xi}_n(t)\|_{p,n}^{p-1} + 2C_2(p-1)\|\check{\xi}_n(t)\|_{p,n}^p.
\end{aligned} \tag{26}$$

Then, from (26) via the Gronwall's inequality in its differential form (see, e.g., [13, Appendix B]), we obtain

$$\|\check{u}_n - \check{v}_n\|_{C(0,T;L^p(\Omega))} = \sup_{t\in[0,T]}\|\check{\xi}_n(t)\|_{p,n} \leq \left(2C_3T\tau + \int_0^T\|Z_n(t)\|_{p,n}dt\right)\exp\left(2C_2T\right). \tag{27}$$

It remains to bound $\int_0^T \|Z_n(t)\|_{p,n} dt$. For this purpose, we use Lemma A.1 (see Section A.1)[3]. Thus, plugging the bound of Lemma A.1(i) into inequality (27), we get the desired conclusion. □

We are now ready to prove our main result.

Proof of Theorem 3.1. (i) Using the triangle inequality, we have

$$\|u - \check{u}_n\|_{C(0,T;L^p(\Omega))} \leq \|u - \check{v}_n\|_{C(0,T;L^p(\Omega))} + \|\check{v}_n - \check{u}_n\|_{C(0,T;L^p(\Omega))}. \tag{28}$$

Since by construction $\hat{K}_n$ is a bounded mapping, we bound the first term on the right-hand side of (28) using [17, Theorem 5.1][4] to get

$$\|u - \check{v}_n\|_{C(0,T;L^p(\Omega))} = O\left(T \exp(O(T))\left(\|K - \hat{K}_n\|_{L^p(\Omega^2)} + \|g - g_n\|_{L^p(\Omega)} + \tau\right)\right), \tag{29}$$

Claim (14) then follows by plugging (29) and Lemma 3.1 into (28).

(ii) Our assumption on q_n together with (7) and (13) entail that

$$\hat{K}_n(x, y) = \sum_{(i,j)\in[n]^2} K_{nij}\chi_{\Omega^{\mathbf{x}}_{nij}}(x, y), \quad K_{nij} = \frac{1}{|\Omega^{\mathbf{X}}_{nij}|}\int_{\Omega^{\mathbf{X}}_{nij}} K(x, y)dxdy$$

Since $g \in \mathrm{Lip}(s, L^q(\Omega))$ and $K \in \mathrm{Lip}(s', L^q(\Omega^2))$, we can invoke Lemma A.3 to get

$$\|K - \hat{K}_n\|_{L^p(\Omega^2)} \leq C(p, q, s')\delta(n)^{s' \min(1,q/p)} \text{ and } \|g - g_n\|_{L^p(\Omega)} \leq C(p, q, s)\delta(n)^{s \min(1,q/p)}. \tag{30}$$

Inserting the bound (30) into (14), and using the fact that $\delta(n) < 1$, yields (15). □

3.2 Networks on graphs generated by random nodes

Let us now turn to the totally random graph model. Consider the following system of difference equations on the totally random graph $G_{q_n}(n, K)$[5] :

$$\begin{cases} \dfrac{u_i^h - u_i^{h-1}}{\tau_h} = \dfrac{1}{n}\displaystyle\sum_{\{j:\ (i,j)\in E(G_{q_n}(n,K))\}} |u_j^{h-1} - u_i^{h-1}|^{p-2}(u_j^{h-1} - u_i^{h-1}), \quad h \in [N] \\ u_i^0 = g_i, i \in [n]. \end{cases} \tag{$\mathcal{P}_n^{r,d}$}$$

As we have done before, we consider the continuum extension of the solution vector $U^h = (u_1^h, u_2^h, \cdots, u_n^h)^\top$, that is a linear interpolation on $\Omega \times [0, T]$

$$\check{u}_n(x, t) = \frac{t_h - t}{\tau_h}u_i^{h-1} + \frac{t - t_{h-1}}{\tau_h}u_i^h \quad \text{if} \quad x \in \Omega^{\mathbf{X}}_{ni}, \quad t \in]t_{h-1}, t_h], \tag{31}$$

[3]This inequality is sharp as can be seen for instance from assertion (ii) of Lemma A.1, at least for $p \geq 2$.

[4]Here, we have made the constant explicit in T compared to the statement in [17, Theorem 5.1].

[5]Recall again from Remark 3.1, that rigorously speaking, each variable involved in the problems and equations of this section should be understood as random.

and a piecewise approximation

$$\bar{u}_n(x,t) = \sum_{i=1}^{n}\sum_{h=1}^{N} u_i^{h-1}\chi_{]t_{h-1},t_h]}(t)\chi_{\Omega_{ni}^{\mathbf{X}}}(x). \tag{32}$$

Then, we have

$$\begin{cases} \frac{\partial}{\partial t}\breve{u}_n(x,t) = -\boldsymbol{\Delta}_p^{\Gamma_n}(\bar{u}_n(x,t)), & x \in \Omega, t > 0, \\ \breve{u}_n(x,0) = g_n(x), & x \in \Omega \end{cases} \tag{$\mathcal{P}_n^r$}$$

where

$$g_n(x) = g_i \quad \text{if} \quad x \in \Omega_{ni}^{\mathbf{X}}, i \in [n],$$

and the random variable Γ_n is such that

$$\Gamma_n(x,y) = \Upsilon_{ij} \quad \text{for} \quad (x,y) \in \Omega_{nij}^{\mathbf{X}}.$$

If conditioned with respect to a realization $\mathbf{x} = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$ of the random vector $\mathbf{X}$, problem ($\mathcal{P}_n^{r,d}$) can be rewritten on $G_{q_n}(n,K)$ in the following form

$$\begin{cases} \dfrac{u_i^h - u_i^{h-1}}{\tau_h} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} \lambda_{ij}\left|u_j^{h-1} - u_i^{h-1}\right|^{p-2}(u_j^{h-1} - u_i^{h-1}), & (i,h) \in [n] \times [N], \\ u_i^0 = g_i, \quad i \in [n]. \end{cases} \tag{$\mathcal{P}_n^d$}$$

By capitalizing on the results obtained for the the case where $\{G_{q_n}(n,K)\}_{n\in\mathbb{N}}$ was generated by the deterministic sequence $\mathbf{x}$, we get the following result.

Theorem 3.2. *Suppose that $p \in]1,+\infty[$, $K \in L^\infty(\Omega^2)$ is a symmetric and measurable mapping, and $g \in L^\infty(\Omega)$. Let u and U_h denote the solutions to ($\mathcal{P}$) and ($\mathcal{P}_n^{r,d}$), respectively. Let $\breve{u}_n$ be the continuum extension of U_h given in (31). Then, the following hold:*

(i) For $T > 0$, there exist positive constants C_1 and C_2, independent of n and T, such that for any $\beta > 0$

$$\left\|u - \breve{u}_n\right\|_{C(0,T;L^p(\Omega))} \leq C_1 T \exp\left(O(T)\right)\left(\left(\beta\frac{\log(n)}{n} + \frac{\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)}{n^{p/2}}\right)^{1/p} + \left\|K - \hat{K}_n\right\|_{L^p(\Omega^2)} + \left\|g - g_n\right\|_{L^p(\Omega)} + \tau\right), \tag{33}$$

with probability at least $1 - n^{-C_2 q_n^{2p-1}\beta}$.

(ii) Suppose furthermore that $g \in \mathrm{Lip}(s, L^q(\Omega))$ and $K \in \mathrm{Lip}(s', L^q(\Omega^2))$, $s, s' \in]0,1]$, and $q_n \left\|K\right\|_{L^\infty(\Omega^2)} \leq 1$. Let $\theta \overset{\text{def}}{=} \min(s,s')\min(1,q/p)$. Then, for $T > 0$, there exist positive constants C_1 and C_2, independent of n and T, such that for any $\beta > 0$ and $t \in]0,e[$

$$\left\|u - \breve{u}_n\right\|_{C(0,T;L^p(\Omega))} \leq C_1 T \exp\left(O(T)\right)\left(\left(\beta\frac{\log(n)}{n} + \frac{\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)^{1/p}}{n^{p/2}}\right) + \left(\frac{t\log(n)}{n}\right)^\theta + \tau\right), \tag{34}$$

with probability at least $1 - \left(n^{-C_2 q_n^{2p-1}\beta} + n^{-t}\right)$.

Remark 3.4.

(i) The dependence of the constant C in the parameters is similar to Remark 3.3(i).

(ii) As observed in Remark 3.3(v), one can take $T = c_2 \log(n)$, in which case $T \exp(c_1 T) = c_2 n^{c_1 c_2} \log(n)$, with $c_1, c_2 > 0$. Consequently, if one sets $q_n = \log(n)^{-\delta/(2p-1)}$, for $\delta \in]0,1[$ (see Remark 3.3(iii)), then the bound in (34) scales as $O\left(\frac{\log(n)^s}{n^{\min(1/p,1/2,\theta)-c_1c_2}}\right)$, for some $s > 0$, which converges to 0 provided that $c_1 c_2 < \min(1/p, 1/2, \theta)$.

As a preparatory step to prove Theorem 3.2, the following lemma is instrumental. It establishes that the spacings between the n uniformly distributed nodes are $O(\log(n)/n)$ with high probability.

Lemma 3.2. *Consider the sequence of random spacings $(\mathbf{X}_{(1)}, \mathbf{X}_{(2)} - \mathbf{X}_{(1)}, \ldots, 1 - \mathbf{X}_{(n)})$, where we recall $\{\mathbf{X}_{(i)}\}_{i=1}^n$ are the order statistics of $\mathbf{X}$. Let $t \in]0, e[$. Then, for any $i \in [n]$*

$$\delta_i \stackrel{\text{def}}{=} \mathbf{X}_{(i)} - \mathbf{X}_{(i-1)} \leq t \frac{\log(n)}{n}, \tag{35}$$

with probability at least $1 - n^{-t}$.

Proof of Lemma 3.2. Since $\mathbf{X}_i$ are i.i.d. uniform random variables on Ω, we have, by virtue of [27, Theorem 1.6.7] that the random variables δ_i, $i \in [n]$, have the same distribution as the random variables $Z_i / \sum_{k=1}^{n+1} Z_k$, where $Z_1, \cdots, Z_{n+1}$ are i.i.d standard exponential random variables. In addition, invoking [27, Lemma 1.6.6], we know that $S_{n+1} \stackrel{\text{def}}{=} \sum_{k=1}^{n+1} Z_k$ is a Gamma random variable with parameters $(1, n+1)$ (thus having the density $f_{S_{n+1}}(s) = e^{-s} s^n / n!$, $s \geq 0$).

Now, combining these two observations, we obtain by straightforward integral calculations that for any $\varepsilon \in [0, 1[$

$$\begin{aligned}
\mathbb{P}(\delta_i \geq \varepsilon) = \mathbb{P}(Z_i \geq \varepsilon S_{n+1}) &= \mathbb{P}((1-\varepsilon) Z_i \geq \varepsilon (S_{n+1} - Z_i)) \\
&= \mathbb{P}\left(Z_{n+1} \geq \frac{\varepsilon}{1-\varepsilon} S_n\right) \\
&= \int_0^{+\infty} \mathbb{P}\left(Z_{n+1} \geq \frac{\varepsilon}{1-\varepsilon} s\right) f_{S_n}(s) ds \\
&= \int_0^{+\infty} e^{-\frac{\varepsilon}{1-\varepsilon} s} e^{-s} \frac{s^{n-1}}{(n-1)!} ds \\
&= (1-\varepsilon)^n.
\end{aligned} \tag{36}$$

The equality of the second line stems from an equality in distribution, since $S_{n+1} - Z_i$ has the same distribution as S_n and Z_i has the same distribution as Z_{n+1}, and the fact that Z_i and $S_{n+1} - Z_i$ are independent. Taking $\varepsilon = t \frac{\log(n)}{n} \in]0, 1[$, and using the standard inequality $\log(1-u) \leq -u$, for $u \in [0, 1]$, we get

$$\mathbb{P}(\delta_i \geq \varepsilon) = (1-\varepsilon)^n = \exp(n \log(1-\varepsilon)) \leq \exp(-n\varepsilon) = n^{-t}.$$

□

Proof of Theorem 3.2. The idea of the proof is to take the conditional probability with respect to a fixed realization $\mathbf{x} = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$ of the random vector $\mathbf{X}$, then use the bound in Theorem 3.1, which is independent of $\mathbf{x}$, and finally integrate with respect to the uniform density on Ω^n.

(i) We have

$$\begin{aligned}
\mathbb{P}\left(\|u - \check{u}_n\|_{C(0,T;L^p(\Omega))} \geq \varepsilon'\right) &= \frac{1}{|\Omega|^n}\int_{\Omega^n} \mathbb{P}\left(\|u - \check{u}_n\|_{C(0,T;L^p(\Omega))} \geq \varepsilon' | \mathbf{X} = \mathbf{x}\right) d\mathbf{x} \\
&\leq \frac{1}{|\Omega|^n}\int_{\Omega^n} n^{-C_2 q_n^{2p-1}\beta} d\mathbf{x} \\
&= n^{-C_2 q_n^{2p-1}\beta},
\end{aligned} \tag{37}$$

with

$$\varepsilon' = C_1 T \exp\left(O(T)\right)\left(\left(\beta\frac{\log(n)}{n} + \frac{\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)}{n^{p/2}}\right)^{1/p} + \|K - \hat{K}_n\|_{L^p(\Omega^2)} + \|g - g_n\|_{L^p(\Omega)} + \tau\right).$$

Thus, (33) follows from the fact that the obtained bound in (14) is independent of the random choice of $\mathbf{x}$.

(ii) In view of (30), we can argue that

$$\mathbb{P}\left(\|K - \hat{K}_n\|_{L^p(\Omega^2)} + \|g - g_n\|_{L^p(\Omega)} \geq \kappa\right) \leq \mathbb{P}\left(\left(C(p,q,s) + C(p,q,s')\right)\delta(n)^\theta \geq \kappa\right).$$

Taking $\kappa = \left(C(p,q,s) + C(p,q,s')\right)\left(t\frac{\log(n)}{n}\right)^\theta$, for $t \in]0, e[$, and applying Lemma 3.2, we deduce that

$$\mathbb{P}\left(\|K - \hat{K}_n\|_{L^p(\Omega^2)} + \|g - g_n\|_{L^p(\Omega)} \geq \kappa\right) \leq n^{-t}.$$

Denote the events

$$\begin{aligned}
A_1 &: \left\{\|\check{v}_n - \check{u}_n\|_{C(0,T;L^p(\Omega))} \leq \varepsilon\right\} \\
A_2 &: \left\{\|K - \hat{K}_n\|_{L^p(\Omega^2)} + \|g - g_n\|_{L^p(\Omega)} \leq \kappa'\right\}
\end{aligned}$$

and their complements A_i^c, where

$$\varepsilon = CT\exp\left(O(T)\right)\left(\left(\beta\frac{\log(n)}{n} + \frac{\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)}{n^{p/2}}\right)^{1/p} + \tau\right)$$

and $\kappa' = CT\exp\left(O(T)\right)\left(t\frac{\log(n)}{n}\right)^\theta$, with C the largest constants among the one in claim (i) and
$\left(C(p,q,s) + C(p,q,s')\right)$. Using the union bound, we get

$$\mathbb{P}\left(\|u - \check{u}_n\|_{C(0,T;L^p(\Omega))} \leq \varepsilon + \kappa'\right) \geq \mathbb{P}\left(\cap_{i=1}^2 A_i\right) = 1 - \mathbb{P}\left(\cup_{i=1}^2 A_i^c\right)$$

$$\geq 1 - \sum_{i=1}^{2} \mathbb{P}\left(A_i^c\right) \geq 1 - \left(n^{-C_2 q_n^{2p-1}\beta} + n^{-t}\right),$$

which yields the desired claim.

□

3.3 Rate regimes

A close inspection of the error bound in (34) (Theorem 3.2) reveals three contributions:

- Spatial discretization and edge sampling: the first contribution is materialized in the first term which scales as (see Remark 3.3(ii))

$$O\left(\left(\frac{\log(n)}{n}\right)^{1/p} + \frac{\max\left(q_n^{-(1-1/p)}, q_n^{-1/2}\right)}{n^{1/2}}\right).$$

  This term represents the spatial discretization error when approximating the continuous evolution equation ($\mathcal{P}$) on the random inhomogeneous graph model $G_{q_n}(n, K)$ generated according to Definition 2.1 with the graphon K.

- Data approximation and node sampling: the second term is $O\left(\left(\frac{\log(n)}{n}\right)^{\theta}\right)$ which captures the error of approximating the initial data g and the graphon K on the grid of size $\delta(n)$ which concentrates around $\log(n)/n$. The presence of the error on K is clearly tied to the nonlocal nature of the evolution equation on graphs. This approximation error depends on the regularity of g and K, and the latter encodes the geometry/structure of the underlying graphs. The more regular g and K are, the faster the convergence rate.

- Time discretization: the last term, which is $O(\tau)$, is classical and corresponds to the time discretization error.

At this stage, one may wonder which of the first two terms dominate, or in other words, what are the different regimes exhibited by the convergence rate as a function of the problem parameters (p, q, s, s'). This is quite important as it will reveal which nonlocal p-Laplacian evolution problems are harder/easier to discretize by highlighting the role of each parameter, and for instance that of p and the impact of nonlocality (i.e. graphon structure).

Toward this goal, we first make the error measure in (34) independent of p and we choose to quantify the error in the classical $L^2(\Omega)$ norm. Consequently, thanks to the classical inequalities (16) and (48), as well as boundedness of the solutions, it is not difficult to see that

$$\left\|u - \check{u}_n\right\|_{C(0,T;L^2(\Omega))} = \begin{cases} O\left(\left(\beta\frac{\log(n)}{n}\right)^{1/p} + \frac{\max\left(q_n^{-(1-1/p)}, q_n^{-1/2}\right)}{n^{1/2}} + \left(\frac{t\log(n)}{n}\right)^{\theta} + \tau\right), & p \in [2, +\infty[\\ O\left(\left(\beta\frac{\log(n)}{n}\right)^{1/2} + \frac{\max\left(q_n^{-(p-1)/2}, q_n^{-p/4}\right)}{n^{p/4}} + \left(\frac{t\log(n)}{n}\right)^{p\theta/2} + \tau^{p/2}\right) & p \in]1, 2], \end{cases} \tag{38}$$

holds with probability at least $1 - \left(n^{-C_2 q_n^{2p-1}\beta} + n^{-t}\right)$.

To make the rest of the discussion more concrete we will take $q_n = \log(n)^{-\delta/(2p-1)}$, with $\delta \in [0,1[$, which covers both dense ($\delta = 0$) and non-dense ($\delta \in]0,1[$) graphs; see Remark 3.3(iii) and Proposition 2.1. Thus, we have

$$\max\left(q_n^{-(1-1/p)}, q_n^{-1/2}\right) = \begin{cases} O\big(\log(n)^{1/2}\big) & p \in [2,+\infty[\\ O\big(\log(n)^{p/4}\big) & p \in]1,2], \end{cases}$$

In turn, the second term in (38) is bounded by

$$\left(\frac{\log(n)}{n}\right)^{\min(p/4,1/2)}, \forall p \in]1,+\infty[. \tag{39}$$

Without loss of generality[6], we also suppose that $s = s'$ and $q \leq p$ so that $\theta = sq/p \in]0, q/p] \subset]0,1]$. In this case, (38) reads

$$\left\|u - \check{u}_n\right\|_{C(0,T;L^2(\Omega))} = O\left(\left(\frac{\log(n)}{n}\right)^{\min(1/p,1/2,sq/p)\min(p/2,1)} + \tau^{\min(p/2,1)}\right).$$

The term depending on n then exhibits four different regimes as a function of p, s and q (see Figure 1). Indeed, it is straightforward to see that it scales as

$$\begin{cases} \left(\frac{\log(n)}{n}\right)^{sq/p} & \text{for} \quad p \geq 2, \quad sq \in]0,1], \\ \left(\frac{\log(n)}{n}\right)^{1/p} & \text{for} \quad p \geq 2, \quad sq \in]1,p], \\ \left(\frac{\log(n)}{n}\right)^{sq/2} & \text{for} \quad p \in]1,2], \quad sq \in]0,p/2], \\ \left(\frac{\log(n)}{n}\right)^{p/4} & \text{for} \quad p \in]1,2], \quad sq \in [p/2,p]. \end{cases}$$

In particular, the convergence rate shows a transition phenomenon at $p = 2$. The rate increases with p for $p \in]2,+\infty[$ while it decreases with p for $p \in]1,2]$ and $sq \in [p/2,p]$. As expected, the dependence of the rate on the initial data g and graphon K is more prominent as they become irregular, i.e. for smaller values of sq. For small sq and $p \in]1,2]$, the rate is independent of p.

A Appendix

A.1 A key deviation result

The following lemma establishes a key deviation inequality for $\int_0^T \left\|Z_n(t)\right\|_{p,n} dt$ where $Z_n(\cdot)$ is a random process defined as

$$Z_{ni}(t) = \frac{1}{n}\alpha_{ij}(t)\sum_{j=1}^{n}(\lambda_{ij} - \gamma_{ij}), \tag{40}$$

where $\sup_{(i,j)\in[n]^2, t\in[0,T]} |\alpha_{ij}(t)| < +\infty$, and the λ_{ij}'s are independent random variables such that $q_n\lambda_{ij}$ is Bernoulli with parameter $q_n\gamma_{ij}$, where $\sup_{i,j}\gamma_{ij} < +\infty$ and q_n satisfies (A.2). It is obvious that this process covers that in (19) as a special case.

[6]This setting is true for many graphons, see, e.g., Remark 3.3(vi).

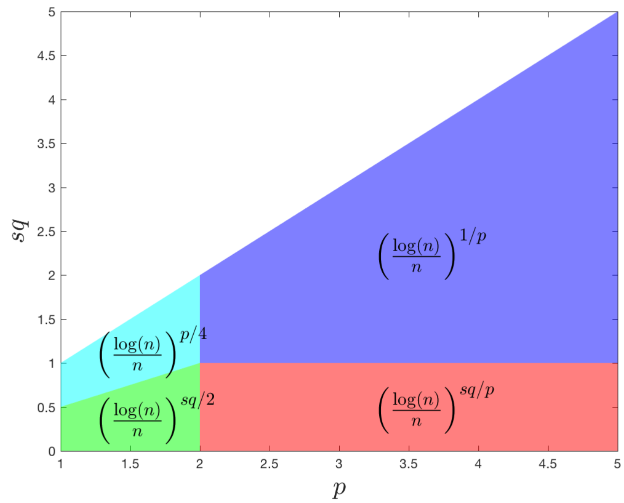

Figure 1: Different regimes according to the values of p and s, and q.

Lemma A.1. *Let $Z_n(\cdot)$ be the random process defined in (40). Then, we have*

(i) For $p \in [1, +\infty[$, $T > 0$, there exists a positive constant C, such that for any $\beta > 0$

$$\mathbb{P}\left(\int_0^T \left\|Z_n(t)\right\|_{p,n} dt \geq \varepsilon\right) \leq n^{-Cq_n^{2p-1}\beta},$$

with

$$\varepsilon = T\left(\beta\frac{\log(n)}{n} + C_3 \max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)\frac{1}{n^{p/2}}\right)^{1/p},$$

where C_3 is a positive constant which will be explicit in the proof.

(ii) For $p \in [2, +\infty[$, suppose that there exists a positive constant C, such that for $t > 0$

$$\inf_{j\in[n]} \frac{1}{n}\sum_{i>j}\frac{\alpha_{ij}^2(t)}{q_n}\gamma_{ij}(1 - q_n\gamma_{ij}) \geq C.$$

Then,

$$\mathbb{E}\left(\int_0^T \left\|Z_n(t)\right\|_{p,n}^p dt\right) \sim \frac{T}{n^{p/2}}.$$

To prove this lemma, we need the following deviation inequalities that we include for the reader convenience.

Rosenthal's inequality [18]. Let n be a positive integer, $\gamma \geq 2$ and $U_1, \ldots, U_n$ be n zero mean independent random variables such that $\sup_{i\in[n]} \mathbb{E}(|U_i|^\gamma) < \infty$. Then there exists a positive constant C such that

$$\mathbb{E}\left(\left|\sum_{i=1}^n U_i\right|^\gamma\right) \leq C \max\left(\sum_{i=1}^n \mathbb{E}(|U_i|^\gamma), \left(\sum_{i=1}^n \mathbb{E}(U_i^2)\right)^{\gamma/2}\right).$$

Bernstein's inequality [23]. Let n be a positive integer and $U_1, \ldots, U_n$ be n zero mean independent random variables such that there exists a positive constant M satisfying $\sup_{i\in[n]} |U_i| \leq M < \infty$. Then, for any $\upsilon > 0$,

$$\mathbb{P}\left(\sum_{i=1}^{n} U_i \geq \upsilon\right) \leq \exp\left(-\frac{\upsilon^2}{2\left(\sum\limits_{i=1}^{n} \mathbb{E}\left(U_i^2\right) + \upsilon M/3\right)}\right).$$

Proof of Lemma A.1. (i) Using the Jensen inequality, we have

$$\mathbb{P}\left(\int_0^T \left\|Z_n(t)\right\|_{p,n} dt \geq \varepsilon\right) \leq \mathbb{P}\left(T^{p-1}\int_0^T \left\|Z_n(t)\right\|_{p,n}^p dt \geq \varepsilon^p\right).$$

Let us first recall that $q_n\lambda_{ij}$ are independent Bernoulli random variables with parameters $q_n\gamma_{ij}$. For the sake of simplicity, set, for $(i,j) \in [n]^2$, $Y_{ni} \stackrel{\text{def}}{=} \int_0^T \left|\frac{1}{n}\sum\limits_{j=1}^{n} U_{nij}(t)\right|^p dt$, where $U_{nij}(t) \stackrel{\text{def}}{=} \alpha_{ij}(t)(\lambda_{ij} - \gamma_{ij})$. We have

$$I \stackrel{\text{def}}{=} \mathbb{P}\left(\int_0^T \left\|Z_n(t)\right\|_{p,n}^p dt \geq T^{1-p}\varepsilon^p\right) = \mathbb{P}\left(\frac{1}{n}\left(\sum_{i=1}^{n} Y_{ni} - \mathbb{E}(Y_{ni})\right) \geq T^{1-p}\varepsilon^p - \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(Y_{ni})\right).$$

It remains now to bound $\mathbb{E}\left(Y_{ni}\right)$. We distinguish the cases where $p \geq 2$ and $p \in]1, 2[$.

- $p \geq 2$. Using the Rosenthal inequality with the independent according to j zero-mean random variables $U_{nij}(t)$, we have

$$\begin{aligned}\mathbb{E}\left(Y_{ni}\right) &= \frac{1}{n^p}\int_0^T \mathbb{E}\left(\left|\sum_{j=1}^{n} U_{nij}(t)\right|^p\right) dt \\ &\leq \frac{C_1 T}{n^p}\sup_{t\in[0,T]}\max\left(\sum_{j=1}^{n}\mathbb{E}(\left|U_{nij}(t)\right|^p), \left(\sum_{j=1}^{n}\mathbb{E}(U_{nij}(t)^2)\right)^{p/2}\right). \qquad (41)\end{aligned}$$

  We have

$$\begin{aligned}\mathbb{E}\left(\left|U_{nij}(t)\right|^p\right) &= q_n^{-p}\left|\alpha_{ij}(t)\right|^p\left|q_n\gamma_{ij}(1-q_n\gamma_{ij})^p + (q_n\gamma_{ij})^p(1-q_n\gamma_{ij})\right| \\ &= q_n^{-(p-1)}\left|\alpha_{ij}(t)\right|^p\gamma_{ij}(1-q_n\gamma_{ij})\left((q_n\gamma_{ij})^{p-1} + (1-q_n\gamma_{ij})^{p-1}\right).\end{aligned}$$

  Taking $p = 2$, we get

$$\mathbb{E}(U_{nij}(t)^2) = q_n^{-1}\alpha_{ij}^2(t)\gamma_{ij}(1-\gamma_{ij}).$$

  Since $\sup_{(i,j)\in[n]^2, t\in[0,T]}|\alpha_{ij}(t)| < +\infty$, and γ_{ij} is also bounded and p being greater than 2, there exists $C_2 > 0$, such that,

$$\max\left(\sum_{j=1}^{n}\mathbb{E}(\left|U_{nij}(t)\right|^p), \left(\sum_{j=1}^{n}\mathbb{E}(U_{nij}(t)^2)\right)^{p/2}\right) \leq C_2\max\left(nq_n^{-(p-1)}, n^{p/2}q_n^{-p/2}\right)$$

$$\le C_2 \max\big(q_n^{-(p-1)}, q_n^{-p/2}\big) n^{p/2}.$$

Therefore

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(Y_{ni}\right) \le C_1 C_2 T \max\left(q_n^{-(p-1)}, q_n^{-p/2}\right) n^{-p/2}. \tag{42}$$

- $p \in [1, 2[$. Observe that by the mutual independence of the random variables $\{\lambda_{ij}\}_{(i,j)\in[n]^2}$, we deduce that $\{U_{nij}(t)\}_{j=1}^{n}$ are independent and zero-mean random variables. Thus

$$\mathbb{E}\left(\left(\sum_{j=1}^{n} U_{nij}(t)\right)^2\right) = \operatorname{Var}\left(\sum_{j=1}^{n} U_{nij}(t)\right) = \sum_{j=1}^{n}\mathbb{E}\left(U_{nij}(t)^2\right). \tag{43}$$

Therefore, applying the Jensen inequality to the concave function $x \mapsto x^{p/2}$, we obtain

$$\begin{aligned}
\mathbb{E}\left(Y_{ni}\right) \le \frac{T}{n^p}\sup_{t\in[0,T]}\mathbb{E}\left(\big|\sum_{j=1}^{n} U_{nij}(t)\big|^p\right) &\le \frac{T}{n^p}\sup_{t\in[0,T]}\left(\mathbb{E}\left(\left(\sum_{j=1}^{n} U_{nij}(t)\right)^2\right)\right)^{p/2}\\
&= \frac{T}{n^p}\sup_{t\in[0,T]}\left(\sum_{j=1}^{n}\mathbb{E}\left(U_{nij}(t)^2\right)\right)^{p/2}\\
&= \frac{T}{n^p}\sup_{t\in[0,T]}\left(\sum_{j=1}^{n}\frac{\alpha_{ij}(t)^2}{q_n}\gamma_{ij}(1-q_n\gamma_{ij})\right)^{p/2}\\
&\le \frac{C_2 T}{q_n^{p/2}} n^{-p/2} \le C_2 T \max\left(q_n^{-(p-1)}, q_n^{-p/2}\right) n^{-p/2}.
\end{aligned} \tag{44}$$

Altogether, we have shown that for any $p \ge 1$,

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(Y_{ni}\right) \le C_3 T \max\left(q_n^{-(p-1)}, q_n^{-p/2}\right) n^{-p/2}, \tag{45}$$

where $C_3 = C_2 \max(1, C_1)$.

Hence, setting $W_{ni} = Y_{ni} - \mathbb{E}\left(Y_{ni}\right)$ and $\kappa = T^{1-p}\varepsilon^p - C_3 T \max\left(q_n^{-(p-1)}, q_n^{-p/2}\right) n^{-p/2}$, we have

$$I \le \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} W_{ni} \ge \kappa\right).$$

Let $\varepsilon > 0$ such that $\kappa > 0$. Observe that the random variables $\{W_{ni}\}_{i=1}^{n}$ are independent, zero-mean, and obey:

▷ $\sup_{i\in[n]}\big|W_{ni}\big| \le 2\sup_{i\in[n]}\big|Y_{ni}\big| \le C_4 T$, since α_{ij} and $q_n\gamma_{ij}$ are both uniformly bounded.

$\triangleright$ $\sum_{i=1}^{n} \mathbb{E}\left(W_{ni}^2\right) = \sum_{i=1}^{n} \mathrm{Var}\left(Y_{ni}\right) \leq \sum_{i=1}^{n} \mathbb{E}\left(Y_{ni}^2\right)$. Using the Jensen inequality with the function $x \mapsto x^2$, and replacing the exponent "p" in inequality (41), by "$2p$" which is greater than 2, we obtain

$$\sum_{i=1}^{n} \mathbb{E}\left(W_{ni}^2\right) \leq \sum_{i=1}^{n} \mathbb{E}\left(Y_{ni}^2\right) \leq C_5 T^2 \max\left(q_n^{-(2p-1)}, q_n^{-p}\right) \frac{1}{n^{p-1}}.$$

We are then in position to apply the Bernstein inequality to $\{W_{ni}\}_{i=1}^{n}$ according to the index i, whence we get, after some elementary algebra

$$\begin{aligned}\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} W_{ni} \geq \kappa\right) &\leq \exp\left(-\frac{n^2\kappa^2}{2\left(\sum_{i=1}^{n} \mathbb{E}\left(W_{ni}^2\right) + n\kappa C_4 T/3\right)}\right) \\ &\leq \exp\left(-\frac{C_6}{2}\min\left(q_n^{2p-1}, q_n^{p}\right)\frac{n\kappa^2}{n^{-p}T^2 + \kappa T}\right).\end{aligned}$$

Taking $\kappa = \beta T \frac{\log(n)}{n} > Tn^{-p}$, for $p \geq 1$, we have after straightforward calculations

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} W_{ni} \geq \kappa\right) \leq \exp\left(-\frac{C_6}{4}\min\left(q_n^{2p-1}, q_n^{p}\right) n\kappa/T\right) = n^{-\frac{C_6}{4}\min\left(q_n^{2p-1}, q_n^{p}\right)\beta}.$$

In turn,

$$I \leq \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} W_{ni} \geq \kappa\right) \leq n^{-C\min\left(q_n^{2p-1}, q_n^{p}\right)\beta}.$$

For this choice of κ, observe that

$$\begin{aligned}\kappa = \beta T\frac{\log(n)}{n} &\Leftrightarrow T^{1-p}\varepsilon^p - C_3 T \max\left(q_n^{-(p-1)}, q_n^{-p/2}\right) n^{-p/2} = \beta T\frac{\log(n)}{n} \\ &\Leftrightarrow \varepsilon = T\left(\beta\frac{\log(n)}{n} + C_3\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)\frac{1}{n^{p/2}}\right)^{1/p}.\end{aligned}$$

Thus

$$\mathbb{P}\left(\int_0^T \left\|Z_n(t)\right\|_{p,n} dt \geq \varepsilon\right) \leq n^{-C\min\left(q_n^{2p-1}, q_n^{p}\right)\beta}. \tag{46}$$

As $q_n \leq 1$ by (A.2) and $2p-1 \geq p$ for $p \in\geq 1$, we obviously have $\min\left(q_n^{2p-1}, q_n^{p}\right) = q_n^{2p-1}$.

(ii) Recalling the notation in the proof of claim (i), we have

$$\forall (i,j) \in [n]^2, \frac{1}{n}\sum_{i=1}^{n} Y_{ni} = \frac{1}{n}\int_0^T \sum_{i=1}^{n}\left|Z_{ni}(t)\right|^p dt = \frac{1}{n^{p+1}}\int_0^T \sum_{i=1}^{n}\left|\sum_{j=1}^{n} U_{nij}(t)\right|^p dt.$$

Thus, for $p \in [2, +\infty[$, applying the Jensen inequality and using (43), we have

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(Y_{ni}) &= \frac{1}{n^{p+1}}\int_0^T \sum_{i=1}^{n}\mathbb{E}\left(\big|\sum_{j=1}^{n}U_{nij}(t)\big|^p\right)dt \\
&\geq \frac{1}{n^{p+1}}\int_0^T \sum_{i=1}^{n}\left(\mathbb{E}\left(\sum_{j=1}^{n}U_{nij}(t)\right)^2\right)^{p/2}dt \\
&= \frac{1}{n^{p+1}}\int_0^T \sum_{i=1}^{n}\left(\mathrm{Var}\left(\sum_{j=1}^{n}U_{nij}(t)\right)\right)^{p/2}dt \\
&= \frac{1}{n^{p+1}}\int_0^T \sum_{i=1}^{n}\left(\sum_{j=1}^{n}\mathrm{Var}(U_{nij}(t))\right)^{p/2}dt \\
&= \frac{1}{n^{p+1}}\int_0^T \sum_{i=1}^{n}\left(\sum_{j=1}^{n}\frac{\alpha_{ij}^2(t)}{q_n}\gamma_{ij}(1-q_n\gamma_{ij})\right)^{p/2}dt \\
&\geq C^{p/2}Tn^{-p-1}n^{p/2+1} \geq \frac{C^{p/2}T}{n^{p/2}}.
\end{aligned}$$

Combining this lower-bound with the upper-bounded (45), we get the claimed equivalence.

□

A.2 Approximation theoretic results

In an effort to make this paper more self-contained we briefly recall some results on functional spaces and approximation theory that our work relies on.

L^p spaces embeddings. Since $|\Omega| = 1$, we have the classical inclusion $L^q(\Omega) \subset L^p(\Omega)$ for $1 \leq p \leq q < +\infty$. More precisely

$$\|F\|_{L^p(\Omega)} \leq \|F\|_{L^q(\Omega)} \leq \|F\|_{L^\infty(\Omega)}. \tag{47}$$

We also have the reverse bound

$$\|F\|_{L^p(\Omega)} \leq \|F\|_{L^\infty(\Omega)}^{1-q/p}\|F\|_{L^q(\Omega)}^{q/p}, \tag{48}$$

for any $1 \leq q < p < +\infty$.

Lipschitz spaces $\mathrm{Lip}(s, L^q(\Omega^d))$ [9, Ch. 2, §6 and 9]. We introduce the Lipschitz spaces $\mathrm{Lip}(s, L^q(\Omega^d))$, for $d \in \{1, 2\}$, which contain functions with, roughly speaking, s "derivatives" in $L^q(\Omega^d)$ [9, Ch. 2, Section 9].

Definition A.1. *For $F \in L^q(\Omega^d)$, $q \in [1, +\infty]$, we define the (first-order) $L^q(\Omega^d)$ modulus of smoothness by*

$$\omega(F, h)_q \stackrel{\text{def}}{=} \sup_{\boldsymbol{z}\in\mathbb{R}^d, |\boldsymbol{z}|<h}\left(\int_{\boldsymbol{x},\boldsymbol{x}+\boldsymbol{z}\in\Omega^d}\big|F(\boldsymbol{x}+\boldsymbol{z}) - F(\boldsymbol{x})\big|^q d\boldsymbol{x}\right)^{1/q}. \tag{49}$$

The Lipschitz spaces $\mathrm{Lip}(s, L^q(\Omega^d))$ *consist of all functions* F *for which*

$$\left|F\right|_{\mathrm{Lip}(s,L^q(\Omega^d))} \stackrel{\mathrm{def}}{=} \sup_{h>0} h^{-s}\omega(F,h)_q < +\infty.$$

We restrict ourselves to values $s \in]0,1]$ as for $s > 1$, only constant functions are in $\mathrm{Lip}(s, L^q(\Omega^d))$. It is easy to see that $\left|F\right|_{\mathrm{Lip}(s,L^q(\Omega^d))}$ is a semi-norm. $\mathrm{Lip}(s, L^q(\Omega^d))$ is endowed with the norm

$$\left\|F\right\|_{\mathrm{Lip}(s,L^q(\Omega^2))} \stackrel{\mathrm{def}}{=} \left\|F\right\|_{L^q(\Omega^2)} + \left|F\right|_{\mathrm{Lip}(s,L^q(\Omega^d))}.$$

The space $\mathrm{Lip}(s, L^q(\Omega^2))$ is the Besov space $\mathbf{B}^s_{q,\infty}$ [9, Ch. 2, Section 10] which are very popular in approximation theory. In particular, $\mathrm{Lip}(1, L^1(\Omega^d))$ contains the space $\mathrm{BV}(\Omega^d)$ of functions of bounded variation on Ω^d, i.e. the set of functions $F \in L^1(\Omega^d)$ such that their variation is finite:

$$V_{\Omega^2}(F) \stackrel{\mathrm{def}}{=} \sup_{h>0} h^{-1} \sum_{i=1}^{d} \int_{\Omega^d} \left|F(\boldsymbol{x} + he_i) - F(\boldsymbol{x})\right| d\boldsymbol{x} < +\infty$$

where $e_i, i \in \{1, d\}$ are the coordinate vectors in $\mathbb{R}^d$; see [9, Ch. 2, Lemma 9.2]. Thus Lipschitz spaces are rich enough to contain functions with both discontinuities and fractal structure.

Let us define the piecewise constant approximation of a function $F \in L^q(\Omega^2)$ (a similar reasoning holds of course on Ω) on a partition of Ω^2 into cells $\Omega_{nij} \stackrel{\mathrm{def}}{=} \left\{]x_{i-1}, x_i]\times]y_{j-1}, y_j] : \ (i,j) \in [n]^2\right\}$ of maximal mesh size $\delta \stackrel{\mathrm{def}}{=} \max\limits_{(i,j)\in[n]^2} \max(\left|x_i - x_{i-1}\right|, \left|y_j - y_{j-1}\right|)$,

$$F_n(x,y) \stackrel{\mathrm{def}}{=} \sum_{i,j=1}^{n} F_{nij}\chi_{\Omega_{nij}}(x,y), \quad F_{ij} = \frac{1}{\left|\Omega_{nij}\right|}\int_{\Omega_{nij}} F(x,y)dxdy.$$

Clearly, F_n is nothing but the orthogonal projection of F on the n^2-dimensional subspace of $L^q(\Omega^2)$ defined as

$$\mathrm{Span}\left\{\chi_{\Omega_{nij}} : \ (i,j) \in [n]^2\right\}.$$

Lemma A.2. *There exists a positive constant* C_s*, depending only on* s*, such that for all* $F \in \mathrm{Lip}(s, L^q(\Omega^d))$*,* $d \in \{1,2\}$*,* $s \in]0,1]$*,* $q \in [1,+\infty]$*,*

$$\left\|F - F_n\right\|_{L^q(\Omega^d)} \leq C_s \delta^s \left|F\right|_{\mathrm{Lip}(s,L^q(\Omega^d))}. \tag{50}$$

Proof . Using the general bound [9, Ch. 7, Theorem 7.3] for the error in spline approximation, and in view of Definition A.1, we have

$$\left\|F - F_n\right\|_{L^q(\Omega^d)} \leq C_s\omega(F,\delta)_q = C\delta^s(\delta^{-s}\omega(F,\delta)_q) \leq C_s\delta^s\left|F\right|_{\mathrm{Lip}(s,L^q(\Omega^d))}.$$

□

An immediate consequence is the following result.

Lemma A.3. *Assume that $F \in L^\infty(\Omega^d) \cap \mathrm{Lip}(s, L^q(\Omega^d))$, $d \in \{1,2\}$, $s \in]0,1]$, $q \in [1,+\infty]$, and let $p \in]1,+\infty[$. Then there exists a positive constant $C(p,q,s)$, depending on p, q and s such that*

$$\left\|F - F_n\right\|_{L^p(\Omega^d)} \leq C(p,q,s)\delta^{s\min(1,q/p)}. \tag{51}$$

Proof . We have

$$\left\|F - F_n\right\|_{L^p(\Omega^d)} \leq \begin{cases} \left\|F - F_n\right\|_{L^q(\Omega)} \leq C\left|F\right|_{\mathrm{Lip}(s,L^q(\Omega))}\delta^s, & \text{if} \quad q \geq p; \\ \left\|F - F_n\right\|_{L^\infty(\Omega^d)}^{1-q/p}\left\|F - F_n\right\|_{L^q(\Omega^d)}^{q/p} \leq C\left(2\left\|F\right\|_{L^\infty(\Omega)}\right)^{1-q/p}\left|F\right|_{\mathrm{Lip}(s,L^q(\Omega^d))}^{q/p}\delta^{sq/p} & \\ \qquad \text{otherwise}, \end{cases}$$

where we used (47) (resp. (48)) and Lemma A.2 in the first (resp. second) case. □

Acknowledgement. This work was supported by the ANR grant GRAPHSIP. JF was partly supported by Institut Universitaire de France. JF and AE would like to acknowledge support within the EU grant No. 777826, the NoMADs project.

References

[1] F. Andreu, J. Mazón, J. Rossi, and J. Toledo. A nonlocal p-laplacian evolution equation with neumann boundary conditions. *Journal de Mathématiques Pures et Appliquées*, 90(2):201 – 227, 2008.

[2] F. Andreu-Vaillo, J. M. Mazón, J. D. Rossi, and J. J. Toledo-Melero. *Nonlocal Diffusion Problems*, volume 165 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence (R.I.), 2010.

[3] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Struct. Algorithms*, 31(1):3–122, Aug. 2007.

[4] B. Bollobás and O. Riordan. Metrics for sparse graphs. In S. Huczynska, J. D. Mitchell, and C. M. E. Roney-Dougal, editors, *Surveys in Combinatorics 2009*, London Mathematical Society Lecture Note Series, pages 211–288. Cambridge University Press, Cambridge, 2009.

[5] B. Bollobás and O. Riordan. Sparse graphs: Metrics and random models. *Random Structures & Algorithms*, 39(1):1–38, 2011.

[6] C. Borgs, J. Chayes, L. Lovász, V. Sós, and K. Vesztergombi. Convergent sequences of dense graphs i: Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219(6):1801 – 1851, 2008.

[7] C. Borgs, J. Chayes, L. Lovász, V. Sós, and K. Vesztergombi. Limits of randomly grown graph sequences. *European Journal of Combinatorics*, 32(7):985 – 999, 2011.

[8] A. Buades, B. Coll, and J.-M. Morel. Neighborhood filters and PDEs. *Numerische Mathematik*, 105(1):1–34, 2006.

[9] R. A. DeVore and G. G. Lorentz. *Constructive Approximation*, volume 303 of *Grundlehren der mathematischen*. Springer-Verlag, Berlin Heidelberg, 1993.

[10] A. Elmoataz, X. Desquesnes, and O. Lezoray. Non-local morphological pdes and p-laplacian equation on graphs with applications in image processing and machine learning. *IEEE Journal of Selected Topics in Signal Processing*, 6(7):764–779, 2012.

[11] A. Elmoataz, M. Toutain, and D. Tenbrinck. On the p-laplacian and ∞-laplacian on graphs with applications in image and data processing. *SIAM Journal on Imaging Sciences*, 8(4):2412–2451, 2015.

[12] P. Erdös and A. Rényi. On the Evolution of Random Graphs. *Publication of The Mathematical Institute of The Hungarian Academy of Sciences*, 5:17–61, 1960.

[13] L. C. Evans. *Partial Differential Equations.* American Mathematical Society, Providence (R.I.), 2010.

[14] E. N. Gilbert. Random Graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.

[15] G. Gilboa and S. Osher. Nonlocal linear image regularization and supervised segmentation. *Multiscale Modeling & Simulation*, 6(2):595–630, 2007.

[16] G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. *Multiscale Modeling & Simulation*, 7(3):1005–1028, 2009.

[17] Y. Hafiène, J. Fadili, and A. Elmoataz. Nonlocal p-laplacian evolution problems on graphs. *SIAM Journal on Numerical Analysis*, 56(2):1064–1090, 2018.

[18] R. Ibragimov and S. Sharakhmetov. The exact constant in the Rosenthal inequality for random variables with mean zero. *Theory of Probability and Its Applications*, 46(1):127–132, 2002.

[19] D. Kaliuzhnyi-Verbovetskyi and G. Medvedev. The semilinear heat equation on sparse random graphs. *SIAM Journal on Mathematical Analysis*, 49(2):1333–1355, 05 2017.

[20] S. Kindermann, S. Osher, and P. W. Jones. Deblurring and denoising of images by nonlocal functionals. *Multiscale Modeling & Simulation*, 4(4):1091–1115, 2005.

[21] L. Lovász. *Large Networks and Graph Limits*, volume 60. American Mathematical Society, 2012.

[22] L. Lovász and B. Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933 – 957, 2006.

[23] P. Massart. *Concentration inequalities and model selection*, volume 1896 of *Ecole d'Eté de Probabilités de Saint-Flour XXXIII - 2003*. Springer Verlag, New York, 2007.

[24] G. S. Medvedev. The nonlinear heat equation on dense graphs. *SIAM Journal on Mathematical Analysis*, 46(4):2743–2766, 2014.

[25] G. S. Medvedev. The nonlinear heat equation on W-random graphs. *Arch. Ration. Mech. Anal.*, 212(3):781–803, 2014.

[26] G. S. Medvedev. The continuum limit of the kuramoto model on sparse random graphs. *arxiv e-prints*, 2018.

[27] R.-D. Reiss. *Approximate Distributions of Order Statistics with Applications to Nonparametric Statistics.* Springer-Verlag, New York, New York, 1989.