Normandie Université

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité MATHEMATIQUES

Préparée au sein de l'Université de Caen Normandie

Continuum limits of evolution and variational problems on graphs

Présentée et soutenue par
Yosra HAFIENE

Thèse soutenue publiquement le 05/12/2018 devant le jury composé de		
M. MARTIN BURGER	Professeur, Université Alexander Erlangen -Allemagne	Rapporteur du jury
Mme CAROLA-BIBIANE SCHONLIEB	Professeur, Université de Cambridge - Royaume Uni	Rapporteur du jury
M. JEAN-FRANCOIS AUJOL	Professeur des universités, Université de Bordeaux	Membre du jury
M. ANTONIN CHAMBOLLE	Directeur de recherche au CNRS, ECOLE POLYTECHNIQUE	Membre du jury
Mme JULIE DELON	Professeur des universités, UNIVERSITE PARIS 5 UNIVERSITE PARIS DESC	Membre du jury
M. NOUREDDINE IGBIDA	Professeur des universités, UNIVERSITE LIMOGES	Président du jury
M. JALAL FADILI	Professeur des universités, 14 ENSI de Caen	Directeur de thèse
M. ABDERRAHIM ELMOATAZ BILLAH	Professeur des universités, UNIVERSITE CAEN NORMANDIE	Co-directeur de thèse

Thèse dirigée par JALAL FADILI et ABDERRAHIM ELMOATAZ BILLAH, Groupe de recherche en informatique, image, automatique et instrumentation

UNIVERSITÉ CAEN NORMANDIE

GREYC

To my beloved parents.

To my dear uncle Faycel.

To Marwen.

To my beloved sisters, brother and nieces.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor Professor Jalal Fadili whose expertise, consistent guidance, ample time spent and boundless help made this thesis a reality. His creative ideas and critical thinking have driven me throughout my PhD study. From him, I not only learned a way of thinking and problem solving , but more importantly not to give up quickly!

Meanwhile, besides being a great mentor, he is also a great friend to me. I really appreciate the conversations we had, the suggestions and understandings he provided and his support especially during my difficult times. Thank you so much, Jalal, for forcing me to look at research and my work in different ways and for opening my mind. Thank you, for everything.

I'm highly indebted to my professor Frédéric Richard thanks to whom I was introduced to Jalal and had the chance to do my PhD with him. It was one of the best experiences of my life.

I would like to express my thanks to my co-supervisor professor Abderrahim Elmoataz for sharing his knowledge and for his support and advices.

I would like also to thank Christophe Chesneau for his assistance and cooperation.

Over the years of my PhD, I was lucky to meet lots of great members and friends in GREYC Lab. Thank you François, Sovann, Thoraya, Sonia, Zeina, Abdelhak, Ali, Shivang, Arbi, Bilel, Imed, Tony, Ryan for being part of my daily life. I wish them a brilliant future in their research journey. Khadija, I am very happy to have met you and become friends with you. I wish you the best!

I would like to thank the secretaries Arielle Perette and Gaëlle Lenogue (GREYC, ENSICAEN), and Marie Meleux (UNICAEN), for all the care they have taken of me.

I'm thankful to my great friend and collaborator Samir from the university Aix-Marseille for all the help and encouragement he gave to me. The fantastic year I spent in the Institute of Mathematics in Marseille, is where all this became possible. I really miss the days in Chateau Gombert with you and I wish you the best!

My gratitude goes out to all those who tried to encourage and support me even with one word or prayer. Thank you Wafa for pushing me to work against all odds and especially to have woken me up every morning for a very long time!

Most importantly, I want to thank my family for their years of support, encouragement, endless love and for giving me the reasons to smile during the tough times. My beloved parents, sisters, brother and nieces and all my extended family, you just cheer me up! Marwen, I feel so blessed having you in my life. I love you all, you mean everything to me.

Projet GRAPHSIP co-financé par l'Union européenne, la Région Normandie dans le cadre du programme opérationnel FEDER/FSE 2014-2020 et l'ANR (ANR-14-CE27-0001)

Abstract

The nonlocal p-Laplacian operator, the associated evolution equation and variational regularization, governed by a given kernel, have applications in various areas of science and engineering. In particular, they are modern tools for massive data processing (including signals, images, geometry), and machine learning tasks such as classification. In practice, however, these models are implemented in discrete form (in space and time, or in space for variational regularization) as a numerical approximation to a continuous problem, where the kernel is replaced by an adjacency matrix of a graph. Yet, few results on the consistency of these discretization are available. In particular it is largely open to determine when do the solutions of either the evolution equation or the variational problem of graph-based tasks converge (in an appropriate sense), as the number of vertices increases, to a well-defined object in the continuum setting, and if yes, at which rate. In this manuscript, we lay the foundations to address these questions.

Combining tools from graph theory, convex analysis, nonlinear semigroup theory and evolution equations, we give a rigorous interpretation to the continuous limit of the discrete nonlocal p-Laplacian evolution and variational problems on graphs. More specifically, we consider a sequence of (deterministic) graphs converging to a so-called limit object known as the *graphon*. If the continuous p-Laplacian evolution and variational problems are properly discretized on this graph sequence, we prove that the solutions of the sequence of discrete problems converge to the solution of the continuous problem governed by the graphon, as the number of graph vertices grows to infinity. Along the way, we provide a consistency/error bounds. In turn, this allows to establish the convergence rates for different graph models. In particular, we highlight the role of the graphon geometry/regularity. For random graph sequences, using sharp deviation inequalities, we deliver nonasymptotic convergence rates in probability and exhibit the different regimes depending on p, the regularity of the graphon and the initial data.

Keywords: nonlocal diffusion, nonlocal regularization, p-Laplacian, graphs, graphon, graph limits, numerical approximation, error bound, convergence rate, convex analysis.

Résumé

L'opérateur du p-Laplacien nonlocal, l'équation d'évolution et la régularisation variationnelle associées régies par un noyau donné ont des applications dans divers domaines de la science et de l'ingénierie. En particulier, ils sont devenus des outils modernes pour le traitement massif des données (y compris les signaux, les images, la géométrie) et dans les tâches d'apprentissage automatique telles que la classification. En pratique, cependant, ces modèles sont implémentés sous forme discrète (en espace et en temps, ou en espace pour la régularisation variationnelle) comme approximation numérique d'un problème continu, où le noyau est remplacé par la matrice d'adjacence d'un graphe. Pourtant, peu de résultats sur la consistence de ces discrétisations sont disponibles. En particulier, il est largement ouvert de déterminer quand les solutions de l'équation d'évolution ou du problème variationnel des tâches basées sur des graphes convergent (dans un sens approprié) à mesure que le nombre de sommets augmente, vers un objet bien défini dans le domaine continu, et si oui, à quelle vitesse. Dans ce manuscrit, nous posons les bases pour aborder ces questions.

En combinant des outils de la théorie des graphes, de l'analyse convexe, de la théorie des semi-groupes nonlinéaires et des équations d'évolution, nous interprétons rigoureusement la limite continue du problème d'évolution et du problème variationnel du p-Laplacien discrets sur graphes. Plus précisément, nous considérons une suite de graphes (déterministes) convergeant vers un objet connu sous le nom de *graphon*. Si les problèmes d'évolution et variationnel associés au p-Laplacien continu nonlocal sont discrétisés de manière appropriée sur cette suite de graphes, nous montrons que la suite des solutions des problèmes discrets converge vers la solution du problème continu régi par le graphon, lorsque le nombre de sommets tend vers l'infini. Ce faisant, nous fournissons des bornes d'erreur/consistance.

Cela permet à son tour d'établir les taux de convergence pour différents modèles de graphes. En particulier, nous mettons en exergue le rôle de la géométrie/régularité des graphons. Pour les séquences de graphes aléatoires, en utilisant des inégalités de déviation (concentration), nous fournissons des taux de convergence nonasymptotiques en probabilité et présentons les différents régimes en fonction de p, de la régularité du graphon et des données initiales.

Mots-clés: Diffusion nonlocale, régularisation nonlocale, p-Laplacien, graphes, graphons, limites de graphes, approximation numérique, borne d'erreur, vitesse de convergence, analyse convexe.

Table of contents

Chapter 1

Introduction

Contents

1.1 Context, motivations and objectives

1.1.1 Context

In recent years, evolutions based on Partial Differential Equations (PDEs) have shown to provide very effective tools in various fields throughout science and engineering such as signal/image processing, machine learning, computer vision and biology [10, 38, 88, 11, 58, 7]. Indeed, many problems to handle end up solving an evolution problem involving different kinds of operators depending on the tasks to carry out. Such PDE-based methods have the advantages of better mathematical modeling, connections with physics and better geometrical approximations. Differential operators involved in these PDEs are classically based on local derivatives, that reflect local interactions in the data. Recently, nonlocal counterparts have been proposed in the context of image processing to design gradient-based regularization functionals and PDEs associated with their minimization [63] for many image processing tasks, such as denoising, deconvolution, segmentation, inpainting, optical-flow and more. Following ideas from graph theory, it has been shown that many PDE-based processes, minimizations and computation methods can be generalized to the nonlocal setting. A main advantage for image processing is the ability to process both structures (geometrical parts) and textures within the same framework.

Among other operators, the nonlocal p-Laplacian operator have become more and more popular both in the setting of Euclidean domains and on discrete graphs, as the p-Laplacian problem has been possessing many important features shared by many practical problems in mathematics, physics, engineering, biology, and economy, such as continuum mechanics, phase transition phenomena, population dynamics [6, 7]. Some closely related applications can be found in image processing, such as spectral clustering [34], computer vision and machine learning [45, 48, 72, 2, 99]. This operator is defined on $L^p(\Omega)$ for a bounded domain Ω, $p \in [1, +\infty]$, being a set-valued mapping for $p = 1$ and $p = \infty$, as follows

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

(**A.1**) $\Omega \subset \mathbb{R}$ is a bounded domain, without loss of generality $\Omega = [0, 1]$.

(**A.2**) $K(\cdot,\cdot)$ is a symmetric, non-negative and bounded function on Ω^2.

It can be seen as the nonlocal analogue of the p-Laplacian operator defined on $W^{1,p}(\Omega)$ for $p \in [1, +\infty[$, being also a set-valued mapping for $p = 1$ and $p = \infty$, as

$$\mathbf{\Delta}_p(u(x)) = \operatorname{div}(\left|\nabla u(x)\right|^{p-2}\nabla u(x)),$$

which occurs also in many mathematical models and physical processes such as nonlinear diffusion/filtration and non-Newtonian flows [19].

The nonlinear diffusion problem (Cauchy problem), known as the nonlocal p-Laplacian evolution problem with homogeneous Neumann boundary conditions [7] associated to $\mathbf{\Delta}_p^K(\cdot)$ is

$$\begin{cases} u_t(x,t) = \frac{\partial}{\partial t}u(x,t) = -\mathbf{\Delta}_p^K(u(x,t)), & a.e.\ x \in \Omega, t > 0, \\ u(x,0) = g(x), & a.e.\ x \in \Omega. \end{cases} \tag{$\mathcal{P}_{\text{nloc}}$}$$

The nonlocal diffusion equation shares many properties with the corresponding local problem. If the kernel K is properly rescaled, it has been shown in [5] that problem ($\mathcal{P}_{\text{nloc}}$) converges strongly in $L^\infty((0,T); L^p(\Omega))$ to the well-known local p-Laplacian evolution equation

$$\begin{cases} u_t(x,t) = \frac{\partial}{\partial t}u(x,t) = \mathbf{\Delta}_p(u(x,t)), & a.e.\ x \in \Omega, t > 0, \\ u(x,0) = g(x), & a.e.\ x \in \Omega, \end{cases} \tag{$\mathcal{P}_{\text{loc}}$}$$

which corresponds for $p = 2$ to the heat equation $u_t(x,t) = \mathbf{\Delta}u(x,t)$, while the extreme case, $p = 1$, corresponds to the total variation flow with homogeneous Neumann boundary conditions. The problem ($\mathcal{P}_{\text{loc}}$) occurs also in many applications such as physics, biology or economy [77, 44].

Particularly, if $K(x, y) = J(x - y)$, where the kernel $J : \Omega \to \mathbb{R}$ is a nonnegative continuous radial function with compact support verifying $J(0) > 0$ and $\int_\Omega J(x)dx = 1$, nonlocal evolution equations of the form

$$u_t(x,t) = J * u(x,t) - u(x,t) = \int_\Omega J(x-y)(u(y,t) - u(x,t))dy, \qquad (\mathcal{P}^*_{\text{nloc}})$$

where $*$ stands for the convolution, have many applications in modeling diffusion processes [6, 14, 15, 35, 57, 110, 56]. As stated in [57], in modeling the dispersal of organisms in space when $u(x, t)$ is their density at the point x at time t, $J(x - y)$ is considered as the probability distribution of jumping from position y to position x, then, the expression $J * u - u$ represents transport due to long-range dispersal mechanisms, that is the rate at which organisms are arriving to location x from any other place. The integration is in Ω, imposing consequently that diffusion takes place only in Ω, there is no flux of individuals across the boundary, from where comes the nonlocal analogue to Neumann boundary conditions.

The evolution problem ($\mathcal{P}_{\text{nloc}}$) can also be interpreted as the gradient flow associated to the Dirichlet energy

$$R_p(v, K) = \frac{1}{2p}\int_{\Omega^2} K(x,y)\left|v(y) - v(x)\right|^p dydx, \qquad (1.1.1)$$

which is the nonlocal analogue to the energy functional $\frac{1}{p}\int_\Omega |\nabla v|^p$ associated to the local p-Laplacian.

On the other hand, in the context of image processing, smoothing and denoising are key filtering processes. Among the existing methods, the variational ones, based on regularization, provide a general framework to design such efficient filter processes. Solutions of variational models can be obtained by minimizing appropriate energy functions: an empirical loss plus a regularization term. The minimization is usually performed by a descent method designed to solve the corresponding Euler-Lagrange equations. In the nonlocal setting, the resulting discrete schemes are closely linked to an important category of neighborhood filters which have shown their efficiency to better preserve fine and repetitive image structures than local ones [76, 31]. Nonlocal regularization problems are much more powerful on real world processing data than local ones due to their self-similarity and long range dependence. Among these variational problems, the nonlocal variational p-Laplacian problem has become more and more popular in the context of image processing for nonlocal (patch-based) regularization of inverse problems and in data processing in graphs. This problem is defined as minimizing the sum of a data fidelity term and a regularization term associated to the nonlocal energy functional (1.1.1), i.e;

$$\min_{u \in L^2(\Omega)} \left\{ E_\lambda(u, g, K) \stackrel{\text{def}}{=} \frac{1}{2\lambda}\left\|u - g\right\|^2_{L^2(\Omega)} + R_p(u, K) \right\}, \qquad (\mathcal{VP}_{\text{nloc}})$$

$\lambda \in]0, +\infty[$ is a regularization parameter specifying the trade-off between the two competing terms.

1.1.2 Motivations

In many real-world problems, data can be represented on a graph. Each vertex of the graph corresponds to a datum, and the edges encode the pairwise relationships or similarities among the data. A typical example of graph data is the web. The vertices are just the web pages, and the edges denote the hyperlinks. In market basket analysis, the items also form a graph by connecting any two items which have appeared in the same shopping basket. For the particular case of images, pixels (represented by nodes) have a specific organization expressed by their spatial connectivity. Therefore, a typical graph used to represent images is a grid graph. For the particular case of unorganized data, a graph can also be associated with by modeling neighborhood relationships between the data elements.

For these practical reasons, recently, there has been a surge of interest in adapting and solving nonlocal PDEs such as ($\mathcal{P}_{\text{nloc}}$) and variational problems such as ($\mathcal{VP}_{\text{nloc}}$) on data which is given by

Figure 1.1: Examples of images that can be represented by weighted graphs as their natural representation.

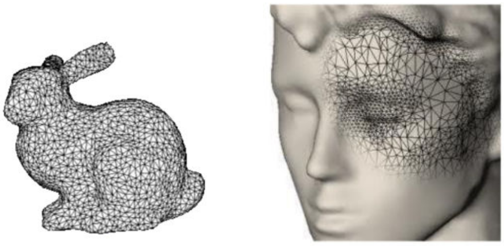


Figure 1.2: Examples of meshes that can be represented by weighted graphs as their natural representation.

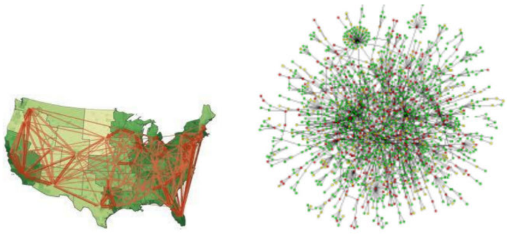


Figure 1.3: Examples of networks that can be represented by weighted graphs as their natural representation.

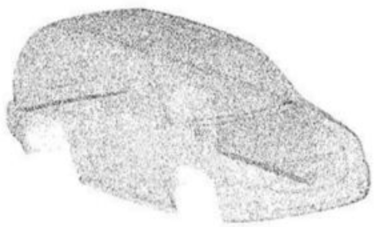


Figure 1.4: Example of point clouds/unorganized data that can be represented by weighted graphs.

arbitrary graphs and networks, since the data in practice is discrete, graphs constitute a natural structure suited to their representation. Using this framework, problems are directly expressed in a discrete setting. This way to proceed encompasses local and nonlocal methods in the same framework by using appropriate graphs topologies and edge weights depending on the data structure and the task to be performed. The demand for such methods is motivated by existing and potential future applications [46, 49]. These practical considerations lead naturally to a discrete time and space approximation of ($\mathcal{P}_{\text{nloc}}$) and a space approximation of ($\mathcal{VP}_{\text{nloc}}$) encoded by the structure of the graph. So that these discrete problems can be applied in the same way to images, meshes or data of any size by simply adapting the topology of the graph and the weight function. Motivated by these practical considerations, much work has been done constructing and analyzing the discrete analogue of the nonlocal continuous evolution/regularization for the p-Laplacian operator on graphs. The proposed framework works on any discrete data represented by weighted graphs which allows to take into account the nonlocal interactions in the data by explicitly introducing discrete nonlocal derivatives and functionals on graphs of arbitrary topologies, to transcribe the continuous setting.

Before going deeper into details, Let us see an example to illustrate the use of the nonlocal p-Laplacian evolution and regularization problems to deal with image processing tasks such as semi-supervised segmentation and denoising relying on the nonlocal heat equation (2-Laplacian) by analyzing the evolution equation in the continuous setting and then discretizing it on an appropriate graph structure to get the desired result. An interesting advantageous of such a method/algorithm is the connection between denoising and segmentation, where the same flow is used for both tasks and only the initial conditions are different (see more details in [62, Section 5]).

(a) Original image (b) Segmented image

Figure 1.5: Segmentation of a textured image by a nonlocal graph. The first column presents the original image with the initial markers super-imposed. The second one presents the result of the segmentation via a nonlocal graph.

Coming back to our discrete analysis. For that, let us consider a partition (not necessarily uniform) $\{t_h\}_{h=1}^N$ of the time interval $[0,T]$. Let $\tau_{h-1} \stackrel{\text{def}}{=} |t_h - t_{h-1}|$ and the maximal size $\tau = \max_{h\in[N]} \tau_h$. A fully discrete counterpart (in space and Forward-Euler in time) of ($\mathcal{P}_{\text{nloc}}$) on a given graph G_n is then given by

$$\begin{cases} \dfrac{u_i^h - u_i^{h-1}}{\tau_{h-1}} = \dfrac{1}{n}\displaystyle\sum_{j=1}^n K_{nij}\left|u_j^{h-1} - u_i^{h-1}\right|^{p-2}(u_j^{h-1} - u_i^{h-1}), \quad (i,h) \in \{1,\cdots,n\} \times \{1,\cdots,N\}, \\ \\ u_i(0) = g_i^0, \quad i \in \{1,\cdots,n\}. \end{cases} \tag{$\mathcal{P}^d_{\text{nloc}}$}$$

Similarly, that of ($\mathcal{VP}_{\text{nloc}}$) is given by

$$\min_{u_n\in\mathbb{R}^n} \left\{ E_{n,\lambda} \stackrel{\text{def}}{=} \frac{1}{2\lambda n}\left\|u_n - g_n\right\|_2^2 + R_{n,p}(u_n, K_n) \right\}, \tag{$\mathcal{VP}^d_{\text{nloc}}$}$$

where

$$R_{n,p}(u_n, K_n) \stackrel{\text{def}}{=} \frac{1}{2n^2 p}\sum_{i,j=1}^n K_{nij}\left|u_{nj} - u_{ni}\right|^p. \tag{1.1.2}$$

K_{nij} can be seen as the adjacency matrix of the graph G_n.

The discrete nonlocal problems ($\mathcal{P}^d_{\text{nloc}}$) and ($\mathcal{VP}^d_{\text{nloc}}$) are just approximations of the underlying continuous problems. Thus, the following legitimate questions have to be answered seperately for each problem:

- (**Q1**) What is the structure of the solution of the discrete problem? is there any continuous limit (as $n \to +\infty$) at all? If yes, in what sense?
- (**Q2**) What is the rate of convergence to this limit and what is its relation to the unique (strong) solution of ($\mathcal{P}_{\mathrm{nloc}}$)/the unique global minimizer of ($\mathcal{VP}_{\mathrm{nloc}}$)?
- (**Q3**) What are the parameters involved in this convergence and what is their influence in the corresponding rate?
- (**Q4**) Can this continuum limit help us get better insight into discrete models/algorithms and design new ones?

In the literature, numerous works have been carried out in the recent years attempting to answer some of these questions. However most of them focus only on certain specific problems. As a consequence, their results are rather limited and cannot be extended to complicated general cases.

1.1.3 Objectives

The main objectives of this work is to answer all questions (**Q1**)-(**Q4**) above for both the nonlocal evolution and variational problems. We begin first by studying the nonlocal p-Laplacian evolution problem ($\mathcal{P}_{\mathrm{nloc}}$). In Chapter 3, we answer question (**Q1**): we study the convergence and stability properties of the numerical solutions of the general discrete problem and give a general error estimate. Based on this error bound, in Chapter 4, we give the rate of convergence to the continuous limit for different graph models and specify the parameters involved in this rate and show their influence, which answers questions (**Q2**)-(**Q3**). Secondly, we turn to study the nonlocal p-Laplacian variational problem ($\mathcal{VP}_{\mathrm{nloc}}$). In Chapter 6, we give a general error estimate for the discrete problem ($\mathcal{VP}^d_{\mathrm{nloc}}$). Next, in Chapter 7, we specify the assumptions under which we are able to answer in detail questions (**Q2**)-(**Q3**).

1.2 Main contributions

1.2.1 The p-Laplacian evolution problem on graphs

Our first main result, which is at the heart of Chapter 3, establishes a general error bound for the fully discretized p-Laplacian evolution problem (in space and time) using forward and backward Euler schemes, respectively. This bound allows to deal with networks on convergent graph sequences and prove the convergence of ($\mathcal{P}^d_{\mathrm{nloc}}$) to ($\mathcal{P}_{\mathrm{nloc}}$) and provide the corresponding rate to answer (**Q1**), (**Q2**) and (**Q3**).

1.2.1.1 A digest of main results

For the nonlocal p-Laplacian evolution problem, we prove the following results:

- (i) Kobayashi type estimates: error estimates to compare two trajectories corresponding to the p-Laplacian evolution problem governed by two different kernels and initial data.
- (ii) Consistency and error estimates of the numerical solutions to the fully discretized problem valid uniformly for $t \in [0, T]$, $T > 0$.
- (iii) Application to dynamical networks on simple and weighted graphs: convergence of discrete approximations on deterministic and random inhomogenous graphs to a continuum limit (governed by graphons).
- (iv) We quantify the corresponding convergence rates and we reveal the role of the data geometry/regularity and of p on these rates.

I - General error bound: Kobayashi type estimates. We consider the *forward* Euler time-discrete approximation to ($\mathcal{P}_{\text{nloc}}$). The space approximation is seen through the use of the subscript n to emphasize the fact that we use a kernel and an initial data depending on n. For that, we take again the time partition mentioned previously

$$\begin{cases} \dfrac{u_n^h(x) - u_n^{h-1}(x)}{\tau_{h-1}} = -\boldsymbol{\Delta}_p^{K_n}(u_n^{h-1}(x)), & a.e.\ x \in \Omega, h \in \{1, \cdots, N\}, \\ u_n^0(x) = g_n^0(x), \quad a.e.\ x \in \Omega. \end{cases} \qquad (\mathcal{P}^f_{\text{nloc},\tau})$$

First, we prove that ($\mathcal{P}^f_{\text{nloc},\tau}$) is well-posed (i.e; starting from $g_n^0(x) \in L^\infty(\Omega)$, there exists a unique accumulation point to the iterates of ($\mathcal{P}^f_{\text{nloc},\tau}$)). Besides the forward Euler scheme, we prove the same result for the *backward* Euler scheme.We consider a time-continuous extension of u_n^h obtained by a time linear interpolation as follows

$$\check{u}_n(x,t) = \frac{t_h - t}{\tau_{h-1}} u_n^{h-1}(x) + \frac{t - t_{h-1}}{\tau_{h-1}} u_n^h(x), \quad t \in]t_{h-1}, t_h], \quad x \in \Omega. \qquad (1.2.1)$$

We prove the following theorem.

Theorem 1.2.1. *Suppose $p \in]1, +\infty[$, $g, g_n^0 \in L^\infty(\Omega)$ and K, K_n are measurable, symmetric and bounded mappings.*

Let u be the unique solution of problem ($\mathcal{P}_{\text{nloc}}$), and $\check{u}_n$ is built as in (1.2.1) from the time-discrete approximation $u_n^h(x)$ defined in ($\mathcal{P}^f_{\text{nloc},\tau}$). Then

$$\left\|u - \check{u}_n\right\|_{C(0,T;L^p(\Omega))} \leq C\left(\left\|g_n - g_n^0\right\|_{L^p(\Omega)} + \left\|g - g_n\right\|_{L^p(\Omega)} + \left\|K - K_n\right\|_{L^p(\Omega^2)}\right) + O(\tau), \qquad (1.2.2)$$

where the constant C is independent of n.

$C(0, T; L^p(\Omega))$ denotes the space of uniformly time continuous functions with values in $L^p(\Omega)$ endowed with the norm $\|\cdot\|_{C(0,T;L^p(\Omega))} \stackrel{\text{def}}{=} \sup_{t\in[0,T]} \|\cdot\|_{L^p(\Omega)}$.

We also obtain convergence in $L^p(\Omega)$ for both time continuous and totally discretized problems. Convergence in $L^2(\Omega)$ norm is thus a corollary. We obtain these results without any extra regularity assumption. In Chapter 4, we apply the above result to analyze the convergence rates of networks on deterministic/random convergent graph sequences as summarized here after.

II - Convergence rates for networks on deterministic graph sequences For networks on simple graph sequences, we show the convergence of the discrete solution to the continuous solution. We provide the corresponding convergence rate. We show how the accuracy of the approximation depends on the regularity of the boundary of support of the graphon.

In addition, for weighted graphs, we give a precise error estimate under the mild assumption that both the kernel K and the initial data g are in Lipschitz spaces, which in particular contain functions of bounded variation (these spaces will be detailed later on in Section 2.3).

Corollary 1.2.2. *Suppose that $p \in]1, +\infty[$, $K : \Omega^2 \to [0, 1]$ is a symmetric and measurable function in $\mathrm{Lip}(s, L^p(\Omega^2))$, and $g \in \mathrm{Lip}(s, L^p(\Omega)) \cap L^\infty(\Omega)$, $s \in]0, 1]$. Then*

$$\left\|u - \check{u}_n\right\|_{C(0,T;L^p(\Omega))} \leq O(n^{-s}) + O(\tau). \qquad (1.2.3)$$

If $\mathrm{Lip}(s, L^p(\Omega^2))$ is replaced with $\mathrm{BV}(\Omega^2)$, then the rate becomes

$$\left\|u - \check{u}_n\right\|_{C(0,T;L^p(\Omega))} \leq O(n^{-1/p}) + O(\tau). \qquad (1.2.4)$$

For this graph model, we also study the limit as $p \to \infty$ and we prove that solutions to the semidiscrete scheme converge uniformly to a nonlocal evolution problem.

III - Convergence rates for networks on random graph sequences. Using sophisticated deviation inequalities, we prove non-asymptotic convergence and give the rate of convergence of the discrete solution to its continuous limit as the number of vertices $n \to \infty$.

To get the corresponding convergence rate, a supplementary assumption is added regarding the kernel K and the initial data g, that is belonging to $\mathrm{Lip}(s', L^q(\Omega^2))$ and $\mathrm{Lip}(s, L^q(\Omega))$, respectively. This measure allows us to identify different asymptotic regimes ($n \to +\infty$) depending on the values of p and the parameters s, s' and q .

Theorem 1.2.3. *Suppose that $p \in]1, +\infty[$, $K \in L^\infty(\Omega^2) \cap \mathrm{Lip}(s', L^q(\Omega^2))$ is a symmetric and measurable mapping with $q_n \|K\|_{L^\infty(\Omega^2)} \leq 1$ and $g \in L^\infty(\Omega) \cap \mathrm{Lip}(s, L^q(\Omega))$, $s, s' \in]0, 1]$. Let $\theta \stackrel{\mathrm{def}}{=} \min(s, s') \min(1, q/p)$. Then, for $T > 0$, there exists a positive constant C, such that for any $\beta > 0$ and $t \in]0, e[$*

$$\|u - \check{u}_n\|_{C(0,T;L^p(\Omega))} \leq C \left(\left(\beta \frac{\log(n)}{n} + \frac{\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)}{n^{p/2}} \right)^{1/p} + \left(\frac{t \log(n)}{n}\right)^{\theta} \right) + O(\tau), \tag{1.2.5}$$

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

1.2.1.2 Relation to previous work

A general error bound. Concerning previous work for this model, the authors of [91] have already obtained a similar conclusion under different but complementary assumptions. Indeed, they dealt with the problem ($\mathcal{P}_{\mathrm{nloc}}$) in which only the case $K(x, y) = J(x - y)$ was treated. First, they considered a semi-discretization in space of this problem using a non-uniform partition of Ω. They showed that the solutions of the obtained ODE system converge uniformly to the continuous one as the mesh size goes to zero. Secondly, by discretizing also the time variable (using only the forward Euler scheme) and presenting a totally discrete method, they showed that solutions to the numerical scheme converge uniformly to the continuous solution as the mesh size and the time step go to zero. The uniform convergence they establish, however, imposes the positivity of the solution which is a stringent assumption. Furthermore, no error estimate was provided in [91], only asymptotic convergence was supplied. Our results are much stronger since we provide a general error estimate that allows us to get an L^p-norm convergence. We go further by adressing both forward and backward Euler schemes.

Networks on convergent graph sequences. Another closely related and important work dealing with networks on graphs is that in [83, 84, 70] which paved the way to study limit phenomena of evolution problems on both deterministic and random (dense and sparse) graphs. In [83], the author focused on a nonlinear (nonlocal) heat evolution equation on graphs, where the operator $\boldsymbol{\Delta}_p^K$ was replaced by the operator $\boldsymbol{D}^W : u \in L^2(\Omega) \to -\int_\Omega W(x, y)D(u(y) - u(x))dy$, with $W(\cdot, \cdot)$ verifying Assumption (**A.2**) and in which the function D was assumed Lipschitz-continuous. This assumption was essential to prove well-posedness (existence and uniqueness follow immediately from the Cauchy Lipschitz Theorem), as well as to study the consistency in L^2-norm of the spatial semi-discrete approximation on simple and weighted graph sequences. Though this seminal work was quite inspiring to us, it differs from our work in many crucial aspects. First, the nonlocal p-Laplacian evolution problem at hand is different and cannot be covered by [83] where the function $x \mapsto x|x|^{p-2}$ lacks Lipschitzianity for $p \in]1, +\infty[$, and thus raises several challenges (including well-posedness and error estimates). Our results on Kobayashi-type estimates are also novel and are of independent interest beyond problems on networks. We also consider both the semi-discrete and fully-discrete versions with both forward and backward Euler approximations, that we fully characterize. For networks on random graphs, in [84] the author dealt with networks on dense random graphs. Again, he considered only the spatial semi-discrete scheme for which he showed the convergence in probability of discrete solutions to the continuous one relying on the central limit theorem (CLT). Thus, those results are asymptotic and no convergence rate was provided. Our result goes much beyond this work by considering a more general random graph model (the dense model is then just a particular case) and by exploiting sophisticated deviation inequalities that permitted us not only to prove nonasymptotic bounds of the error between the discrete model and the continuum one, but also to quantify the corresponding convergence rate.

1.2.2 The p-Laplacian variational problem on graphs

Turning to the variational problem, our major result which is the main of Chapter 6 establishes a general error bound for the discretized p-Laplacian variational problem. This result answers question (**Q1**). By exploiting this general error bound combined with a key regularity result of the solution that we also provide in Chapter 6, we deal in Chapter 7 with networks on convergent graph sequences and prove the convergence of ($\mathcal{VP}^d_{\text{nloc}}$) to ($\mathcal{VP}_{\text{nloc}}$) as well as quantify the corresponding convergence rate.

1.2.2.1 A digest of main results

For the nonlocal p-Laplacian variational problem, we prove the following results:

(i) General (L^2-norm) error estimate to compare the unique solution of the discrete problem ($\mathcal{VP}^d_{\text{nloc}}$) and the one of the continuum one ($\mathcal{VP}_{\text{nloc}}$).

(ii) Application to dynamical networks on simple and weighted graphs: capitalizing on (i), we show convergence of discrete approximations on deterministic and random inhomogeneous graphs to a continuum limit (governed by graphons).

(iii) We quantify the corresponding convergence rates and we reveal the role of the data geometry/regularity on these rates.

I -A general error estimate. We begin by studying the consistency of ($\mathcal{VP}_{\text{nloc}}$) in which we investigate functionals with a nonlocal regularization term corresponding to the p-Laplacian operator. We first give a general error estimate controlling the convergence and regularity properties of the numerical solutions for the general discrete variational problem ($\mathcal{VP}^d_{\text{nloc}}$). Under the assumption $p \in [1, +\infty[$, as $n \to +\infty$, we prove that the solution to this problem, that can be regarded as a discrete approximation of the initial problem via the kernel and the initial data discretization, converges to a nonlocal variational problem. In addition, we obtain convergence in the L^2 norm. We obtain these results without any extra regularity assumption.

Theorem 1.2.4. *Suppose that $g \in L^2(\Omega)$ and K is a nonnegative measurable, symmetric and bounded mapping. Let $u^\star$ and $u^\star_n$ be the unique minimizers of ($\mathcal{VP}_{\text{nloc}}$) and ($\mathcal{VP}^d_{\text{nloc}}$), respectively. Then, we have the following error bound.*

(i) If $p \in [1, 2]$, then

$$\left\|I_n u^\star_n - u^\star\right\|^2_{L^2(\Omega)} \leq C\left(\left\|g - I_n g_n\right\|^2_{L^2(\Omega)} + \left\|g - I_n g_n\right\|_{L^2(\Omega)} + \left\|K - I_n K_n\right\|_{L^{\frac{2}{2-p}}(\Omega^2)} + \left\|u^\star - I_n P_n u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)}\right), \tag{1.2.6}$$

where C is a positive constant independent of n.

(ii) If $\inf_{(x,y)\in\Omega^2} K(x,y) \geq \kappa > 0$, then for any $p \in [1, +\infty[$,

$$\left\|I_n u^\star_n - u^\star\right\|^2_{L^2(\Omega)} \leq C\left(\left\|g - I_n g_n\right\|^2_{L^2(\Omega)} + \left\|g - I_n g_n\right\|_{L^2(\Omega)} + \left\|K - I_n K_n\right\|_{L^\infty(\Omega^2)} + \left\|u^\star - I_n P_n u^\star\right\|_{L^p(\Omega)}\right), \tag{1.2.7}$$

where C is a positive constant independent of n.

As we do for the evolution problem, the solution of ($\mathcal{VP}^d_{\text{nloc}}$) being discrete, to be able to compare it with the continuum one of ($\mathcal{VP}_{\text{nloc}}$), we are in need of an intermediate function (also for the continuum correspondings of K_{nij} and g_n), that is why we define the injector I_n and the projector P_n to get this intermediate continuum function (these operators are defined in details in Chapter 6).

II - Convergence rates for networks on deterministic graph sequences. Secondly, we apply these results, using the graph limits theory, to dynamical networks on simple and weighted dense graphs to show that the approximation of minimizers of the discrete problems on simple and weighted graph sequences converge to those of the continuous problem. Specifically, for simple graph sequences, we show how the accuracy of the approximation depends on the regularity of the boundary of support of the graphon. For networks on weighted graphs we give a precise error estimate under the mild assumption that both the kernel K and the initial data g are in Lipschitz spaces, $\text{Lip}(s, L^q(\Omega))$, $\text{Lip}(s', L^q(\Omega^2))$, respectively.

Theorem 1.2.5. *Let $p \in [1, 2[$, and assume that $g \in L^\infty(\Omega) \cap \text{Lip}(s, L^q(\Omega))$, with $s \in]0, 1]$ and $q \in [2/(3-p), 2]$. Suppose moreover that $K \in \text{Lip}(s', L^{q'}(\Omega^2))$, $(s', q') \in]0, 1] \times [1, +\infty[$ and $K(x, y) = J(|x-y|)$, $\forall (x, y) \in \Omega^2$, with J a nonnegative bounded measurable mapping on Ω. Let $u^\star$ and $u^\star_n$ be the unique minimizers of ($\mathcal{VP}_{\text{nloc}}$) and ($\mathcal{VP}^d_{\text{nloc}}$), respectively. Then, the following error bounds hold.*

$$\left\|I_n u^\star_n - u^\star\right\|^2_{L^2(\Omega)} \leq C n^{-\min\{sq/2, s', s'q'(1-p/2)\}}. \tag{1.2.8}$$

where C is a positive constant independent of n.

III - Convergence rates for networks on random graph sequences. Then, relying on the same error estimate, we study networks on inhomogeneous random graphs. More precisely, using sophisticated deviation inequalities, we prove convergence and give the rate of convergence of the discrete solution to its continuous limit with high probability under the same assumptions as for deterministic graphs on the Kernel K and the initial data g.

Theorem 1.2.6. *Suppose that $p \in [1, 2[$, $g \in L^2(\Omega)$ and K is a nonnegative measurable, symmetric and bounded mapping. Let $u^\star$ and $u^\star_n$ be the unique minimizers of ($\mathcal{VP}_{\text{nloc}}$) and ($\mathcal{VP}^d_{\text{r,nloc}}$), respectively. Let $p' = \frac{2}{2-p}$.*

(i) There exist positive constants C and C_1 that do not depend on n, such that for any $\beta > 0$

$$\begin{aligned}\left\|I_n u^\star_n - u^\star\right\|^2_{L^2(\Omega)} \leq C \Bigg(& \left(\beta \frac{\log(n)}{n} + \frac{\max\left(q_n^{-(p'-1)}, q_n^{-p'/2}\right)}{n^{p'/2}}\right)^{1/p'} + \left\|g - I_n g_n\right\|^2_{L^2(\Omega)} \\ & + \left\|g - I_n g_n\right\|_{L^2(\Omega)} + \left\|K - I_n \hat{K}^{\mathbf{X}}_n\right\|_{L^{p'}(\Omega^2)} + \left\|u^\star - I_n P_n u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)} \Bigg),\end{aligned} \tag{1.2.9}$$

with probability at least $1 - 2n^{-C_1 \min\left(q_n^{(2p'-1)}, q_n^{p'}\right)\beta}$.

(ii) Assume moreover that $g \in L^\infty(\Omega) \cap \text{Lip}(s, L^q(\Omega))$, with $s \in]0, 1]$ and $q \in [2/(3-p), 2]$, that $K(x, y) = J(|x-y|)$, $\forall (x, y) \in \Omega^2$, with J a nonnegative bounded measurable mapping on Ω, that $K \in \text{Lip}(s', L^{q'}(\Omega^2))$, $(s', q') \in]0, 1] \times [p', +\infty]$ and $q_n\left\|K\right\|_{L^\infty(\Omega^2)} \leq 1$. Then there exist positive constants C and C_1 that do not depend on n, such that for any $\beta > 0$ and $t \in]0, e[$

$$\left\|I_n u^\star_n - u^\star\right\|^2_{L^2(\Omega)} \leq C \left(\left(\beta \frac{\log(n)}{n} + \frac{\max\left(q_n^{-(p'-1)}, q_n^{-p'/2}\right)}{n^{p'/2}}\right)^{1/p'} + \left(t \frac{\log(n)}{n}\right)^{\min(sq/2, s')} \right), \tag{1.2.10}$$

with probability at least $1 - \left(2n^{-C_1 \min\left(q_n^{(2p'-1)}, q_n^{p'}\right)\beta} + n^{-t}\right)$.

1.2.2.2 Relation to previous work

Nonlocal neighborhood filters. Since the work of Buades and Morel [31] on image filtering by non-local means, several recent works have shown the interest of introducing non-local regularization functions to take into account a more complex interactions and introduce more flexibility in the regularization functions [8, 33, 61, 36, 108]. Kindermann, Osher and Jones [103] interpreted non local means and neighborhood filters as regularization based on non local functionals. Gilboa and Osher [62] have proposed a non local quadratic functional of weighted differences for image regularization and semi-supervised segmentation. These works can be considered as the non local analogues of Total Variation models for image regularization. Most of the proposed regularization processes have been proposed in the context of image processing where images are considered as continuous functions on continuous domains. Then, one considers a continuous energy functional which is classically solved by the corresponding Euler-Lagrange equation or its associated gradient flow. However, the discretization of the underlying differential operators is difficult for high dimensional data and for image and data defined on irregular domains.

Networks on graphs. In [59] the authors studied the consistency of a variational problem given in terms of minimizing a functional corresponding to the total variation on random graphs. They looked at the limit of the discrete total variation on graphs representing point clouds as the number of data points goes to infinity. The limit was considered in the Γ-convergence sense [29]. Based on this result, in [99], the authors considered a discrete p-Laplacian regularization problem on random geometric graphs to carry out a semi-supervised learning task. Their aim was to assign real-valued labels to a set of n sample points, provided a small training subset of N labeled points. To do so, they investigated a family of regression problems and studied the asymptotic behavior when the number of unlabeled points increases. To solve the regression problem, they considered a discrete objective functional discretized in an appropriate way to encode the structure of the graph. Relying on tools of calculus of variations and optimal transportation, they showed the (locally) uniform convergence of minimizers of these nonlinear functionals in random discrete setting to the minimizers of the continuum energy functional corresponding to the local p-Laplacian operator. These results on asymptotic behavior of minimizers do not provide any error estimates for finite n.

For local variational problems, the authors of [109] have studied the numerical approximation of the Rudin-Osher-Fatemi image smoothing model consisting of minimizing the following energy functional

$$E(v) \stackrel{\text{def}}{=} \frac{1}{2\lambda}\|u-g\|^2_{L^2(\Omega)} + |v|_{\mathrm{BV}(\Omega)},$$

$|v|_{\mathrm{BV}(\Omega)}$ denotes the bounded variation seminorm. They bound the difference between the continuous solution and the solutions to various finite-difference approximations to this model. They give a bound of the L^2-norm of the difference between these two solutions.

However, to the best of our knowledge, there is no rigorous study of numerical approximations for the nonlocal variational problem ($\mathcal{VP}_{\mathrm{nloc}}$).

1.3 Organisation of the manuscript

This manuscript consists of two parts and eight chapters.

Chapter 2: This chapter collects the necessary mathematical material used throughout the manuscript.

Chapter 3: In this chapter, we present our main result: the global error estimate for the discrete p-Laplacian evolution problem. Our results include two main parts: the consistency of the time-continuous problem (Theorem 3.3.1) and the consistency of the time discrete problem (Theorem 3.4.4). We end up this chapter by a brief discussion of the relation of our estimates to Kobayashi type estimates.

Chapter 4: In this chapter we present our results on networks on convergent graph sequences. We apply our result on the consistency of the p-Laplacian evolution problem to networks on convergent graph sequences. We deal first with deterministic dense graphs (simple and weighted graphs). In Section 4.4, we generalize the above analysis to cover networks on random inhomogeneous graphs.

Chapter 5: In this chapter, we deal with the normalized p-Laplacian evolution problem on graphs. We recall first the basic definitions and properties of the normalized p-Laplacian operator on graphs. Next, in Section 5.4, we study the consistency of its associated diffusion problem. We finish this chapter by showing some experiments related to data processing (filtering images/3D point clouds) to illustrate the use of this operator.

Chapter 6: In this chapter, we present our main result for the variational p-Laplacian problem: the error bound for the discrete problem. We also provide a key regularity result that will be useful for the next chapter dealing with networks on convergent graph sequences.

Chapter 7: In this chapter, we present our results on networks on convergent graph sequences for the nonlocal variational p-Laplacian problem. Doing the same way as for the evolution problem, we deal first with dense deterministic graphs (simple and weighted) and then with random inhomogeneous graphs in Section 7.3.

Chapter 8: This last chapter summarizes our contributions and draws important conclusions. It also discusses several interesting perspectives and open problems.

Chapter 2

Mathematical Background

Contents

In this chapter, we collect the necessary mathematical material used in the manuscript.

Let $\mathbb{R}$ denote the set of real numbers, $\mathbb{R}^+$ the set of nonnegative reals, $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ the extended real line and $\mathbb{R}^n$ the n-dimensional real Euclidean space. We denote by $\mathbb{N}$ the set of non-negative integers, by $\mathbb{N}^*$, the set of positive integers. We use the notation $[n] = \{1, \cdots, n\}$. For a set $\mathcal{C}$, $|\mathcal{C}|$ denotes its cardinality.

2.1 Tools from graph limits theory

We present some definitions and 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 [80, 25] and then further developed in a series of papers by Borgs et al. [23, 24]. 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'.

2.1.1 Preliminaries

An undirected *graph* is a pair $G = (V(G), E(G))$ satisfying $E(G) \subset V(G) \times V(G)$. $V(G)$ stands for the set of vertices (or nodes, or points), each node $i \in V(G)$ is an abstract representation of an element of the data structure represented by the graph. $E(G)$ denotes the edges (or lines) set and is composed of pairs of vertices (i, j). An edge represents the connection between two vertices. It is said then that these two vertices are adjacent, or neighbors which is denoted by $i \sim j$. In this manuscript, we consider graphs without loops or parallel edges in which the edges are symmetric (these kind of graphs are called *simple*). We can therefore define the set $E(G)$ such that:

$$E(G) \stackrel{\text{def}}{=} \{(i, j) \in V(G) \times V(G) | i \sim j \quad \text{and} \quad i \neq j\}. \tag{2.1.1}$$

The usual way to picture a graph is by drawing a dot (or a cercle) for each vertex and joining two of these dots by a line if the corresponding two vertices form an edge. Just the way these dots and lines are drawn is irrelevant: all that matters is the information which pairs of vertices form an edge and which do not.

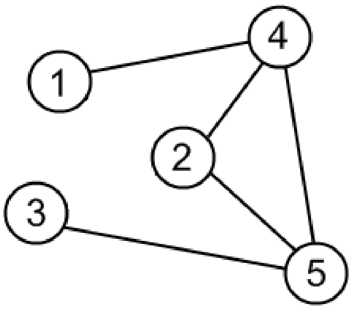
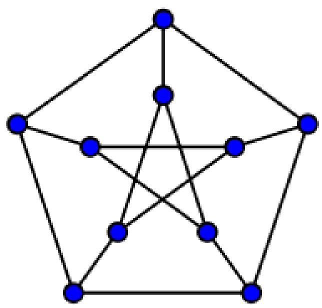


Figure 2.1: Example of an undirected simple graph G with $V(G) = \{1, \cdots, 5\}$ nodes with edge set $E(G) = \{(1, 4), (4, 2), (4, 5), (5, 3)\}$.

For a graph G, the adjacency matrix is a square $|V(G)| \times |V(G)|$ matrix such that its elements indicate whether pairs of vertices are adjacent or not in the graph. In the special case of a finite simple graph, the adjacency matrix is a $(0, 1)$-matrix with zeros on its diagonal since edges from a vertex to itself (loops) are not allowed in simple graphs. If the graph is undirected, the adjacency matrix is symmetric. A non-standard way of visualizing graphs using another version of the adjacency matrix is the so-called *pixel picture*. On the left of Figure 2.2 we see a graph (the Petersen graph). In the middle, we see its adjacency matrix. On the right, we see another version of its adjacency matrix, where the 0's are replaced by white pixels and the 1's are replaced by black pixels.The whole picture is on the unit square.



0	1	0	0	1	1	0	0	0	0
1	0	1	0	0	0	1	0	0	0
0	1	0	1	0	0	0	1	0	0
0	0	1	0	1	0	0	0	1	0
1	0	0	1	0	0	0	0	0	1
1	0	0	0	0	0	0	1	1	0
0	1	0	0	0	0	0	0	1	1
0	0	1	0	0	1	0	0	0	1
0	0	0	1	0	1	1	0	0	0
0	0	0	0	1	0	1	1	0	0

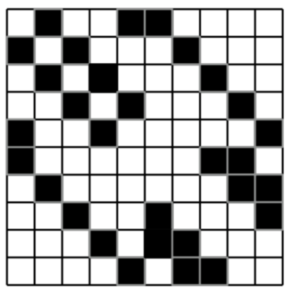


Figure 2.2: The Petersen graph, its adjacency matrix, and its pixel picture.

A weighted graph G is a graph with weight $\beta(i, j)$ associated to each edge (i, j). We see in Figure 2.3 a picture of a weighted graph. The weight function represents the similarity between the vertices of the graph. It is defined as $\beta : V(G) \times V(G) \to \Omega \subset \mathbb{R}^+$ (we restrict ourselves to the values in $\Omega = [0, 1]$ in this manuscript). Since we are dealing with undirected graphs, the weight function is symmetric: $\forall (i, j) \in V(G)^2, \beta(i, j) = \beta(j, i)$. The adjacency matrix of a weighted graph is obtained by replacing the 1's in the adjacency matrix by the weights of the edges. An unweighted graph is a weighted graph where all the edge weights are 1.

For more information on graphs we refer the reader to [43].

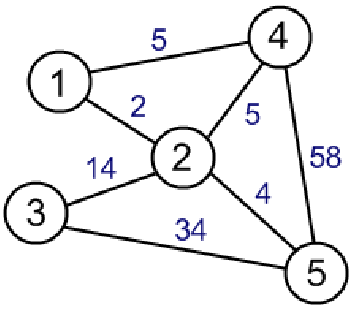


Figure 2.3: Example of weighted graph with $V(G) = \{1, \cdots, 5\}$ nodes with edge set $E(G) = \{(1,2),(2,3),(1,4),(4,2),(4,5),(5,3)\}$ and $\{2,\ 4,\ 5,\ 5,\ 14,\ 34,\ 58\}$ are weights assigned to edges.

2.1.2 Convergence of graph sequences

Let $G_n = (V(G_n), E(G_n))$, $n \in \mathbb{N}^*$, be a sequence of dense, i.e., $|E(G_n)| = O(|V(G_n)|^2)$ finite, and simple graphs.

For two simple (unweighted) graphs F and G, $\hom(F,G)$ indicates the number of homomorphisms (adjacency-preserving maps) from $V(F)$ to $V(G)$. This number is normalized to get the *homomorphism density*

$$t(F,G) = \frac{\hom(F,G)}{|V(G)|^{|V(F)|}}.$$

This quantity can be interpreted as the probability that a random map of $V(F)$ into $V(G)$ is a homomorphism.

This notion is extended to weighted graphs. To every homomorphism $\phi : V(F) \to V(G)$, we let

$$\hom_\phi(F,G) \stackrel{\text{def}}{=} \prod_{(i,j)\in E(F)} \beta_G(\phi(i),\phi(j)).$$

Then the homomorphism function is defined by

$$\hom(F,G) = \sum_{\phi:V(F)\to V(G)} \hom_\phi(F,G)$$

and the homomorphism density as defined for simple graphs

$$t(F,G) = \frac{\hom(F,G)}{|V(G)|^{|V(F)|}}.$$

Suppose that the number of nodes of G_n tends to infinity. Suppose that the graphs G_n become more and more similar in the sense that $t(F,G_n)$ tends to a limit $t(F)$ for every simple graph F. Based on this, the following notion of convergence is defined

Definition 2.1.1. The sequence of graphs $\{G_n\}_{n\in\mathbb{N}^*}$ is called convergent if $t(F,G_n)$ is convergent for every simple graph F.

Remark 2.1.2. Note that $t(F,G_n) = O(1)$ if $|E(G_n)| = O(|V(G_n)|^2)$ so that this definition is meaningful only for sequences of dense graphs and otherwise the limit is 0 for every simple graph F

with at least one edge. In the theory of graph limits, convergence in Definition 2.1.1 is called *left-convergence*. Since this is the only convergence of graph sequences that we use, we would refer to the left-convergent sequence as convergent (see [23, Section 2.5]).

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 $\Omega_1^{(n)}, \cdots, \Omega_n^{(n)}$, and for every $x \in \Omega_i^{(n)}$, $y \in \Omega_j^{(n)}$ define

$$K_{G_n}(x,y) = \begin{cases} 1 & \text{if} \quad (i,j) \in E(G_n), \\ 0 & \text{otherwise.} \end{cases} \tag{2.1.2}$$

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

$$K_{G_n}(x,y) = \begin{cases} \beta(i,j) & \text{if} \quad (i,j) \in E(G_n), \\ 0 & \text{otherwise.} \end{cases} \tag{2.1.3}$$

This construction is not unique, however given a graph, the set of pixel kernels arising via (2.1.2) can be considered to be equivalent via the weakly isomorphic relation (to be defined shortly).

Convergent graph sequences have a limit object, which can be represented as a measurable function. Let $\mathcal{K}$ denote the space of all measurable bounded functions $K : [0,1] \to \mathbb{R}$ such that $K(x,y) = K(y,x)$ for all $x, y \in [0,1]$. We also define $\mathcal{K}_0 = \{K \in \mathcal{K} : 0 \leq K \leq 1\}$. The functions of this space are called ***graphons***.

The main motivation behind introducing graphons is that they provide a much more explicit representation for this limit object as the following theorem shows.

Theorem 2.1.3 ([25, Theorem 2.1]). (i) *For every convergent graph sequence $\{G_n\}_{n\in\mathbb{N}^*}$, there is a function $K \in \mathcal{K}_0$ such that $t(F, G_n) \to t(F, K)$ for every simple graph F, i.e.,*

$$t(F, G_n) \to t(F, K) \overset{def}{=} \int_{\Omega^{V(F)}} \prod_{(i,j)\in E(F)} K(x_i, x_j)dx. \tag{2.1.4}$$

(ii) *This function K is uniquely determined up to measure-preserving transformation in the following sense: for every other limit K' there are measure-preserving maps $\phi, \psi : \Omega \to \Omega$ such that $K(\phi(x), \phi(y)) = K'(\psi(x), \psi(y))$.*

(iii) *Every function $K \in \mathcal{K}_0$ arises as the limit of a convergent graph sequence, i.e., for every $K \in \mathcal{K}_0$, there is a sequence of graphs $\{G_n\}_{n\in\mathbb{N}^*}$ satisfying (2.1.4).*

Remark 2.1.4. The above theorem gives a result of existence and uniqueness of the limit but it is not a constructive result. In fact, there is a natural "limit object" in the form of a symmetric measurable function $K : \Omega \to \Omega^2$ which arises as a limit of an appropriate graph sequence but this limit is not expilicitly known for every graph sequence.

Remark 2.1.5. The homomorphism density (for the graphon) $t(F, K)$ defined in (2.1.4) can be seen as the extension of the homomorphism density on graphs. We can think of the interval Ω as the set of nodes, and of the value $K(x, y)$ as the weight of the node (x, y). Then the formula

$$t(F, K) = \int_{\Omega^{V(F)}} \prod_{(i,j)\in E(F)} K(x_i, x_j) \prod_{i\in V(F)} dx_i$$

is an infinite analogue of weighted homomorphism numbers.

We now introduce the cut norm which is used to construct the cut distance and define convergence for graph sequences. In fact, an appropriate notion of distance between two arbitrary, possibly different number of nodes graphs can be defined, such that convergent sequences are Cauchy in this metric and vice versa. The completion of the metric space of graphs relative to this metric can be described, and its elements, i.e., limit objects for convergent graph sequences, can be characterized in various ways. Let K be a graphon, the *cut norm* of K is defined by

$$\|K\|_{\Box} \stackrel{\text{def}}{=} \sup_{S,T\in\mathcal{L}_{\Omega}} \Big|\int_{S\times T} K(x,y)dxdy\Big|,$$

where $K \in L^1(\Omega^2)$ and $\mathcal{L}_{\Omega}$ stands for the set of all Lebesgue measurable subsets of Ω. The cut norm is a norm; this is easy to prove using standard arguments. Given a graphon K and a map ϕ from Ω to Ω, we define the ϕ *pull-back* of K by $K^{\phi} : (x,y) \mapsto K(\phi(x),\phi(y))$. Let $S(\Omega)$ denote the set of measure-preserving maps from Ω into Ω. Then the *cut distance* between two graphons K and W is defined by

$$\mathrm{d}_{\Box}(K,W) \stackrel{\text{def}}{=} \inf_{\phi\in S(\Omega)} \|K - W^{\phi}\|.$$

This 'distance' function is only a pseudo-distance function since different graphons can have distance zero. This issue can be rectified by considering the quotient space of weakly isomorphic graphons (to be shortly defined) (see also [78, Sections 8.2 and 10.7] for more details).

Definition 2.1.6 ([23, Theorem 2.6]). The sequence of dense graphs $\{G_n\}_{n\in\mathbb{N}^*}$ is said to *converge* if $\{K_{G_n}\}_n$ is a Cauchy sequence with respect to the cut distance.

An interesting consequence of this definition is that the space of graphs, or equivalently pixel kernels, is not closed under the cut distance. The space of graphons (larger than the space of graphs) defines the completion of this space.

Informally a graphon can be thought of as a generalization of the adjacency matrix of a (weighted) graph which has a continuum number of vertices. It should be noted that the cut norm and cut distance definitions extend naturally to the larger space of graphons. Hence, geometrically, the graphon K can be interpreted as the limit of K_{G_n} defined in (2.1.2) (and (2.1.3)) for the cut-norm.

We now define convergence of a graph sequence to a kernel via the cut distance.

Definition 2.1.7 ([23, Theorem 2.6]). Let $\{G_n\}_{n\in\mathbb{N}^*}$ be a sequence of graphs and let $K \in \mathcal{K}_0$. Then $G_n \to K$ as $n \to \infty$ if and only if $\mathrm{d}_{\Box}(K_{G_n}, K) \to 0$ as $n \to \infty$.

A useful observation that will be used throughout this manuscript is the following:

$$\mathrm{d}_{\Box}(K,W) \le \|K - W\|_{\Box} \le \|K - W\|_{L^1(\Omega^2)} \le \|K - W\|_{L^p(\Omega^2)} \le 1 \quad \forall p \in [1,+\infty]. \tag{2.1.5}$$

Thus, convergence of the sequence of pixel kernels $\{K_{G_n}\}_{n\in\mathbb{N}^*}$ (recall constructions (2.1.2) and (2.1.3)) in the L^p-norm implies the convergence of the graph sequence $\{G_n\}_{n\in\mathbb{N}^*}$ ([25, Theorem 2.3]). In other words, convergence in L^2-norm of $\{K_{G_n}\}_{n\in\mathbb{N}^*}$ is sufficient to prove convergence of a sequence of graphs with respect to Definition 2.1.7.

We now introduce the weakly isomorphic relation, denoted $\approx$, which identifies sets of graphons which all have a cut distance of zero apart [78, Corollary 10.34]. Let K, $W \in \mathcal{K}_0$ be two graphons, we say that K and W are *weakly isomorphic* if and only if

$$\mathrm{d}_{\Box}(K,W) = 0.$$

We can just take the easy way out, and call two graphons K and W weakly isomorphic if $t(F,K) = t(F,W)$ for every simple graph F.

From these definitions, an important consequence (observation) is that every point in the completion is defined by a Cauchy sequence, which tends to a graphon K. Two Cauchy sequences define the same point of the completion if and only if merging them we get a Cauchy sequence, which implies that they have the same limit graphon (up to weak isomorphism). Conversely, every graphon is the limit of a Cauchy sequence and so it corresponds to a point in the completion.

Before we move on to give some illustrative examples, an important remark is in order.

Remark 2.1.8. In this manuscript, we focused only in exposing results of graph sequences convergence with respect to a single metric (the cut-metric) among the metrics defined in [26, 27, 23, 24, 79, 80]. In fact, in these papers, the authors introduced several natural metrics for graphs (we can cite in addition to the cut-metric d_{cut}, the count (or subgraph) metric d_{sup}, and the partition-metric d_{part}), and showed that they are equivalent, in that if $\{G_n\}_{n\in\mathbb{N}^*}$ is a sequence of graphs with $|V(G_n)| \to \infty$, then if $\{G_n\}_{n\in\mathbb{N}^*}$ is Cauchy with respect to one of these metrics then it is Cauchy with respect to all of them.

Example 2.1.9 (Half graphs). Let $G_{n,n}$ denote the bipartite graph on $2n$ nodes $\{1, \cdots, n, 1', \cdots, n'\}$, where i is connected to j' if and only if $i \leq j$. It is easy to see that this sequence is convergent and its limit is the function

$$K(x,y) = \begin{cases} 1, & \text{if} \quad |x-y| \geq 1/2, \\ 0, & \text{otherwise}. \end{cases} \tag{2.1.6}$$

Figure 2.4 shows an example of the half-graph for $n = 16$, its pixel picture and the corresponding graphon.

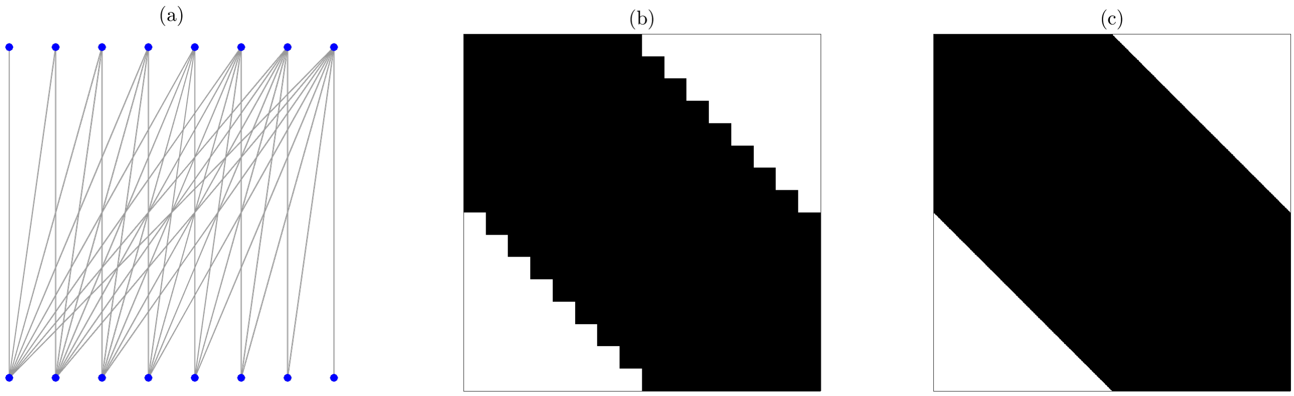


Figure 2.4: (a) A half-graph of 16 vertices. (b) The plot of its pixel picture. (c) The corresponding graphon.

Example 2.1.10 (Nearest neighbor graph). Let $V = [n]$. The nearest neighbor (nn) of i is a point j, $j \neq i$ with minimum distance for a given similarity metric from i . To make the nearest neighbor unique we choose the point j with maximum index in case of ties and denote by $nn(i)$ the set of neighbors of vertex i. By nature, the neighborhood relations of a nn-graph are not necessarily symmetric. In order to preserve the property of symmetry of the edges, we use in this manuscript a symmetric (or reciprocal) version of the nn-graph. In this version, the E set of edges is defined by

$$E(G) \stackrel{\text{def}}{=} \{(i,j) | i \in nn(j) \quad \text{or} \quad j \in nn(i)\}.$$

The nn-graph plays a prominent role in non-local data analysis and processing methods, and in particular in non-local models for image processing. It will then be of particular interest in applications.

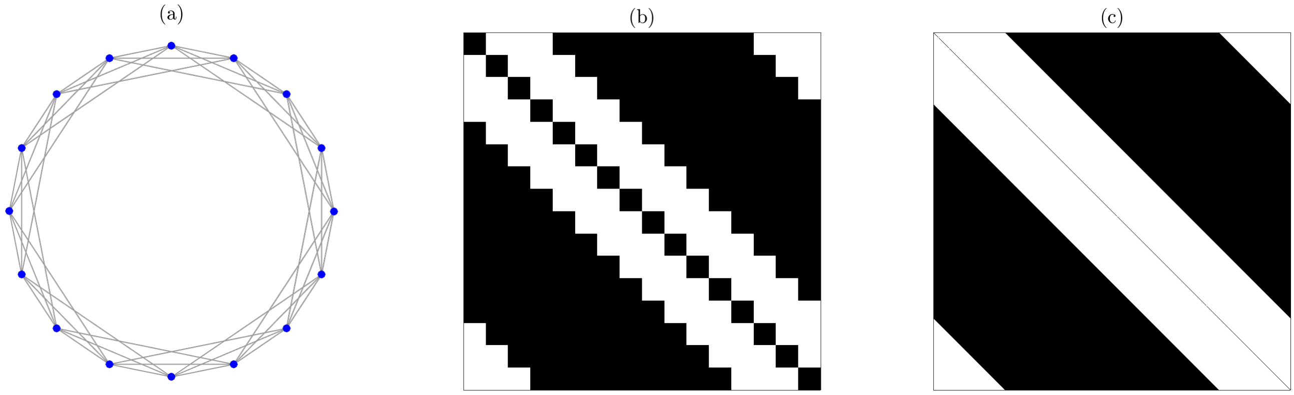


Figure 2.5: (a) A nearest-neighbour graph with 16 vertices. (b) The plot of its pixel picture. (c) The corresponding graphon.

Example 2.1.11 (Simple threshold graphs). These graphs are defined on the set $[n]$ by connecting i and j if and only if $i+j \leq n$. These graphs converge to the graphon defined by $K(x,y) = \mathbb{1}_{(x+y\leq 1)}$, which we call *the simple threshold graphon.*

Figure 2.6 displays an example of the threshold graph for $n = 16$ vertices, its pixel picture and the corresponding graphon.

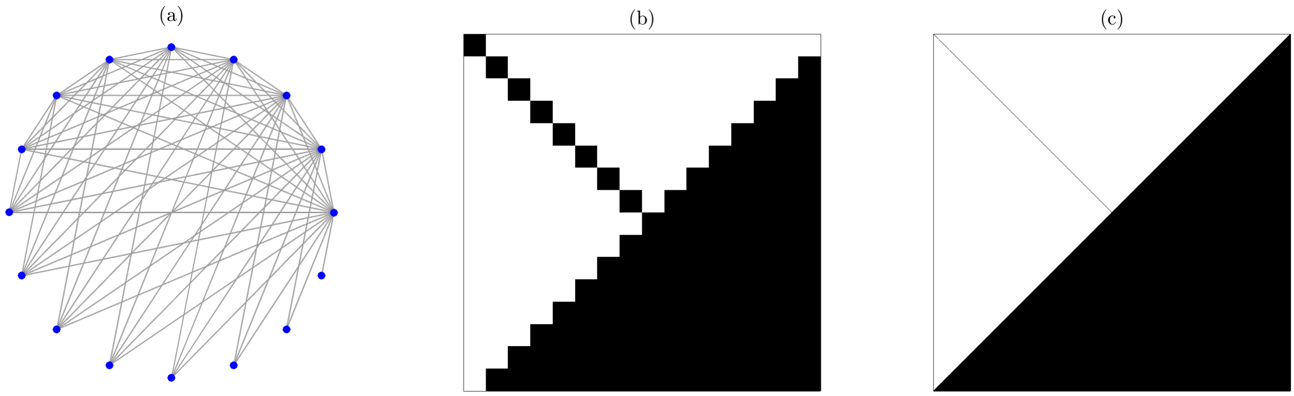


Figure 2.6: (a) A simple-threshold graph with 16 vertices. (b) The plot of its pixel picture. (c) The corresponding graphon.

2.1.3 Determinisitc graph models

In this section, we present the deterministic graph models that will be used in Chapters 4 and 7 when we treat networks on convergent graph sequences. These models are chosen to illustrate our results on the consistency of the nonlocal p-Laplacian evolution and variational problems ($\mathcal{P}_{\mathrm{nloc}}$) and ($\mathcal{VP}_{\mathrm{nloc}}$), respectively. These models are of interest in their own and were constructed in [83].

2.1.3.1 Simple graphs

We fix $n \in \mathbb{N}^*$, divide Ω into n intervals

$$\Omega_1^{(n)} = \left[0, \frac{1}{n}\right[, \Omega_2^{(n)} = \left[\frac{1}{n}, \frac{2}{n}\right[, \ldots, \Omega_j^{(n)} = \left[\frac{j-1}{n}, \frac{j}{n}\right[, \ldots, \Omega_n^{(n)} = \left[\frac{n-1}{n}, 1\right[,$$

and let $\mathcal{Q}_n$ be the partition of Ω, $\mathcal{Q}_n = \{\Omega_i^{(n)}, i \in [n]\}$. Denote $\Omega_{ij}^{(n)} \stackrel{\text{def}}{=} \Omega_i^{(n)} \times \Omega_j^{(n)}$.

First, we consider the case of a sequence of simple graphs converging to $\{0,1\}$-graphon.

We define a sequence of simple graphs $G_n = (V(G_n), E(G_n))$ such that $V(G_n) = [n]$ and

$$E(G_n) = \left\{(i,j) \in [n]^2 : \ \Omega_{ij}^{(n)} \cap \overline{\mathrm{supp}(K)} \neq \emptyset\right\},$$

where

$$\mathrm{supp}(K) = \left\{(x,y) \in \Omega^2 : \ K(x,y) \neq 0\right\}. \tag{2.1.7}$$

As we mentioned before, the kernel K represents the corresponding graphon, that is the limit as $n \to \infty$ of the function $K_{G_n} : \Omega^2 \to \{0,1\}$ such that

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

As $n \to \infty$, $\{K_{G_n}\}_{n \in \mathbb{N}^*}$ converges to the $\{0,1\}$-valued mapping whose support is defined by (2.1.7).

2.1.3.2 Weighted graphs

We now review a more general class of graph sequences. We consider two sequences of weighted graphs generated by a given graphon K.

Let $K : \Omega^2 \to [a,b]$ $a,b > 0$, be a symmetric measurable function which will be used to assign weights to the edges of the graphs considered below.

Next, we define the quotient of K and $\mathcal{Q}_n$ denoted $K/\mathcal{Q}_n$ as a weighted graph with n nodes

$$K/\mathcal{Q}_n = \left([n], [n] \times [n], \hat{K}_n\right).$$

As before, weights $(\hat{K}_n)_{ij}$ are obtained by averaging K over the sets in $\mathcal{Q}_n$

$$(\hat{K}_n)_{ij} = n^2 \int_{\Omega_i^{(n)} \times \Omega_j^{(n)}} K(x,y) dx dy. \tag{2.1.8}$$

The second sequence of weighted graphs is constructed as follows

$$\mathbb{G}(X_n, K) = \left([n], [n] \times [n], \breve{K}_n\right),$$

where

$$X_n = \left\{\frac{1}{n}, \frac{2}{n}, \cdots, 1\right\}, \qquad (\breve{K}_n)_{ij} = K\left(\frac{i}{n}, \frac{j}{n}\right). \tag{2.1.9}$$

One can easily see that $\hat{K}_n$ is the projection of K on the space of piecewise constant functions and $\breve{K}_n$ is nothing but the sampling of K at the vertices of the graph.

2.1.4 Random graphs

The theory of random graphs was founded in the late 1950s and early 1960s by Erdös and Rényi [50], who started the systematic study of the space $\mathcal{G}(n, M)$ of graphs with n labeled vertices and $M = M(n)$ edges, with all graphs equiprobable. Nearly the same time, Gilbert [60] introduced the closely related model $\mathcal{G}(n,p)$ of random graphs on n labeled vertices: a random $G(n,p) \in \mathcal{G}(n,p)$ is obtained by selecting edges independently, each with probability $p = p(n)$. As Erdös and Rényi are the founders of the theory of random graphs, it is not surprising that both $G(n,p)$ and $G(n,M)$ are now known as Erdös-Rényi random graphs.

The theory of random graphs lies at the intersection between graph theory and probability theory. Let V be a set of n points, say $V = [n]$. The aim is to turn the set $\mathcal{G}$ of all graphs on V into a probability space. Intuitively we should be able to generate $G \in \mathcal{G}$ randomly as follows: for each $e \in V \times V$ we decide by some random experiments wether or not e shall be an edge of G, these experiments are performed independently and for each the probability of accepting e as an edge of G is equal to some fixed number $p \in [0,1]$.

Later, Lovász et al. [25] defined a more general random graph model. Given any symmetric measurable function $K : \Omega^2 \to \Omega$ and an integer $n > 0$, we can generate a random graph $G(n, K)$ on node set V as follows. Generate n independent numbers $\mathbf{X}_1, ..., \mathbf{X}_n$ from the uniform distribution on Ω, and then connect nodes i and j with probability $K(\mathbf{X}_i, \mathbf{X}_j)$. As a special case, if K is the constant p-valued function, we get $G(n, p)$. This sequence is convergent almost surely, and in fact it converges to the weighted graph with one node and one loop with weight p.

We now present some canonical examples of graph sequences which converge to a given *graphons*.

Example 2.1.12 (The Erdös-Renyi graphs.). Let $p \in]0,1[$ and consider the sequence of random graphs $G(n,p) = (V(G(n,p)), E(G(n,p)))$ such that $V(G(n,p)) = [n]$ and the probability $\mathbb{P}\{(i,j) \in E(G(n,p))\} = p$ for any $(i,j) \in [n]^2$. Then for any simple graph F, $t(F, G(n,p))$ converges almost surely to $p^{\left|E(F)\right|}$ as $n \to \infty$ [23] and $\{G(n,p)\}$ converges almost surely to the p-constant graphon.

Figure 2.7 shows a realization of the Erdös-Renyi graph model for $n = 16$, its pixel picture and the corresponding graphon.

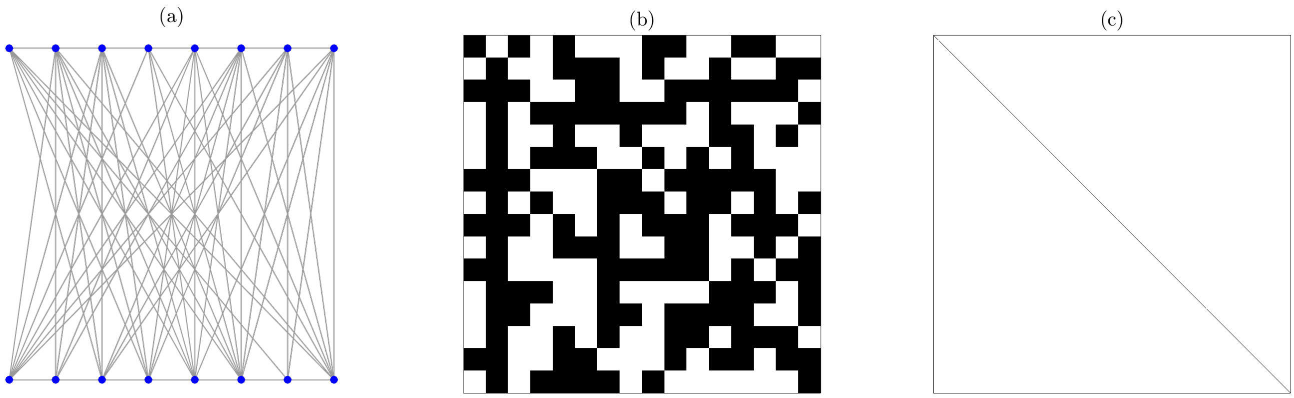


Figure 2.7: (a) A realization of the Erdös-Renyi random graph model with $p = 0.5$. (b) Its pixel picture. (c) The corresponding graphon.

Example 2.1.13 (Uniform attachement graphs). We define a (dense) uniform attachment graph sequence as follows: if we have a current graph G_n with n nodes, then we create a new isolated node, and then for every pair of previously nonadjacent nodes, we connect them with probability $1/n$.

One can prove that with probability 1, the sequence $\{G_n\}_{n \in \mathbb{N}^*}$ has a limit, which is the function $K(x,y) = \min(x,y)$ [78, Proposition 11.40].

Figure 2.8 shows an example of the uniform attatchment graph for $n = 16$, its pixel picture and the corresponding graphon.

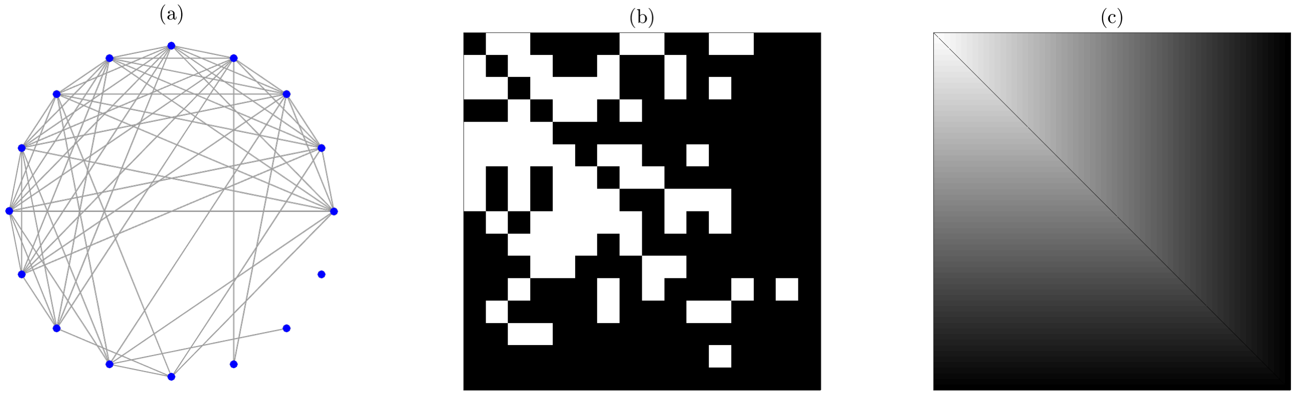


Figure 2.8: (a) A realization of the uniform attachment graph random model. (b) Its pixel picture. (c) The corresponding graphon.

Example 2.1.14 (Small World random graphs). Let $X_n = \{x_1, \cdots, x_n\}$ be a sequence of n points from Ω and let $K \in \mathcal{K}_0$ be a $\{0,1\}$-*graphon*. We assume that K is almost everywhere continuous on Ω^2 and its support has a positive Lebesgue measure. Next, define

$$K_p(x,y) \stackrel{\text{def}}{=} (1-p)K(x,y) + p(1-K(x,y)), \quad p \in [0, 0.5]. \tag{2.1.10}$$

The Small World random graph sequence $G_n([n], E(G_n))$ is constructed as follows. For every $(i,j) \in [n]^2$, $i \neq j$

$$\mathbb{P}((i,j) \in E(G_n)) = K_p(x_i, x_j).$$

The decision whether to include (i,j) to $E(G)$ is made independently for each pair $(i,j) \in [n]^2$, $i \neq j$. Note that for $p = 0.5$, this graph becomes the Erdös-Renyi graph with parameter $p = 1/2$.

Figure 2.9 shows an example of the small world random graph for $n = 16$, its pixel picture and the corresponding graphon.

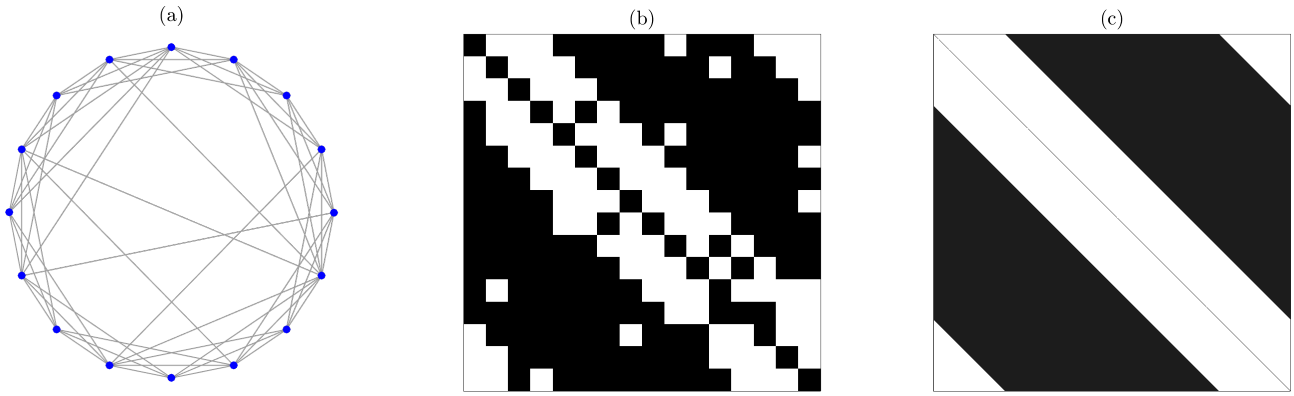


Figure 2.9: (a) A realization of the small world random graph model with $p = 0.1$. (b) Its pixel picture. (c) The corresponding graphon.

2.1.5 The random inhomogeneous graph model

The classical random graph models defined previously (and various other models) 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 $G(n,p)$ or in $G(n,M)$ do not vary very much: their distribution is close to a Poisson distribution. However, many large real-world graphs are highly inhomogeneous. One reason is that the vertices may have been 'born' at different times, with old and new vertices having very different properties. Experimentally, the spread of degrees is often very large. In particular, in many examples the degree distribution follows a power law. In the last few years, this has led to the introduction and analysis of many new random graph models designed to incorporate or explain these features.

We describe in this section the model of inhomogeneous random graphs that will be used throughout. The construction of this inhomogeneous random graph model was proposed in [20, 21, 22].

Definition 2.1.15. Fix $n \in \mathbb{N}^*$ and let K be a symmetric measurable function on Ω^2. Generate the 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 $(\mathbf{X}_1, \cdots, \mathbf{X}_n) \stackrel{\text{def}}{=} \mathbf{X}$ from the uniform distribution on Ω. Let $\left\{\mathbf{X}_{(i)}\right\}_{i\in[n]}$ 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 $(i, j) \in [n]^2$, $i \neq j$,
$$\mathbb{P}\left((i, j) \in E(G_n)|\mathbf{X}\right) = q_n \hat{K}{}^{\mathbf{X}}_{nij}, \tag{2.1.11}$$
where
$$\hat{K}{}^{\mathbf{X}}_{nij} \stackrel{\text{def}}{=} \min\left(\frac{1}{\left|\Omega^{\mathbf{X}}_{nij}\right|}\int_{\Omega^{\mathbf{X}}_{nij}} K(x, y)dxdy, 1/q_n\right), \tag{2.1.12}$$
and
$$\Omega^{\mathbf{X}}_{nij} \stackrel{\text{def}}{=}]\mathbf{X}_{(i-1)}, \mathbf{X}_{(i)}]\times]\mathbf{X}_{(j-1)}, \mathbf{X}_{(j)}] \tag{2.1.13}$$
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 $\mathbf{X}$.

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

Assumption 2.1.16. *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 $K \in L^\infty(\Omega^2)$;*

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

There is no loss of generality in taking 1 in the bound of (A.2).

Although we shall give general results in Sections 4 and 7 that hold under (A.1)-(A.2), it is helpful to bear in mind one particular example of the general class of models we shall study. This example is inspired by the so-called *almost dense (or non uniform)* random graphs (see [21, Section 3.4]).

Proposition 2.1.17. *Suppose $K \in L^\infty(\Omega^2)$ 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 [21, Lemma 3.5 and Lemma 3.8], that were designed for the graph model described in Remark 4.4.1 (given later on in Section 4.4.1), can be adapted to cover our graph model with (2.1.11) 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 [21, Section 2.1] for details about this metric). This shows (A.1). (A.2) is trivially verified. □

Remark 2.1.18. The graph model of Proposition 2.1.17 encompasses the dense random graph model (i.e., with $\Theta(n^2)$ edges) extensively studied in [80, 25], 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 graphs (but not too sparse), i.e., with $o(n^2)$ but $\omega(n)$ edges. For example, one can take $g_n = C\log(n)^{-\delta}$, where $\delta \in]0, 1[$, in which case one has $q_n = \exp(-C\log(n)^{1-\delta}) = o(1)$.

2.2 Tools from analysis

2.2.1 Convex analysis on Hilbert spaces

We here collect some important results from convex analysis which will be used in the up coming chapters. A comprehensive account on convex analysis on Hilbert spaces can be found in [16]. Denote $\mathcal{H}$ a real Hilbert space endowed with inner product $\langle \cdot, \cdot \rangle$ and associated norm $\| \cdot \|$.

Definition 2.2.1 (Convex set). A set $\mathcal{S}$ of $\mathcal{H}$ is *convex*, if

$$\forall x, x' \in \mathcal{S}, \forall t \in]0,1[, \quad tx + (1-t)x' \in \mathcal{S}.$$

Let $\mathcal{S} \subseteq \mathcal{H}$ be a non-empty set and function $F : \mathcal{S} \to \overline{\mathbb{R}}$. The domain of F is

$$\mathrm{dom}(F) \stackrel{\text{def}}{=} \{x \in \mathcal{S} \ : \ F(x) < +\infty\}.$$

F is called *proper* if $-\infty \notin F(\mathcal{S})$ and $\mathrm{dom}(F) \neq \emptyset$.

Definition 2.2.2 (Convex function). A function $F : \mathcal{H} \to \overline{\mathbb{R}}$ is convex if

$$\forall x, x' \in \mathcal{H}, \forall t \in [0,1], \quad F(tx + (1-t)x') \leq tF(x) + (1-t)F(x').$$

Definition 2.2.3 (Strongly convex function). A function $F : \mathcal{H} \to \overline{\mathbb{R}}$ is strongly convex with parameter $m > 0$ if

$$\forall x, x' \in \mathcal{H}, \forall t \in [0,1], \quad F(tx + (1-t)x') \leq tF(x) + (1-t)F(x') - \frac{m}{2}t(1-t)\left\|x - x'\right\|^2.$$

Definition 2.2.4 (Lower semi-continuous function). Given a function $F : \mathcal{H} \to \overline{\mathbb{R}}$ and a point $x \in \mathcal{H}$. F is *lower-semi continuous (lsc)* at x if

$$\lim_{x \to x'} \inf F(x') \geq F(x).$$

The class of *proper*, *convex* and *lsc* functions on $\mathcal{H}$ is denoted as $\Gamma_0(\mathcal{H})$.

Definition 2.2.5 (Indicator function). Let $\mathcal{S} \subseteq \mathcal{H}$ be a convex non-empty closed set, *the indicator function* of $\mathcal{S}$, $i_{\mathcal{S}} \in \Gamma_0(\mathcal{H})$, is defined by

$$i_{\mathcal{S}} = \begin{cases} 0, & \text{if } x \in \mathcal{S}, \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.2.1}$$

Definition 2.2.6 (Subdifferential). Let $F \in \Gamma_0(\mathcal{H})$ the subdifferential of F is the set-valued operator $\partial F : \mathcal{H} \to 2^{\mathcal{H}}$ such that for $x \in \mathcal{H}$

$$\partial F(x) \stackrel{\text{def}}{=} \left\{\eta \in \mathcal{H} : \ F(x') - F(x) \geq \left\langle \eta, x' - x \right\rangle, \quad \forall x' \in \mathcal{H} \right\}. \tag{2.2.2}$$

F is *subdifferentiable* at x if $\partial F(x) \neq \emptyset$, and an element of $\partial F(x)$ is called a *subgradient*.

We have the following result whose proof can be found in [16, Proposition 17.26(i)].

Lemma 2.2.7. *Let $F : \mathcal{H} \to \bar{R}$ be proper and convex, and let $x \in \mathrm{dom}(F)$. Suppose that F is Gâteaux differentiable at x. Then $\partial F(x) = \{\nabla F(x)\}$.*

The proof of this result can be found in [16, Proposition 17.26(i)].

In plain words, a Gâteaux differentiable function at x is subdifferentiable there with its gradient as its unique subgradient.

Definition 2.2.8 (Normal cone). Let $\mathcal{S} \subseteq \mathcal{H}$ be a non-empty closed convex set. The normal cone operator is the subdifferential of the indicator function of $\mathcal{S}$, i.e.,

$$N_{\mathcal{S}}(x) \stackrel{\text{def}}{=} \partial i_{\mathcal{S}}(x) = \begin{cases} \left\{\eta \in \mathcal{H} : \ \left\langle \eta, x' - x \right\rangle \leq 0, \forall x' \in \mathcal{H}\right\} & \text{if } x \in \mathcal{S}, \\ \emptyset & \text{otherwise.} \end{cases} \tag{2.2.3}$$

2.2.2 Accretive operators and non-linear semi-groups

All the definitions and results with proofs can be found for instance in [7].

Let $(\mathcal{X}, \|\cdot\|)$ be a Banach space. Let $A : \mathcal{X} \to 2^{\mathcal{X}}$ be a set-valued operator. For notational convenience, the operator will be sometimes identified with its graph by denoting $(x, y) \in A$ for $y \in A(x)$. $\mathbf{Dom}(A) \stackrel{\text{def}}{=} \{x \in X \ : \ Ax \neq \emptyset\}$ is called the *domain* of A and $\mathbf{R}(A) \stackrel{\text{def}}{=} \{Ax \ : \ x \in \mathbf{Dom}(A)\}$ its range.

Definition 2.2.9 (Accretive operator). An operator A in $\mathcal{X}$ is *accretive* if

$$\|x - \hat{x}\| \leq \|x - \hat{x} + \lambda(y - \hat{y})\| \quad \text{whenever} \quad \lambda > 0 \quad \text{and} \quad (x, y), (\hat{x}, \hat{y}) \in A.$$

Definition 2.2.10 (Non-expansive operator). An operator $A : \mathcal{X} \to \mathcal{X}$ is called *non-expansive* if it is 1-Lipschitz continuous, *i.e.*

$$\|A(x) - A(\hat{x})\| \leq \|x - \hat{x}\|, \ \forall x, \hat{x} \in \mathcal{X}.$$

Definition 2.2.11 (Resolvent). Let $A : \mathcal{X} \to 2^{\mathcal{X}}$ and $\gamma > 0$. The *resolvent* of A is defined by

$$J_{\gamma A} \stackrel{\text{def}}{=} (\mathbf{I} + \gamma A)^{-1}.$$

We have the following equivalent characterization of accretivity, whose proof can be found in e.g., [96].

Lemma 2.2.12. *The operator A is accretive if and only if its resolvent is a single-valued non-expansive map on $\mathbf{Dom}(J_{\lambda A})$ for $\lambda > 0$.*

Definition 2.2.13 (m-accretive operator). An operator $A : \mathcal{X} \to 2^{\mathcal{X}}$ is m-accretive if it is accretive and $\mathbf{Dom}(J_{\lambda A}) = \mathcal{X}$ for some (hence all) $\lambda > 0$.

In the Hilbertian case, the notion of m-accretivity coincides with maximal monotonicity which is the celebrated Minty theorem.

Crandall and Liggett introduced in [40] the following limit:

$$S(t)x_0 = \lim_{n \to \infty} (J_{t/nA})^n.$$

Under some closedness assumptions on the operator A, they proved that this limit exists and defines a strongly continuous semigroup $\{S(t)\}_{t \geq 0}$ on $\mathcal{X}$. This semigroup plays an important role for proving solution existence and uniqueness of the abstract Cauchy problem

$$\begin{cases} \dot{x} + Ax \ni 0, \\ x(t_0) = x_0. \end{cases} \tag{2.2.4}$$

More precisely, $x(t) \stackrel{\text{def}}{=} S(t - t_0)x_0$ is the unique strong solution to the abstract Cauchy problem (2.2.4). In the context of the non-local p-Laplacian evolution equation that will be at the heart of Part I, this exponential formula will be instrumental to prove not only for well-posedness, but also to establish Lipschitz continuity of the solution as a function of the initial data. A key step to prove this is to show that the nonlocal p-Laplacian operator belongs to a rich family of operators known as m-completely accretive operators.

In [18], Ph. Bénilan and M. G Crandall introduced a class of operators named completely accretive, for which the semigroup $S(t)$ (see [18]) is order-preserving and non-expansive in every L^p, $p \in [1, +\infty]$. Here we outline some of the main ideas given in [18].

Let Θ be an open set of $\mathbb{R}^N$ and let $\mathcal{M}(\Theta)$ be the space of measurable functions from Θ into $\mathbb{R}$. For $u, v \in \mathcal{M}(\Theta)$, we write

$$u \ll u \quad \text{if and only if} \quad \int_{\Theta} j(u)dx \leq \int_{\Theta} j(u)dx$$

for all $j \in \mathcal{J}_0 \stackrel{\text{def}}{=} \{j : \mathbb{R} \to [0, +\infty], j \text{ convex, lsc, } j(0) = 0\}$.

Definition 2.2.14 (Completely accretive operator). Let A be an operator in $\mathcal{M}(\Theta)$. We say that A is *completely accretive* if

$$u - \hat{u} \ll u - \hat{u} + \lambda(v - \hat{v}) \quad \text{for all } \lambda > 0 \text{ and all } (u, \hat{u}), (v, \hat{v}) \in A.$$

The definition of completely accretive operators does not refer explicitly to topologies or norms. However, if A is completely accretive in $\mathcal{M}(\Theta)$ and $A \subset L^p(\Theta) \times L^p(\Theta)$, $p \in [1, \infty]$ then A is accretive in $L^p(\Theta)$.

Definition 2.2.15 (m-completely accretive operator). An operator A on $\mathcal{X}$ is completely accretive if it is completely accretive and $\operatorname{dom}(J_A) = \mathcal{X}$, A is said m-completely accretive.

2.2.3 Mean value theorem for continuous functions

In this section we state a lemma that is a generalization of the Lagrange mean value theorem retaining only the continuity assumption, but weakening the differentiability hypothesis. But before this, we state the following classical lemma which is useful throughout the manuscript.

Lemma 2.2.16. *For $\alpha \in]0, 1]$ and $a, b \geq 0$, we have*

$$(a + b)^\alpha \leq a^\alpha + b^\alpha.$$

Lemma 2.2.17. *Suppose that the real-valued function f is continuous on $[a, b]$, where $a < b$, both a and b being finite. If the right and left-derivatives f'_+ and f'_- exist as extended-valued functions on $]a, b[$, then there exists $c \in]a, b[$ such that either*

$$f'_+(c) \leq \frac{f(b) - f(a)}{b - a} \leq f'_-(c)$$

or

$$f'_-(c) \leq \frac{f(b) - f(a)}{b - a} \leq f'_+(c).$$

If moreover f'_+ and f'_- coincide on $]a, b[$, then f is differentiable at c and

$$f(b) - f(a) = f'(c)(b - a).$$

PROOF : From [42, p. 115] (see also [114]), we have under the sole continuity assumption of f on $[a, b]$ that either

$$\frac{f(c + h) - f(c)}{h} \leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(c - d)}{d}$$

or

$$\frac{f(c) - f(c - d)}{d} \leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(c + h) - f(c)}{h},$$

for all $h > 0$ and $d > 0$ such that $(c + h, c - d) \in]a, b[^2$. Passing to the limit as $h \to 0^+$ and $d \to 0^+$ (the limits exist in $[-\infty, +\infty]$ by assumption), we get our inequalities. When f'_+ and f'_- coincide on $]a, b[$, and in particular at c, the inequalities become an equality $f'_+(c) = f'_-(c) = \frac{f(b)-f(a)}{b-a}$, and the derivative at c is finite, whence differentiability follows. □

Let us apply this result to $f : t \in \mathbb{R} \mapsto |t|^{p-2}t$, $p > 1$. f is a continuous[1] monotonically increasing and odd function on $\mathbb{R}$. It is moreover everywhere differentiable for $p \geq 2$, and for $p \in]1, 2[$ it is differentiable except at 0, where $f'_+(0) = f'_-(0) = +\infty$. For all $c \neq 0$, we have $f'(c) = (p - 1)|c|^{p-2}$. Thus applying Lemma 2.2.17, we get the following corollary.

[1]Observe that f is not even continuous at 0 when $p = 1$, and thus Lemma 2.2.17 cannot be applied when $0 \in [a, b]$.

Corollary 2.2.18. *Let $a < b$, both a and b being finite. Then, for any $p > 1$, there exists $c \in]a, b[\setminus\{0\}$ such that*

$$|b|^{p-2}b - |a|^{p-2}a = (p-1)|c|^{p-2}(b-a).$$

2.2.4 Embeddings of L^p spaces on bonded domains

Since Ω has finite Lebesgue measure, we have the classical inclusion $L^q(\Omega) \subset L^p(\Omega)$ for $1 \leq p \leq q < +\infty$. More precisely assume without loss of generality that $|\Omega| = 1$, Then

$$\|f\|_{L^p(\Omega)} \leq |\Omega|^{1/p-1/q}\|f\|_{L^q(\Omega)} = \|f\|_{L^q(\Omega)} \leq \|f\|_{L^\infty(\Omega)}, \tag{2.2.5}$$

We also have the following useful (reverse) bound.

Lemma 2.2.19. *For any $1 \leq q < p < +\infty$ we have*

$$\|f\|_{L^p(\Omega)} \leq \|f\|_{L^\infty(\Omega)}^{1-q/p}\|f\|_{L^q(\Omega)}^{q/p}.$$

In particular, for $q = 1$

$$\|f\|_{L^p(\Omega)} \leq \|f\|_{L^\infty(\Omega)}^{1-1/p}\|f\|_{L^1(\Omega)}^{1/p}.$$

Proof : Using Hölder inequality, we have

$$\begin{aligned}\|f\|_{L^p(\Omega)} &= \left(\int_\Omega |f|^q|f|^{p-q}\right)^{1/p} \\ &\leq \left(\left(\int_\Omega |f|^q\right)\|f\|_{L^\infty(\Omega)}^{p-q}\right)^{1/p} \\ &= \|f\|_{L^\infty(\Omega)}^{1-q/p}\|f\|_{L^q(\Omega)}^{q/p}.\end{aligned}$$

□

2.3 Lipschitz spaces on bounded domains

In this section, we introduce the Lipschitz spaces $\mathrm{Lip}(s, L^p(\Omega^d))$, for $d \in \{1, 2\}$, which contain functions with, roughly speaking, s "derivatives" in $L^p(\Omega^d)$ [41, Ch. 2, Section 9]. These spaces will be a key tool for us to study networks on convergent graph sequences as we will be able to get non-asymptotic error estimates for different graph models when adding the assumption of belonging to these spaces to the kernel $K(\cdot,\cdot)$ and the initial condition $g(\cdot)$ in ($\mathcal{P}_{\text{nloc}}$) and ($\mathcal{VP}_{\text{nloc}}$).

Definition 2.3.1. For $F \in L^p(\Omega^d)$, $p \in [1, +\infty]$, we define the (first-order) $L^p(\Omega^d)$ modulus of smoothness by

$$\omega(F, h)_p \stackrel{\text{def}}{=} \sup_{\boldsymbol{z}\in\mathbb{R}^d, |\boldsymbol{z}|<h} \left(\int_{\boldsymbol{x}, \boldsymbol{x}+\boldsymbol{z}\in\Omega^d} |F(\boldsymbol{x}+\boldsymbol{z}) - F(\boldsymbol{x})|^p d\boldsymbol{x}\right)^{1/p}. \tag{2.3.1}$$

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

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

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

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

The space $\mathrm{Lip}(s, L^p(\Omega^2))$ is the Besov space $\mathbf{B}^s_{p,\infty}$ [41, 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:

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

where $e_i, i \in \{1, d\}$ are the coordinate vectors in $\mathbb{R}^d$; see [41, 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 $F \in L^p(\Omega^2)$ (a similar reasoning holds on Ω),

$$\hat{F}_n(x, y) \stackrel{\text{def}}{=} \frac{1}{|\Omega^{(n)}_{ij}|} \sum_{ij} \left(\int_{\Omega^2} F(x', y') \chi_{\Omega^{(n)}_{ij}}(x', y') dx' dy' \right) \chi_{\Omega^{(n)}_{ij}}(x, y),$$

where $\chi_{\Omega^{(n)}_{ij}}$ is the characteristic function of $\Omega^{(n)}_{ij}$. Clearly, $\hat{F}_n$ is nothing but the projection $\mathbf{P}_{V_{n^2}}(F)$ of F on the n^2-dimensional subspace V_{n^2} of $L^p(\Omega^2)$ defined as $V_{n^2} = \mathrm{Span}\left\{\chi_{\Omega^{(n)}_{ij}} : (i, j) \in [n]^2\right\}$.

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 $\Omega_{nij} \stackrel{\text{def}}{=} \{]x_{i-1}, x_i] \times]y_{j-1}, y_j] : (i, j) \in [n]^2\}$ of maximal mesh size $\delta \stackrel{\text{def}}{=} \max_{(i,j)\in[n]^2} \max(|x_i - x_{i-1}|, |y_j - y_{j-1}|)$,

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

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

$$\|F - F_n\|_{L^q(\Omega^d)} \leq C_s \delta^s |F|_{\mathrm{Lip}(s, L^q(\Omega^d))}. \tag{2.3.2}$$

Proof : Using the general bound [41, Ch. 7, Theorem 7.3] for the error in spline approximation, and in view of Definition 2.3.1, we have

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

□

An immediate consequence is the following result.

Lemma 2.3.3. *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*

$$\|F - F_n\|_{L^p(\Omega^d)} \leq C(p, q, s) \delta^{s \min\{1, q/p\}}. \tag{2.3.3}$$

Proof : We have

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

where we used (2.2.5) (resp. Lemma 2.2.19) and Lemma 2.3.2 in the first (resp. second) case. □

2.4 Tools from probability theory

We here provide two well-known deviation inequalities that will be play a key role in establishing non-asymptotic (sharp) deviation bounds when studying our models on networks on random inhomogeneous graphs in Sections 4.4 and 7.3.

Rosenthal's inequality [69]. Let n be a positive integer, $\gamma \geq 2$ and $U_1, \ldots, U_n$ be n zero mean independent random variables such that $\sup_{i \in \{1, \cdots, n\}} \mathbb{E}(\left|U_i\right|^\gamma) < \infty$. Then there exists a constant $C > 0$ 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 [102, Theorem 6]. Let n be a positive integer and $U_1, \ldots, U_n$ be n zero mean independent random variables such that there exists a constant $M > 0$ satisfying $\sup_{i \in [n]} |U_i| \leq M < \infty$. Then, for any $\upsilon > 0$,

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

Part I

The Nonlocal p-Laplacian Evolution Problem

Chapter 3

General Error Bound

Main contributions of this chapter

- Kobayashi type estimates: Error estimates to compare two trajectories corresponding to the p-Laplacian governed by two kernels and initial data (Theorem 3.3.1).
- Consistency and error estimates of the numerical solutions to the fully-discretized problem valid uniformly for $t \in [0, T]$, where $T > 0$ (Theorem 3.4.4)

The content of this chapter appeared in [66].

Contents

In this chapter, we present a consistency analysis for the nonlocal p-Laplacian evolution problem. Our results include three main parts: well-posedness, consistency of the time-continuous problem and that of the time-discrete problem. For the time-discrete problem, both forward and backward Euler schemes for time discretization are addressed. We prove the convergence of these schemes before we compare the corresponding problems to the continuous one. The obtained error bound will be used in the next chapter to analyze networks on convergent graph sequences. Finally, the usefulness of our results is illustrated by applying them to a coupled nonlocal evolution system with a source term to establish its consistency.

3.1 Introduction

3.1.1 Problem statement

Let us recall now the nonlinear diffusion problem ($\mathcal{P}_{\text{nloc}}$) introduced in Section 1.1.1:

$$\begin{cases} u_t(x,t) = \frac{\partial}{\partial t}u(x,t) = -\boldsymbol{\Delta}_p^K(u(x,t)), & a.e.\ x \in \Omega, t > 0, \\ u(x,0) = g(x), & a.e.\ x \in \Omega, \end{cases} \tag{$\mathcal{P}_{\text{nloc}}$}$$

where p is a fixed but arbitrary number in $]1,+\infty[$ and $\boldsymbol{\Delta}_p^K$ is the nonlocal Laplacian operator:

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

$\Omega \subset \mathbb{R}$ is a bounded domain, without loss of generality $\Omega = [0,1]$, and is a symmetric, non-negative and bounded function. As we precise in Section 1.1.1, for ($\mathcal{P}_{\text{nloc}}$) we are dealing with Neumann boundary conditions. Indeed, since we are integrating in Ω we are imposing that diffusion takes place only in Ω. There is no flux across the boundary. Hence, we are dealing here with the nonlocal analogue to Neumann boundary conditions.

When dealing with local evolution equations, two models of nonlinear diffusion have been extensively studied in the literature, the porous medium equation $v_t = \Delta(\left|v\right|^{m-1}v)$ and the p-Laplacian evolution $v_t = \operatorname{div}(\left|\nabla v\right|^{p-2}\nabla v)$. For the first case, the nonlocal analogous equation was studied in [7, Chapter 5]. The nonlocal analog of the p-Laplacian equation was studied as well in [7] for the particular case $K(x,y) = J(x-y)$. Together with the study of existence and uniqueness of the solution, an important result is proved in [7], that is, if the kernel J is rescaled in an appropriate way, the corresponding solutions of the nonlocal p-Laplacian evolution problems converge strongly in $L^\infty((0,T);L^p(\Omega))$ to the solution of the local p-Laplacian evolution problem. Our main goal in this chapter is to study first the existence and uniqueness of the solution for problem $(\mathcal{P}_{\text{nloc}})$ governed by the bi-variate symmetric kernel and then study its consistency.

3.1.2 Relation to prior work

The authors of [91] have already studied numerical approximations of $(\mathcal{P}_{\text{nloc}})$ under different but complementary assumptions. Indeed, in that paper, only the case $K(x,y) = J(x-y)$ was considered. The authors showed that solutions to the numerical scheme converge to the continuous solution for both semi-discrete and totally discrete approximations. However, the convergence is only uniform and requires the positivity of the solution.

3.2 Well-posedness

We begin by studying the well posedness of $(\mathcal{P}_{\text{nloc}})$. To do so, we treat problem $(\mathcal{P}_{\text{nloc}})$ from the point of view of nonlinear semigroup theory (see Section 2.2.2). For that, we start by giving some preliminary properties of the nonlocal p-Laplacian operator $\mathbf{\Delta}_p^K$.

Proposition 3.2.1.

(i) $\mathbf{\Delta}_p^K$ *is positively homogenous of degree* $p-1$ *;*

$$\mathbf{\Delta}_p^K(\alpha u(x,t)) = \alpha^{p-1}\mathbf{\Delta}_p^K(u(x,t)), \qquad \alpha > 0.$$

(ii) $L^{p-1}(\Omega) \subset \mathbf{Dom}(\mathbf{\Delta}_p^K)$ *if* $p > 2$*;*

(iii) *For* $1 < p \leq 2$ *,* $\mathbf{Dom}(\mathbf{\Delta}_p^K) = L^1(\Omega)$ *and* $\mathbf{\Delta}_p^K$ *is closed in* $L^1(\Omega) \times L^1(\Omega)$*;*

(iv) *For* $p \in]1,+\infty[$*,* $\mathbf{\Delta}_p^K$ *is completely accretive and satisfies the range condition*

$$L^p(\Omega) \subset \mathbf{R}(\mathbf{I} + \mathbf{\Delta}_p^K). \tag{3.2.1}$$

Consequently, the resolvent $J_{\lambda\mathbf{\Delta}_p^K}$ *is single-valued and nonexpansive in* $L^q(\Omega)$ *for* $q \in [1,+\infty]$*.*

Proof : Statements (i) and (ii) are immediate.

(iii) In fact $\mathbf{\Delta}_p^K$ is closed in $L^1(\Omega) \times L^1(\Omega)$ if its graph is closed in $L^1(\Omega) \times L^1(\Omega)$. That is, if $u_n \in \mathbf{Dom}(\mathbf{\Delta}_p^K)$ such that $u_n \overset{L^1(\Omega)}{\longrightarrow} u$ and $\mathbf{\Delta}_p^K u_n \overset{L^1(\Omega)}{\longrightarrow} f$, then $u \in \mathbf{Dom}(\mathbf{\Delta}_p^K)$ and $f = \mathbf{\Delta}_p^K u$, which arises automatically from the continuity of the operator $\mathbf{\Delta}_p^K$.

(iv) See [7, THeorem 6.7]. □

Remark 3.2.2. Arguments are more intricate for $p = 1$ (we still have complete accretivity but the range condition becomes only $L^\infty(\Omega) \subset \mathbf{Dom}(J_{\lambda\mathbf{\Delta}_1^K})$). The problem is still open for $p = +\infty$.

Solutions of $(\mathcal{P}_{\text{nloc}})$ will be understood in the following sense:

Definition 3.2.3. A solution of $(\mathcal{P}_{\text{nloc}})$ in $[0,T]$ is a function

$$u \in W^{1,1}(0,T;L^1(\Omega)),$$

that satisfies $u(x,0) = g(x)$ a.e. $x \in \Omega$ and

$$u_t(x,t) = -\mathbf{\Delta}_p^K(u(x,t)) \quad \text{a.e. in } \Omega\times]0,T[.$$

Remark 3.2.4. Observe that since $u \in W^{1,1}(0,T;L^1(\Omega))$, we have that u is also a **strong** solution (see [7, Definition A.3]). Indeed,

$$\left.\begin{array}{r} C(0,T;L^1(\Omega)) \subset W^{1,1}(0,T;L^1(\Omega)) \\ \\ W^{1,1}(0,T;L^1(\Omega)) \subset W^{1,1}_{loc}(0,T;L^1(\Omega)) \end{array}\right\} \Rightarrow u \in C(0,T;L^1(\Omega)) \cap W^{1,1}_{loc}(0,T;L^1(\Omega)).$$

We are now in position to study well-posedness of problem ($\mathcal{P}_{\mathrm{nloc}}$).

Theorem 3.2.5. *Suppose $p \in]1,+\infty[$ and let $g \in L^p(\Omega)$.*

(i) For any $T > 0$, there exists a unique strong solution in $[0,T]$ of ($\mathcal{P}_{\mathrm{nloc}}$).

(ii) Moreover, for $q \in [1,+\infty]$, if $g_i \in L^q(\Omega)$, $i = 1,2$, and u_i is the solution of ($\mathcal{P}_{\mathrm{nloc}}$) with initial condition g_i, then

$$\left\|u_1(t) - u_2(t)\right\|_{L^q(\Omega)} \leq \left\|g_1 - g_2\right\|_{L^q(\Omega)}, \quad \forall t \in [0,T]. \tag{3.2.2}$$

Remark 3.2.6. For $p \in [1,+\infty]$, taking the initial data in $L^p(\Omega)$, one can show existence and uniqueness of a mild but not a strong solution as $L^1(\Omega)$ and $L^\infty(\Omega)$ are not reflexive spaces and thus do not have the Radon-Nikodym property (see [7, Theorem A.29 and Proposition A.35]). For $p = 1$, one can still establish uniqueness by studying the limit of ($\mathcal{P}_{\mathrm{nloc}}$) as $p \to \infty$ (see [7, Chapter 7, Theorem 7.2], however, (3.2.2) is not verified anymore.

The proof of Theorem 3.2.5 is an extension of that of [7, Theorem 6.8] to the case of a symmetric, nonnegative and bounded kernel K as in our setting (see [7, Remark 6.9]). For this, we only need to show the corresponding versions of [7, Lemmas 6.5 and 6.6] (which are stated there without a proof).

The first lemma to prove is that corresponding to [7, Lemmas 6.5]. It consists in an integration for the nonlocal p-Laplacian operator, which plays the same role as the integration by parts for the local p-Laplacian.

Lemma 3.2.7. *For every $u, v \in L^p(\Omega)$,*

$$\begin{aligned} &-\int_\Omega \int_\Omega K(x,y)\left|u(y) - u(x)\right|^{p-2}(u(y) - u(x))dyv(x)dx \\ &= \frac{1}{2}\int_\Omega \int_\Omega K(x,y)\left|u(y) - u(x)\right|^{p-2}(u(y) - u(x))(v(y) - v(x))dydx. \end{aligned}$$

Proof : Let Ω' be a bounded subset of $\mathbb{R}$ and let $\Gamma \subset \mathbb{R} \setminus \mathrm{int}(\Omega')$.

For $\alpha : (\Omega' \cup \Gamma) \times (\Omega' \cup \Gamma) \to \mathbb{R}$, $u : \Omega' \cup \Gamma \to \mathbb{R}$, and $f : (\Omega' \cup \Gamma) \times (\Omega' \cup \Gamma) \to \mathbb{R}$. We define as in [64] the following generalized nonlocal operators

(a) **Generalized gradient**

$$\mathcal{G}(u)(x,y) \stackrel{\text{def}}{=} (u(y) - u(x))\alpha(x,y), \quad x, y \in \Omega' \cup \Gamma,$$

(b) **Generalized nonlocal divergence**

$$\mathcal{D}(f)(x,y) \stackrel{\text{def}}{=} \int_{\Omega' \cup \Gamma} (f(x,y)\alpha(x,y) - f(y,x)\alpha(y,x))dy, \quad x \in \Omega',$$

(c) **Generalized normal component**

$$\mathcal{N}(f)(x,y) \stackrel{\text{def}}{=} -\int_{\Omega' \cup \Gamma} (f(x,y)\alpha(x,y) - f(y,x)\alpha(y,x))dy, \quad x \in \Gamma.$$

With the above notation in place, the authors in [64] prove that for $v : \Omega' \cup \Gamma \to \mathbb{R}$ and $s : \Omega' \cup \Gamma \times \Omega' \cup \Gamma \to \mathbb{R}$, the following identity holds

$$\int_{\Omega'} v\mathcal{D}(s)dx + \int_{\Omega'\cup\Gamma}\int_{\Omega'\cup\Gamma} s\mathcal{G}(v)dydx = \int_{\Gamma} v\mathcal{N}(s)dx. \tag{3.2.3}$$

Let $\mu : (\Omega' \cup \Gamma) \times (\Omega' \cup \Gamma) \to \mathbb{R}$ be given by

$$\mu(x,y) \overset{\text{def}}{=} |\alpha(x,y)|^p.$$

In our particular case μ is the kernel $K(\cdot,\cdot)$, so that we suppose that α is symmetric. Hence, the following identity

$$\mathcal{D}(|\mathcal{G}(u)|^{p-2}\mathcal{G}(u)) = \mathcal{L}_p u \overset{\text{def}}{=} 2\int_{\Omega'\cup\Gamma} |u(y)-u(x)|^{p-2}(u(y)-u(x))\mu(x,y)dy$$

was also shown in [64, (5.3)] for $p = 2$. The general case was proved in [68], that is

$$\mathcal{L}_p u = \mathcal{D}(\left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u)). \tag{3.2.4}$$

The equality holds whenever both sides are finite.

Applying (3.2.3) with $s(x,y) = \left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u)(x,y)$ and using the identity (3.2.4), we obtain

$$\int_{\Omega'} \mathcal{L}_p(u)vdx + \int_{\Omega'\cup\Gamma}\int_{\Omega'\cup\Gamma} (\left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u)).\mathcal{G}(v)dxdy = \int_{\Gamma} \mathcal{N}(\left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u))vdx.$$

Hence

$$\begin{aligned}
\int_{\Omega'} \mathcal{L}_p vdx &= -\int_{\Omega'\cup\Gamma}\int_{\Omega'\cup\Gamma} (\left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u))\mathcal{G}(v)dxdy + \int_{\Gamma} v\mathcal{N}(\left|\mathcal{G}(u)\right|^{p-2} \\
&= -\int_{\Omega'\cup\Gamma}\int_{\Omega'\cup\Gamma} (\left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u))\mathcal{G}(v)dxdy \\
&+ \int_{\Gamma}\left(-\int_{\Omega'\cup\Gamma} \left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u)(x,y)\alpha(x,y) - \left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u)(y,x)\alpha(y,x)dy\right)vdx \\
&= -\int_{\Omega'\cup\Gamma}\int_{\Omega'\cup\Gamma} \left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u))\mathcal{G}(v)dxdy \\
&- \int_{\Gamma}\int_{\Omega'\cup\Gamma} \alpha(x,y)\left(\left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u)(x,y) - \left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u)(y,x)\right)dyvdx \\
&= -\int_{\Omega'\cup\Gamma}\int_{\Omega'\cup\Gamma} \left|\mathcal{G}(u)\right|^{p-2}\mathcal{G}(u))\mathcal{G}(v)dxdy - \int_{\Gamma} \mathcal{L}_p(u)vdx.
\end{aligned}$$

Thus

$$\int_{\Omega'\cup\Gamma} \mathcal{L}_p(u)vdx = -\int_{\Omega'\cup\Gamma}\int_{\Omega'\cup\Gamma} |\mathcal{G}(u)|^{p-2}\mathcal{G}(u))\mathcal{G}(v)dxdy. \tag{3.2.5}$$

Replacing $\mathcal{G}$ with its form in (3.2.5) and taking $\Omega = \Omega' \cup \Gamma$ as this nonlocal integration formula does not contain any boundary terms, so that, the values of u could be nonzero on the domain Γ without affecting the formula, we get the desired result. □

From this lemma the following monotonicity result can be deduced.

Lemma 3.2.8. *Let $T : \mathbb{R} \to \mathbb{R}$ be a nondecreasing function. Then*

(i) For every $u, v \in L^p(\Omega)$ such that $T(u-v) \in L^p(\Omega)$, we have

$$\begin{aligned}
&\int_{\Omega} (\mathbf{\Delta}_p^K u(x) - \mathbf{\Delta}_p^K v(x))T(u(x)-v(x))dx \\
&= \frac{1}{2}\int_{\Omega}\int_{\Omega} K(x,y)(T(u(y)-v(y)) - T(u(x)-v(x))) \\
&\times \left(\left|u(y)-u(x)\right|^{p-2}(u(y)-u(x)) - \left|v(y)-v(x)\right|^{p-2}(v(y)-v(x))\right)dydx.
\end{aligned} \tag{3.2.6}$$

(ii) Moreover, if T is bounded (3.2.6) holds for every $u, v \in \mathbf{Dom}(\mathbf{\Delta}_p^K)$.

PROOF :

(i) We have

$$\begin{aligned}
&\int_\Omega (\mathbf{\Delta}_p^K u(x) - \mathbf{\Delta}_p^K v(x)) T(u(x)-v(x))dx \\
&= \int_\Omega \left(- \int_\Omega K(x,y)\big|u(y)-u(x)\big|^{p-2}(u(y)-u(x))dy \right) T(u(x)-v(x))dx \\
&+ \int_\Omega \left(\int_\Omega K(x,y)\big|v(y)-v(x)\big|^{p-2}(v(y)-v(x))dy \right) T(u(x)-v(x))dx \\
&= -\int_\Omega\int_\Omega K(x,y)(\big|u(y)-u(x)\big|^{p-2}(u(y)-u(x)) - \\
&\big|v(y)-v(x)\big|^{p-2}(v(y)-v(x)))dyT(u(x)-v(x))dx \\
&= -\int_\Omega\int_\Omega K(x,y)\big|u(y)-u(x)\big|^{p-2}(u(y)-u(x))dyT(u(x)-v(x))dx - \\
&- \int_\Omega\int_\Omega K(x,y)\big|v(y)-v(x)\big|^{p-2}(v(y)-v(x))dyT(u(x)-v(x))dx \\
&= \frac{1}{2}\int_\Omega\int_\Omega K(x,y)\big|u(y)-u(x)\big|^{p-2}(u(y)-u(x))(T(u(y)-v(y)) - T(u(x)-v(x))dxdy \\
&- \frac{1}{2}\int_\Omega\int_\Omega K(x,y)\big|v(y)-v(x)\big|^{p-2}(v(y)-v(x))(T(u(y)-v(y)) - T(u(x)-v(x))dxdy \\
&= -\frac{1}{2}\int_\Omega\int_\Omega K(x,y)(\big|u(y)-u(x)\big|^{p-2} - |v(y)-v(x)|^{p-2}(v(y)-v(x))) \\
&\times (T(u(y)-v(y)) - T(u(x)-v(x))dxdy.
\end{aligned}$$

(ii) If T is bounded , we have

$$\forall u, v \in \mathbf{Dom}(\mathbf{\Delta}_p^K), \qquad T(u-v) \in L^p(\Omega).$$

□

In order to make this manuscript as self-contained as possible, we now give a sketch of the proof of Theorem 3.2.5.

PROOF OF THEOREM 3.2.5: The first step is the fact that the operator $\mathbf{\Delta}_p^K$ verifies Proposition 3.2.1 (precisely the fourth statement)

In short this means that for any $\phi \in L^p(\Omega)$ there is a unique solution to the problem $u + \mathbf{\Delta}_p^K(u) = \phi$ and the resolvent $J_{\mathbf{\Delta}_p^K}$ is a non-expansive mapping in $L^q(\Omega)$ for all $1 \leq q \leq +\infty$. Combining this with [7, Theorem A.29], we get the existence of a mild solution to ($\mathcal{P}_{\text{nloc}}$) On the other hand this mild solution is a strong solution under the hypothesis of the theorem thanks to the complete accretivity of $\mathbf{\Delta}_p^K$ and the range condition (3.2.1) using ([7, Proposition A.35]). Finally the stability principle (3.2.2) is a consequence of [7, Theorem A.28]. □

3.3 Consistency of the time-continuous problem

We begin our study by giving a general consistency result from which we shall extract particular consistency bounds for every specific model of convergent graph sequences that we will treat in Chapter 4. To do this, let us consider the following Neumann evolution problem as ($\mathcal{P}_{\text{nloc}}$)

$$\begin{cases} \frac{\partial}{\partial t} u_n(x,t) = -\mathbf{\Delta}_p^{K_n}(u_n(x,t)), & (x,t) \in \Omega \times]0,T] \\ u_n(x,0) = g_n(x), & x \in \Omega. \end{cases} \qquad (\mathcal{P}_{\text{nloc}}^n)$$

Though not needed in this chapter, the use of the subscript n is a matter of notation and emphasizes the fact that K_n and g_n depend on the parameter n. This will be clear in the application to graph sequences in Chapter 4.

Now we state and prove our main consistency and convergence theorem.

Theorem 3.3.1. *Suppose $p \in]1,+\infty[$, $g, g_n \in L^\infty(\Omega)$ and K, K_n are measurable, symmetric and bounded mappings. Then ($\mathcal{P}_{\rm nloc}$) and ($\mathcal{P}^n_{\rm nloc}$) have unique solutions, respectively, u and u_n. Moreover the following hold.*

(i) We have the error estimate

$$\|u-u_n\|_{C(0,T;L^p(\Omega))} \le C\left(\|g-g_n\|_{L^p(\Omega)} + \|K-K_n\|_{L^p(\Omega^2)}\right), \tag{3.3.1}$$

where the constant C is independent of n.

(ii) Moreover, if $g_n \to g$ and $K_n \to K$ as $n \to \infty$, almost everywhere on Ω and Ω^2, respectively, then

$$\|u-u_n\|_{C(0,T;L^p(\Omega))} \underset{n\to\infty}{\longrightarrow} 0.$$

Proof : In the proof, C_i is any absolute constant independent of n (but may depend on p). Existence and uniqueness of the solutions u and u_n in the sense of Definition 3.2.3 is a consequence of Theorem 3.2.5.

(i) For $1 < p < +\infty$, we define the function

$$\Psi : x \in \mathbb{R} \mapsto |x|^{p-2}x = \text{sign}(x)|x|^{p-1}.$$

Denote $\xi_n(x,t) = u_n(x,t) - u(x,t)$, by subtracting ($\mathcal{P}_{\rm nloc}$) from ($\mathcal{P}^n_{\rm nloc}$), we have a.e.

$$\begin{aligned}\frac{\partial \xi_n(x,t)}{\partial t} &= \int_\Omega K_n(x,y)\{\Psi(u_n(y,t)-u_n(x,t)) - \Psi(u(y,t)-u(x,t))\}dy \\ &+ \int_\Omega (K_n(x,y)-K(x,y))\Psi(u(y,t)-u(x,t))dy.\end{aligned} \tag{3.3.2}$$

Next, we multiply both sides of (3.3.2) by $\Psi(\xi_n(x,t))$ and integrate over Ω to get

$$\begin{aligned}\frac{1}{p}\int_\Omega \frac{\partial}{\partial t}|\xi_n(x,t)|^p dx &= \int_{\Omega^2} K_n(x,y)\{\Psi(u_n(y,t)-u_n(x,t)) - \Psi(u(y,t)-u(x,t))\}\Psi(\xi_n(x,t))dxdy \\ &+ \int_{\Omega^2} (K_n(x,y)-K(x,y))\Psi(u(y,t)-u(x,t))\Psi(\xi_n(x,t))dxdy.\end{aligned} \tag{3.3.3}$$

We estimate the first term on the right-hand side of (3.3.3) using the fact that K_n is bounded so that there exists a positive constant M independent of n, such that, $\|K_n\|_{L^\infty(\Omega^2)} \le M$,

$$\begin{aligned}&\left|\int_{\Omega^2} K_n(x,y)\{\Psi(u_n(y,t)-u_n(x,t)) - \Psi(u(y,t)-u(x,t))\}\Psi(\xi_n(x,t))dxdy\right| \\ &\le M\int_{\Omega^2} \left|\Psi(u_n(y,t)-u_n(x,t)) - \Psi(u(y,t)-u(x,t))\right||\xi_n(x,t)|^{p-1}dxdy.\end{aligned}$$

Now, applying Corollary 2.2.18 with $a = u_n(y,t) - u_n(x,t)$ and $b = u(y,t) - u(x,t)$ (without loss of generality we assume that $b > a$), we get

$$\begin{aligned}&\int_{\Omega^2} \left|\Psi(u_n(y,t)-u_n(x,t)) - \Psi(u(y,t)-u(x,t))\right||\xi_n(x,t)|^{p-1}dxdy \\ &\le (p-1)\int_{\Omega^2} |\xi_n(y,t)-\xi_n(x,t)||\eta(x,y,t)|^{p-2}|\xi_n(x,t)|^{p-1}dxdy,\end{aligned} \tag{3.3.4}$$

where $\eta(x,y,t)$ is an intermediate value between a and b. As we have supposed that $g \in L^\infty(\Omega)$ and $g_n \in L^\infty(\Omega)$, and as $|\Omega|$ is finite, so that $L^\infty(\Omega) \subset L^p(\Omega)$, we deduce from (3.2.2) in

Theorem 3.2.5 that for any $(x,y)\in\Omega^2$ and $t\in[0,T]$, we have

$$\begin{cases} |\eta(x,y,t)|^{p-2} \le |u(y,t)-u(x,t)|^{p-2} \le \left(2\|u(t)\|_{L^\infty(\Omega)}\right)^{p-2} \le C_1 & \text{for } p\in[2,+\infty[, \\ |\eta(x,y,t)|^{p-2} \le |u_n(y,t)-u_n(x,t)|^{p-2} \le \left(2\|u_n(t)\|_{L^\infty(\Omega)}\right)^{p-2} \le C_1' & \text{for } p\in]1,2[. \end{cases} \tag{3.3.5}$$

Inserting (3.3.5) into (3.3.4), and then using the Hölder and triangle inequalities, it follows that

$$\begin{aligned} &M\int_{\Omega^2} \left|\Psi(u_n(y,t)-u_n(x,t))-\Psi(u(y,t)-u(x,t))\right| |\xi_n(x,t)|^{p-1}dxdy \\ &\le M(p-1)C_1\int_{\Omega^2} |\xi_n(y,t)-\xi_n(x,t)||\xi_n(x,t)|^{p-1}dxdy \\ &= C_2\int_{\Omega^2} |\xi_n(y,t)-\xi_n(x,t)||\xi_n(x,t)|^{p-1}dxdy \\ &\le C_2\left(\int_{\Omega^2}|\xi_n(y,t)-\xi_n(x,t)|^p dxdy\right)^{\frac{1}{p}} \times \left(\int_\Omega |\xi_n(x,t)|^p dx\right)^{\frac{p-1}{p}} \\ &\le 2C_2\|\xi_n(t)\|^p_{L^p(\Omega)}. \end{aligned} \tag{3.3.6}$$

We bound the second term on the right-hand side of (3.3.3) as follows

$$\begin{aligned} &\left|\int_{\Omega^2}(K_n(x,y)-K(x,y))\Psi(u(y,t)-u(x,t))\Psi(\xi_n(x,t))dxdy\right| \\ &= \left|\int_{\Omega^2}(K_n(x,y)-K(x,y))\times \operatorname{sign}(u(y,t)-u(x,t))|u(y,t)-u(x,t)|^{p-1}\Psi(\xi_n(x,t))dxdy\right| \\ &\le 2^{p-1}\|u(t)\|^{p-1}_{L^\infty(\Omega)}\left|\int_{\Omega^2}|K_n(x,y)-K(x,y)||\xi_n(x,t)|^{p-1}dxdy\right| \\ &\le 2^{p-1}\|u(t)\|^{p-1}_{L^\infty(\Omega)}\left(\int_\Omega|\xi_n(x,t)|^p dx\right)^{\frac{p-1}{p}} \times \left(\int_{\Omega^2}|K_n(x,y)-K(x,y)|^p dxdy\right)^{\frac{1}{p}} \\ &\le 2C_3\|\xi_n(t)\|^{p-1}_{L^p(\Omega)}\|K_n-K\|_{L^p(\Omega^2)}. \end{aligned} \tag{3.3.7}$$

Bringing together (3.3.6) and (3.3.7), and using standard arguments to switch the derivation and integration signs (Leibniz rule), we have

$$\frac{d}{dt}\|\xi_n(t)\|^p_{L^p(\Omega)} \le 2pC_2\|\xi_n(t)\|^p_{L^p(\Omega)} + 2pC_3\|K_n-K\|_{L^p(\Omega^2)}\|\xi_n(t)\|^{p-1}_{L^p(\Omega)}. \tag{3.3.8}$$

Let $\varepsilon>0$ be arbitrary but fixed, and set

$$\psi_\varepsilon(t) = \left(\|\xi_n(t)\|^p_{L^p(\Omega)}+\varepsilon\right)^{1/p}.$$

By (3.3.8),

$$\frac{d}{dt}\psi_\varepsilon(t)^p \le 2pC_2\psi_\varepsilon(t)^p + 2pC_3\|K_n-K\|_{L^p(\Omega)}\psi_\varepsilon(t)^{p-1}. \tag{3.3.9}$$

Since $\psi_\varepsilon(t)$ is positive on $[0,T]$, from (3.3.9), we have

$$\frac{d}{dt}\psi_\varepsilon(t) \le 2C_2\psi_\varepsilon(t) + 2C_3\|K_n-K\|_{L^p(\Omega^2)}, \qquad t\in[0,T].$$

We apply Gronwall's inequality for $\psi_\varepsilon(t)$ on $[0,T]$ to get

$$\sup_{t\in[0,T]}\psi_\varepsilon(t) \le \left(\psi_\varepsilon(0)+2C_3T\|K_n-K\|_{L^p(\Omega^2)}\right)\exp\{2C_2T\}. \tag{3.3.10}$$

Since $\varepsilon>0$ is arbitrary, (3.3.10) implies

$$\sup_{t\in[0,T]}\|\xi_n(t)\|_{L^p(\Omega)} \le \left(\|g-g_n\|_{L^p(\Omega)}+2C_3T\|K_n-K\|_{L^p(\Omega^2)}\right)\exp\{2C_2T\}. \tag{3.3.11}$$

The desired result holds.

(ii) Since $g_n, g \in L^\infty(\Omega) \subset L^p(\Omega)$ and $|\Omega|$ is finite, the dominated convergence theorem implies that $\lim_{n\to+\infty} \|g_n\|_{L^p(\Omega)} = \|g\|_{L^p(\Omega)}$. The same reasoning applies to K_n and K. Passing to the limit in (3.3.1) and using the Scheffé-Riesz theorem (see [74, Lemma 2]), we get the claim.

□

Remark 3.3.2. Observe that, since $|\Omega|$ is finite, we have the classical inclusion $L^p(\Omega) \subset L^2(\Omega)$ for $p \geq 2$, which leads to the following bound

$$\|u - u_n\|_{C(0,T;L^2(\Omega))} \leq |\Omega|^{\frac{1}{2}-\frac{1}{p}} \|u - u_n\|_{C(0,T;L^p(\Omega))} = \|u - u_n\|_{C(0,T;L^p(\Omega))},$$

as $|\Omega| = 1$. For $p \in]1,2]$, we have, thanks to Lemma 2.2.19, boundedness of the solutions and Jensen inequality,

$$\|u - u_n\|^2_{C(0,T;L^2(\Omega))} = O\left(\|u - u_n\|^p_{C(0,T;L^p(\Omega))}\right) = O\left(\|g - g_n\|^p_{L^p(\Omega)} + \|K - K_n\|^p_{L^p(\Omega^2)}\right).$$

In summary, there is also convergence with respect to the L^2-norm.

3.4 Consistency of the time-discrete problem

3.4.1 Forward Euler discretization

We now consider the following time-discrete approximation of $(\mathcal{P}_{\text{nloc}})$, the forward Euler discretization applied to $(\mathcal{P}^n_{\text{nloc}})$. For that, let us consider a partition (not necessarily uniform) $\{t_h\}_{h=1}^N$ of the time interval $[0,T]$. Let $\tau_{h-1} \stackrel{\text{def}}{=} |t_h - t_{h-1}|$ and the maximal size $\tau = \max_{h\in[N]} \tau_h$, and denote $u_n^h(x) \stackrel{\text{def}}{=} u_n(x,t_h)$. Then, consider

$$\begin{cases} \dfrac{u_n^h(x) - u_n^{h-1}(x)}{\tau_{h-1}} = -\boldsymbol{\Delta}_p^{K_n}(u_n^{h-1}(x)), & x \in \Omega, h \in [N], \\ u_n^0(x) = g_n^0(x), & x \in \Omega. \end{cases} \qquad (\mathcal{P}^f_{\text{nloc},\tau})$$

Before turning to the consistency result, one may wonder whether $(\mathcal{P}^f_{\text{nloc},\tau})$ is well-posed. In the following result, we show that for $p \in]1,+\infty[$, and starting from $g_n^0 \in L^\infty(\Omega)$, there exists a unique weak accumulation point to the iterates of $(\mathcal{P}^f_{\text{nloc},\tau})$. In turn, in the case of practical interest where the problem is finite-dimensional (in fact Euclidean case) as for the application to graphs, we do have existence and uniqueness. Recall the function R_p from (1.1.1).

Lemma 3.4.1. *Consider problem $(\mathcal{P}^f_{\text{nloc},\tau})$. Assume that $g_n^0 \in L^\infty(\Omega)$. Let $\tau_h = \dfrac{\alpha_h}{\max(\|\boldsymbol{\Delta}_p^{K_n}(u_n^h)\|_{L^2(\Omega)}, 1)}$, and suppose that $\sum_{h=1}^{+\infty} \alpha_h = +\infty$ and $\sum_{h=1}^{+\infty} \alpha_h^2 < +\infty$. Then, the iterates of problem $(\mathcal{P}^f_{\text{nloc},\tau})$, starting from g_n^0, have a unique weak accumulation point $u^\star$. Moreover, there are constants $\beta, \varepsilon > 0$ such that*

$$\min_{0\leq i\leq h} R_p(u_n^i, K_n) - R_p(u^\star, K_n) \leq \max(\beta, 1) \frac{\varepsilon^2 + \sum_{i=0}^h \alpha_i^2}{2\sum_{i=0}^h \alpha_i}.$$

Remark 3.4.2. (a) Our condition on the time-step τ_h is reminiscent of the subgradient method. It can be seen as a non-linear CFL-type condition which depends on the data since $\boldsymbol{\Delta}_p^{K_n}$ is not Lipschitz-continuous but only locally so, hence the dependence of τ_h on $\|\boldsymbol{\Delta}_p^{K_n}(u_n^h)\|_{L^2(\Omega)}$.

(b) The rate of convergence on R_p depends on the choice of $\{\alpha_h\}_h$. If one performs N steps on the interval $[0,T]$, one can take

$$\alpha_h = \frac{\varepsilon}{(N+1)^{1/2+\nu}}, h = 0, \ldots, N, \quad \text{with} \quad \nu \in]0, 1/2[,$$

which entails a convergence rate of $\frac{\max(\beta,1)\varepsilon^2}{(N+1)^{1/2-\nu}}$. The smaller ν the faster the rate.

Before proving Lemma 3.4.1 recall from Definition 2.2.6 the subdifferential of a function $F \in \Gamma_0(L^2(\Omega))$. Let $F : L^2(\Omega) \to \mathbb{R}\cup\{+\infty\}$ be a proper lower-semicontinuous and convex function. The subdifferential of F at $u \in L^2(\Omega)$ is the set-valued operator $\partial F : L^2(\Omega) \to 2^{L^2(\Omega)}$ given by

$$\partial F(u) = \left\{\eta \in L^2(\Omega) :\ F(v) - F(u) \geq \langle \eta, u - v\rangle, \quad \forall v \in L^2(\Omega)\right\},$$

where $\langle .,.\rangle$ denotes the inner product in $L^2(\Omega)$.

Proof : Since $p > 1$, we consider in the Hilbert space $L^2(\Omega)$ the subdifferential $\partial R_p(\cdot, K_n)$ whose graph is in $L^2(\Omega) \times L^2(\Omega)$. It is immediately seen that R_p is convex and Gâteaux-differentiable, and thus $\partial R_p(u) = \left\{\mathbf{\Delta}_p^{K_n}(u)\right\}$ (from Lemma 2.2.7). Moreover, it is maximal monotone (or equivalently m-accretive on $L^2(\Omega)$), see [7, p. 198]. Consequently, using that $g_n^0 \in L^\infty(\Omega) \subset L^2(\Omega)$, and so is u_n^h by induction, a solution to $(\mathcal{P}^f_{\text{nloc},\tau})$ coincides with that of

$$\begin{cases} u_n^h(x) = u_n^{h-1}(x) - \tau_{h-1}\eta^{h-1}, & \eta^{h-1} \in \partial R_p(u_n^h, K_n) \\ u_n^0(x) = g_n^0(x), & x \in \Omega, \end{cases}$$

i.e. the subgradient method with initial point g_n^0. Observe that $(\partial R_p(\cdot, K_n))^{-1}(0) \neq \emptyset$ (0 is in it). Thus with the prescribed choice of τ_h, we deduce from [3, Theorem 1] that the sequence of iterates u_n^h has a unique weak accumulation $u^\star \in (\partial R_p(\cdot, K_n))^{-1}(0)$.

The claim on the rate is classical [1]. We here provide a simple and self-contained proof. Since R_p is continuous and convex on $L^2(\Omega)$, it is locally Lipschitz continuous [16, Theorem 8.29]. Moreover, the sequence $\left\{u_n^h\right\}_h$ is bounded, and hence, $\exists \varepsilon > 0$ such that $\left\|u_n^h - u^\star\right\|_{L^2(\Omega)} \leq \varepsilon$, $\forall h \geq 0$. In turn, R_p is Lipschitz continuous around $u^\star$ with Lipschitz constant, say β. Denote $r_n^h = u_n^h - u^\star$. We have

$$\begin{aligned} \|r_n^h\|^2_{L^2(\Omega)} &= \left\|r_n^{h-1} - \tau_{h-1}\eta^{h-1}\right\|^2_{L^2(\Omega)} \\ &= \left\|r_n^{h-1}\right\|^2_{L^2(\Omega)} - 2\frac{\alpha_{h-1}}{\max\left(\left\|\eta^{h-1}\right\|_{L^2(\Omega)},1\right)}\left\langle \eta^{h-1}, r_n^{h-1}\right\rangle + \alpha_{h-1}^2 \\ &\leq \left\|r_n^{h-1}\right\|^2_{L^2(\Omega)} - 2\frac{\alpha_{h-1}}{\max\left(\left\|\eta^{h-1}\right\|_{L^2(\Omega)},1\right)}\Big(R_p(u_n^{h-1}, K_n)) - R_p(u^\star, K_n)\Big) + \alpha_{h-1}^2, \end{aligned}$$

where we used the subdifferential inequality above to get that

$$R_p(u^\star, K_n) \geq F(u_n^{h-1}, K_n) - \left\langle \eta^{h-1}, r_n^{h-1}\right\rangle.$$

Summing up these inequalities we obtain

$$2\sum_{i=0}^{h} \alpha_i \left(R_p(u_n^i, K_n) - R_p(u^\star, K_n)\right) \leq \max(\beta, 1)\left(\left\|r_n^0\right\|^2_{L^2(\Omega)} + \sum_{i=0}^{h}\alpha_i^2\right),$$

whence we deduce

$$\min_{0\leq i\leq h} R_p(u_n^i) - R_p(u^\star) \leq \max(\beta,1)\frac{\varepsilon^2 + \sum_{i=0}^h \alpha_i^2}{2\sum_{i=0}^h \alpha_i}.$$

□

Since the aim is to compare the solutions of problems $(\mathcal{P}_{\text{nloc}})$ and $(\mathcal{P}^f_{\text{nloc},\tau})$, the solution of $(\mathcal{P}^f_{\text{nloc},\tau})$ being discrete, so that it is convenient to introduce an intermediate model which is the continuous extension of the discrete problem using the discrete function $u_n(x) = (u_n^1(x), \cdots, u_n^N(x))$. Therefore, we consider a time-continuous extension of u_n^h obtained by a time linear interpolation as follows

$$\breve{u}_n(x,t) = \frac{t_h - t}{\tau_{h-1}}u_n^{h-1}(x) + \frac{t - t_{h-1}}{\tau_{h-1}}u_n^h(x), \quad t \in]t_{h-1}, t_h], \quad x \in \Omega, \tag{3.4.1}$$

[1] See e.g. [86, Theorem 3.2.2] in finite dimension with a slightly different normalization of the step size τ_h.

and a time piecewise constant approximation

$$\bar{u}_n(x,t) = \sum_{h=1}^{N} u_n^{h-1}(x)\chi_{]t_{h-1},t_h]}(t). \tag{3.4.2}$$

Then, by construction of $\check{u}_n(x,t)$ and $\bar{u}_n(x,t)$, we have the following evolution problem

$$\begin{cases} \frac{\partial}{\partial t}\check{u}_n(x,t) = -\boldsymbol{\Delta}_p^{K_n}(\bar{u}_n(x,t)), & (x,t) \in \Omega\times]0,T] \\ \check{u}_n(x,0) = g_n^0(x), & x \in \Omega. \end{cases} \tag{3.4.3}$$

Lemma 3.4.3. *Assume that $g_n^0 \in L^\infty(\Omega)$. Let $\check{u}_n$ and $\bar{u}_n$ be the functions defined in (3.4.1) and (3.4.2), respectively, then*

$$\left\|\bar{u}_n(t) - \check{u}_n(t)\right\|_{L^p(\Omega)} = O(\tau), \qquad t \in [0,T]. \tag{3.4.4}$$

Proof : It is easy to see that for $t \in]t_{h-1}, t_h]$,

$$\begin{aligned} \left\|\bar{u}_n(t) - \check{u}_n(t)\right\|_{L^p(\Omega)} \leq (t_h - t)\left\|\frac{u_n^h - u_n^{h-1}}{\tau_{h-1}}\right\|_{L^p(\Omega)} \leq \tau\left\|\frac{u_n^h - u_n^{h-1}}{\tau_{h-1}}\right\|_{L^p(\Omega)} = \tau\left\|\boldsymbol{\Delta}_p^{K_n}(u_n^{h-1})\right\|_{L^p(\Omega)} \\ \leq \tau\left\|\boldsymbol{\Delta}_p^{K_n}(u_n^{h-1})\right\|_{L^\infty(\Omega)} \leq \tau 2^{p-1}\left\|u_n^{h-1}\right\|_{L^\infty(\Omega)}^{p-1}. \end{aligned}$$

By induction, for all $h \geq 1$, we have (see Lemma 3.4.1)

$$\left\|u_n^h\right\|_{L^\infty(\Omega)} \leq \left\|u_n^{h-1}\right\|_{L^\infty(\Omega)} + \alpha 2^{p-1}\left\|u_n^{h-1}\right\|_{L^\infty(\Omega)}^{p-1} < +\infty,$$

where $\alpha = \sup_{h\geq 1} \alpha_h < +\infty$. Since t is arbitrary, we obtain a global estimate for all $t \in [0,T]$. □

We are in position now to give our consistency result for the time-discrete problem.

Theorem 3.4.4. *Suppose $p \in]1,+\infty[$, $g, g_n^0 \in L^\infty(\Omega)$ and K, K_n are measurable, symmetric and bounded mappings.*

Let u be the unique solution of problem ($\mathcal{P}_{\rm nloc}$), and $\check{u}_n$ is built as in (3.4.1) from the time-discrete approximation u_n^{h-1} defined in ($\mathcal{P}^f_{\rm nloc,\tau}$). Then

$$\left\|u - \check{u}_n\right\|_{C(0,T;L^p(\Omega))} \leq C\left(\left\|g_n - g_n^0\right\|_{L^p(\Omega)} + \left\|g - g_n\right\|_{L^p(\Omega)} + \left\|K - K_n\right\|_{L^p(\Omega^2)}\right) + O(\tau), \tag{3.4.5}$$

where the constant C is independent of n.

Proof : We follow the same lines as in the proof of Theorem 3.3.1. Denote $\check{\xi}_n(x,t) = \check{u}_n(x,t) - u_n(x,t)$ and $\bar{\xi}_n(x,t) = \bar{u}_n(x,t) - u_n(x,t)$. We thus have a.e.

$$\frac{\partial \check{\xi}_n}{\partial t} = \int_\Omega K_n(x,y)\{\Psi(\bar{u}_n(y,t) - \bar{u}_n(x,t)) - \Psi(u_n(y,t) - u_n(x,t))\}dy. \tag{3.4.6}$$

Next, we multiply both sides of (3.4.6) by $\Psi(\check{\xi}_n(x,t))$ and integrate over Ω using the relation (3.4.3) to get

$$\frac{1}{p}\int_\Omega \frac{\partial}{\partial t}\left|\check{\xi}_n(x,t)\right|^p dx = \int_{\Omega^2} K_n(x,y)\{\Psi(\bar{u}_n(y,t) - \bar{u}_n(x,t)) - \Psi(u_n(y,t) - u_n(x,t))\}\Psi(\check{\xi}_n)(x,t)dxdy. \tag{3.4.7}$$

Similarly to the proof of Theorem 3.3.1, we bound the term on the right-hand side of (3.4.7) using the fact that K_n is bounded, then applying Corollary 2.2.18 between $\bar{u}_n(y,t) - \bar{u}_n(x,t)$ and $u_n(y,t) -$

$u_n(x,t)$, inequality (3.3.5), and finally using Hölder and triangle inequalities. Altogether, this yields

$$\begin{aligned}
&\Big|\int_{\Omega^2} K_n(x,y)\{\Psi(\bar{u}_n(y,t)-\bar{u}_n(x,t))-\Psi(u_n(y,t)-u_n(x,t))\}\Psi(\check{\xi}_n)(x,t)dxdy\Big| \\
&\le C_2\int_{\Omega^2}\left|\bar{\xi}_n(y,t)-\bar{\xi}_n(x,t)\right|\left|\check{\xi}_n(x,t)\right|^{p-1}dxdy \\
&\le C_2\left(\int_{\Omega^2}\left|\bar{\xi}_n(y,t)-\bar{\xi}_n(x,t)\right|^p dxdy\right)^{\frac{1}{p}}\times\left(\int_\Omega\left|\xi_n(x,t)\right|^p dx\right)^{\frac{p-1}{p}} \\
&\le 2C_2\left\|\bar{\xi}_n(t)\right\|_{L^p(\Omega)}\left\|\check{\xi}_n(t)\right\|^{p-1}_{L^p(\Omega)}.
\end{aligned} \tag{3.4.8}$$

By virtue of Lemma 3.4.3 and the triangle inequality for $\bar{\xi}_n(\cdot,\cdot)$, there exists a positive constant C' such that

$$\begin{aligned}
\left\|\bar{u}_n(t)-u_n(t)\right\|_{L^p(\Omega)} &\le \left\|\bar{u}_n(t)-\check{u}_n(t)\right\|_{L^p(\Omega)}+\left\|\check{u}_n(t)-u_n(t)\right\|_{L^p(\Omega)} \\
&\le C'\tau+\left\|\check{\xi}_n(t)\right\|_{L^p(\Omega)}.
\end{aligned} \tag{3.4.9}$$

Hence, bringing together (3.4.8) and (3.4.9), we obtain

$$\frac{d}{dt}\left\|\check{\xi}_n(t)\right\|^p_{L^p(\Omega)} \le 2pC_2\left\|\check{\xi}_n(t)\right\|^p_{L^p(\Omega)}+2pC'\tau\left\|\check{\xi}_n(t)\right\|^{p-1}_{L^p(\Omega)}. \tag{3.4.10}$$

Arrived at this stage, we proceed in the same way using the Gronwall's lemma as in the proof of Theorem 3.3.1, to get

$$\sup_{t\in[0,T]}\left\|\check{\xi}_n(t)\right\|_{L^p(\Omega)} \le \left(\left\|g_n^0-g_n\right\|_{L^p(\Omega)}+2C'T\tau\right)\exp\{2C_2T\}. \tag{3.4.11}$$

Then,

$$\left\|\check{u}_n-u_n\right\|_{C(0,T;L^p(\Omega))} \le C\left\|g_n^0-g_n\right\|_{L^p(\Omega)}+C''\tau. \tag{3.4.12}$$

Using the triangle inequality and (3.3.1) in Theorem 3.3.1, we get

$$\begin{aligned}
\left\|\check{u}_n-u\right\|_{C(0,T;L^p(\Omega))} &\le \left\|\check{u}_n-u_n\right\|_{C(0,T;L^p(\Omega))}+\left\|u_n-u\right\|_{C(0,T;L^p(\Omega))} \\
&\le C''\tau+C\left(\left\|g_n^0-g_n\right\|_{L^p(\Omega)}+\left\|g-g_n\right\|_{L^p(\Omega)}+\left\|K-K_n\right\|_{L^p(\Omega^2)}\right).
\end{aligned} \tag{3.4.13}$$

□

3.4.2 Backward Euler discretization

Our result in Theorem 3.4.4 also holds when we deal with the backward Euler discretization

$$\begin{cases}
\dfrac{u_n^h(x)-u_n^{h-1}(x)}{\tau_{h-1}} = -\boldsymbol{\Delta}_p^{K_n}(u_n^h(x)), & x\in\Omega, h\in[N], \\
u^0(x)=g_n^0(x), & x\in\Omega,
\end{cases} \tag{$\mathcal{P}^b_{n,\tau}$}$$

which can also be rewritten as the implicit update

$$\begin{cases}
u_n^h(x)=J_{\tau_{h-1}\boldsymbol{\Delta}_p^{K_n}}(u_n^{h-1})(x), & x\in\Omega, h\in[N], \\
u^0(x)=g_n^0(x), & x\in\Omega,
\end{cases}$$

and the resolvent $J_{\tau_{h-1}\boldsymbol{\Delta}_p^{K_n}} \stackrel{\text{def}}{=} \left(\mathbf{I}+\tau_{h-1}\boldsymbol{\Delta}_p^{K_n}\right)^{-1}$ is a single-valued non-expansive operator on $L^p(\Omega)$ since $\boldsymbol{\Delta}_p^{K_n}$ is m-accretive [71]. In addition, problem ($\mathcal{P}^b_{n,\tau}$) is well-posed as we state now.

Lemma 3.4.5. *Let $g_n^0\in L^p(\Omega)$. Suppose that $\underline{\tau}\stackrel{\text{def}}{=}\inf_h \tau_h>0$ or $\sum_{h=1}^{+\infty}\tau_h^{\max(2,p)}=+\infty$, then the iterates of ($\mathcal{P}^b_{n,\tau}$), starting from g_n^0, have a unique weak accumulation point $u^\star\in(\boldsymbol{\Delta}_p^K)^{-1}(0)$. Moreover, if*

$\underline{\tau} > 0$, then for $h \geq 1$

$$\left\|\mathbf{\Delta}_p^{K_n}(u_n^h)\right\|_{L^p(\Omega)} \leq \frac{2\left\|g_n^0 - u^\star\right\|_{L^p(\Omega)}}{(\underline{\tau}C_p)^{1/\max(p,2)}h^{1/\max(p,2)}}.$$

Proof : $\mathbf{\Delta}_p^{K_n}$ is accretive on $L^p(\Omega)$ (see the proof of [7, Theorem 6.7]). Moreover, it is well-known that for $p \in]1,+\infty[$, $L^p(\Omega)$ is a uniformly convex and a uniformly smooth Banach space, whose convexity modulus verifies

$$\delta_{L^p(\Omega)}(\varepsilon) \geq \begin{cases} p^{-1}2^{-p}\varepsilon^p & p \in [2,+\infty[, \\ (p-1)\varepsilon^2/8 & p \in]1,2]. \end{cases}$$

Thus, we are in position to apply [95, Theorem 3] to get uniqueness of the weak accumulation point.

Let us turn to the rate. By m-accretiveness $\mathbf{\Delta}_p^{K_n}$, $J_{\tau_{h-1}\mathbf{\Delta}_p^{K_n}}$ is a single-valued operator on the entire $L^p(\Omega)$, and verifies for any $v, w \in L^p(\Omega)$ and $\lambda \in [0,1]$,

$$\left\|J_{\tau_{h-1}\mathbf{\Delta}_p^{K_n}}(v) - J_{\tau_{h-1}\mathbf{\Delta}_p^{K_n}}(w)\right\|_{L^p(\Omega)} \leq \left\|\lambda(v-w) + (1-\lambda)(J_{\tau_{h-1}\mathbf{\Delta}_p^{K_n}}(v) - J_{\tau_{h-1}\mathbf{\Delta}_p^{K_n}}(w))\right\|_{L^p(\Omega)}. \quad (3.4.14)$$

We now evaluate (3.4.14) at $v = u_n^{h-1}$, $w = u^\star$ and $\lambda = 1/2$, and combine it with [111, Corollary 2]. This leads us to consider two possible cases.

(a) $p \in]2,+\infty[$: since $u_n^h = J_{\tau_{h-1}\mathbf{\Delta}_p^{K_n}}(u_n^{h-1})$ and $u^\star$ is a fixed point of $J_{\tau_{h-1}\mathbf{\Delta}_p^{K_n}}$, and in view of [111, Corollary 2, (3.4)], we have

$$\begin{aligned}\left\|u_n^h - u^\star\right\|_{L^p(\Omega)}^p &\leq \left\|\tfrac{1}{2}(u_n^{h-1} - u^\star) + \tfrac{1}{2}(u_n^h - u^\star)\right\|_{L^p(\Omega)}^p \\ &\leq \tfrac{1}{2}\left\|u_n^{h-1} - u^\star\right\|_{L^p(\Omega)}^p + \tfrac{1}{2}\left\|u_n^h - u^\star\right\|_{L^p(\Omega)}^p - 2^{-p}c_p\left\|u_n^{h-1} - u_n^h\right\|_{L^p(\Omega)}^p \\ &\leq \left\|u_n^{h-1} - u^\star\right\|_{L^p(\Omega)}^p - 2^{-p}c_p\left\|u_n^h - u_n^{h-1}\right\|_{L^p(\Omega)}^p,\end{aligned}$$

where we used non-expansiveness of $J_{\tau_{h-1}\mathbf{\Delta}_p^{K_n}}$ to get the last inequality. $c_p = (1+\nu_p^{p-1})(1+\nu_p)^{1-p}$, where ν_p is the unique solution to $(p-2)\nu^{p-1} + (p-1)\nu^{p-2} = 1$, for $\nu \in]0,1[$. Summing up these inequalities and using the fact that

$$\left\|u_n^{h+1} - u_n^h\right\|_{L^p(\Omega)} \leq \left\|u_n^h - u_n^{h-1}\right\|_{L^p(\Omega)}$$

again by non-expansiveness of $J_{\tau_{h-1}\mathbf{\Delta}_p^{K_n}}$, we arrive at

$$\underline{\tau}h\left\|\mathbf{\Delta}_p^{K_n}(u_n^h)\right\|_{L^p(\Omega)}^p \leq h\left\|u_n^h - u_n^{h-1}\right\|_{L^p(\Omega)}^p \leq \sum_{i=1}^h \left\|u_n^i - u_n^{i-1}\right\|_{L^p(\Omega)}^p \leq 2^p\left\|g_n^0 - u^\star\right\|_{L^p(\Omega)}^p/c_p.$$

(b) $p \in]1,2]$: using now [111, Corollary 2, (3.7)] and similar arguments to the first case, we get the inequality

$$\left\|u_n^h - u^\star\right\|_{L^p(\Omega)}^2 \leq \left\|u_n^{h-1} - u^\star\right\|_{L^p(\Omega)}^2 - 2^{-2}(p-1)\left\|u_n^h - u_n^{h-1}\right\|_{L^p(\Omega)}^2.$$

Summing up again we end up with

$$\underline{\tau}h\left\|\mathbf{\Delta}_p^{K_n}(u_n^h)\right\|_{L^p(\Omega)}^2 \leq h\left\|u_n^h - u_n^{h-1}\right\|_{L^p(\Omega)}^2 \leq \sum_{i=1}^h \left\|u_n^i - u_n^{i-1}\right\|_{L^p(\Omega)}^2 \leq 4\left\|g_n^0 - u^\star\right\|_{L^p(\Omega)}^2/(p-1).$$

□

Remark 3.4.6. (a) Observe that the assumption on the initial condition in Lemma 3.4.5 is weaker than that of Lemma 3.4.1.

(b) As expected, the stability constraint needed on the time-step sequence is less restrictive than for the explicit/forward discretization.

(c) Given that $\left\{\left\|u_n^{h+1}-u_n^h\right\|_{L^p(\Omega)}^p\right\}_h$ is a decreasing and summable sequence, one can show that the rate $\left\|\mathbf{\Delta}_p^{K_n}(u_n^h)\right\|_{L^p(\Omega)} = O(h^{-1/\max(p,2)})$ is in fact $\left\|\mathbf{\Delta}_p^{K_n}(u_n^h)\right\|_{L^p(\Omega)} = o(h^{-1/\max(p,2)})$.

Equipped with this result, the proof of an analogue to Theorem 3.4.4 in the implicit case is similar to that of the explicit case modulo the following change

$$\bar{u}_n(x,t) = \sum_{h=1}^{N} u_n^h(x)\chi_{]t_{h-1},t_h]}(t).$$

3.4.3 Relation to Kobayashi type estimates

Consider the evolution problem

$$\begin{cases} u_t + A(t)u(t) \ni f(t), \\ u(0) = g. \end{cases} \tag{CP}$$

A problem of the form (CP) is called an abstract Cauchy problem. The evolution problem ($\mathcal{P}_{\text{nloc}}$) we deal with can be viewed as a particular case of (CP) in its autonomous-homogeneous case, i.e. the operator $A(t) \equiv \mathbf{\Delta}_p^K$ does not depend on time and the source term $f \equiv 0$.

Problem (CP) in the autonomous-homogeneous case was studied by Kobayashi in [73], where he constructed sequences of approximate solutions which converge in an appropriate sense to a solution to the differential inclusion. He provided an inequality that estimates the distance between arbitrary points of two independent sequences generated by the so called proximal iterations, from which, he derived quantitative estimates to compare the continuous and discrete trajectories using the backward Euler scheme. These estimates have similar flavour to ours when $K = K_n$. Later on, these results were generalized to the non-autonomous case as well as to the case where the trajectories are defined by two differential inclusions systems (i.e. different operators A); see [4] and references therein for a thorough review. The latter bounds, expressed in our notation, are provided only in terms of $\left\|\mathbf{\Delta}_p^K(v) - \mathbf{\Delta}_p^{K_n}(v)\right\|_{L^p(\Omega)}$. We go further by exploiting the properties of our operators to get sharp estimates in terms of the data $\left\|K - K_n\right\|_{L^p(\Omega^2)}$. This is more meaningful in our context where we recall that the goal is to study the fully discretized nonlocal p-Laplacian problem on graphs.

3.5 Application to a coupled nonlocal p-Laplacian evolution system

Here we present an illustration of how the consistency results that we get for problem ($\mathcal{P}_{\text{nloc}}$) can be applied in a more general context. In particular we show the consistency of a nonlocal evolution system introduced in [54].

Throughout the section, we consider the following norm

$$\forall (u,v) \in (L^p(\Omega))^2, \quad \left\|(u,v)\right\|_{C\left(0,T;(L^p(\Omega))^2\right)} = \sup_{t\in[0,T]} \left(\left\|u(t)\right\|_{L^p(\Omega)} + \left\|v(t)\right\|_{L^p(\Omega)}\right), \quad p \in]1,+\infty[, \ T > 0,$$

where u et v are

3.5.1 Problem formulation

In [54], the authors propose to study the following nonlocal evolution system:

$$\begin{cases} u_t(x,t) = -\mathbf{\Delta}_p^K(u(x,t)) - 2\lambda v(x,t), & a.e.\ x \in \Omega, t > 0, \\ v_t(x,t) = -\mathbf{\Delta}_2^K(v(x,t)) - (f(x) - u(x,t)), & a.e.\ x \in \Omega, t > 0, \\ u(x,0) = f(x), \quad v(x,0) = 0, \quad a.e\ x \in \Omega, \end{cases} \tag{$\mathcal{S}_{\text{nloc}}$}$$

where $\lambda > 0$. Here the kernel $K : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ is a nonnegative continuous smooth functions with compact support contained in $\Omega \times B(0,d) \subset \mathbb{R}^N \times \mathbb{R}^N$ with

$$0 < \sup_{y \in B(0,d)} K(x,y) = R(x) \in L^\infty(\Omega). \tag{3.5.1}$$

Furthermore, K satisfies

$$\int_{\mathbb{R}^N} K(x,y)dx = 1.$$

In [54, Theorem 2.1], the authors prove the existence and uniqueness of the solution to ($\mathcal{S}_{\text{nloc}}$) that is the couple $(u,v) \in \left[C(0,T;L^1(\Omega)) \cap W^{1,1}(0,T;L^1(\Omega))\right]^2$.

3.5.2 Consistency of the semidiscrete scheme

Let us consider the following coupled system with Neumann boundary conditions as ($\mathcal{S}_{\text{nloc}}$)

$$\begin{cases} \frac{\partial}{\partial t} u_n(x,t) = -\boldsymbol{\Delta}_p^{K_n}(u_n(x,t)) - 2\lambda v_n(x,t), & a.e.\ x \in \Omega,\ t \in [0,T], \\ \frac{\partial}{\partial t} v_n(x,t) = -\boldsymbol{\Delta}_2^{K_n}(v_n(x,t)) - (f_n(x) - u_n(x,t)), & a.e.\ x \in \Omega,\ t \in [0,T], \\ u_n(x,0) = f_n(x), \quad v_n(x,0) = 0, \quad a.e.\ x \in \Omega. \end{cases} \tag{$\mathcal{S}^n_{\text{nloc}}$}$$

As we have done in Section 3.3, the main goal is to compare the couple of solutions of ($\mathcal{S}^n_{\text{nloc}}$) to that of ($\mathcal{S}_{\text{nloc}}$) and get a uniform error bound. This is the statement of the following theorem.

Theorem 3.5.1. *Suppose $p \in]1,+\infty[$, $f, f_n \in L^\infty(\Omega)$ and K, K_n are measurable, symmetric and bounded mappings. Then ($\mathcal{S}_{\text{nloc}}$) and ($\mathcal{S}^n_{\text{nloc}}$) have unique solutions, respectively, (u,v) and (u_n,v_n). Moreover the following hold.*

(i) We have the error estimate

$$\left\|(u-u_n, v-v_n)\right\|_{C\left(0,T;(L^p(\Omega))^2\right)} \leq C\left(\left\|f-f_n\right\|_{L^p(\Omega)} + \left\|K-K_n\right\|_{L^p(\Omega^2)}\right), \tag{3.5.2}$$

where the constant C is independent of n.

(ii) Moreover, if $f_n \to f$ and $K_n \to K$ as $n \to \infty$, almost everywhere on Ω and Ω^2, respectively, then

$$\left\|(u-u_n, v-v_n)\right\|_{C\left(0,T;(L^p(\Omega))^2\right)} \underset{n\to\infty}{\longrightarrow} 0.$$

Proof : In the proof, C_i is any absolute constant independent of n (but may depend on p).

(i) For $1 < p < +\infty$, we define the function

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

Denote $\xi_n(x,t) = u_n(x,t) - u(x,t)$ and $\zeta_n(x,t) = v_n(x,t) - v(x,t)$, by subtracting ($\mathcal{S}_{\text{nloc}}$) from ($\mathcal{S}^n_{\text{nloc}}$), we have

$$\begin{aligned} \frac{\partial \xi_n}{\partial t} &= \int_\Omega K_n(x,y)\{\Psi(u_n(y,t) - u_n(x,t)) - \Psi(u(y,t) - u(x,t))\}dy \\ &+ \int_\Omega (K_n(x,y) - K(x,y))\Psi(u(y,t) - u(x,t))dy - 2\lambda\zeta_n(x,t). \end{aligned} \tag{3.5.3}$$

$$\begin{aligned} \frac{\partial \zeta_n}{\partial t} &= \int_\Omega K_n(x,y)\{(v_n(y,t) - v_n(x,t)) - (v(y,t) - v(x,t))\}dy \\ &+ \int_\Omega (K_n(x,y) - K(x,y))(v(y,t) - v(x,t))dy - (f_n - f)(x) + \xi_n(x,t). \end{aligned} \tag{3.5.4}$$

Next, we multiply both sides of (3.5.3) and (3.5.4) by $\Psi(\xi_n(x,t))$ and $\Psi(\zeta_n(x,t))$, respectively, and integrate over Ω

$$\begin{aligned}\frac{1}{p}\int_\Omega \frac{\partial}{\partial t}|\xi_n(x,t)|^p dx &= \int_{\Omega^2} K_n(x,y)\{\Psi(u_n(y,t)-u_n(x,t)) - \Psi(u(y,t)-u(x,t))\}\Psi(\xi_n(x,t))dxdy \\ &+ \int_{\Omega^2}(K_n(x,y)-K(x,y))\Psi(u(y,t)-u(x,t))\Psi(\xi_n(x,t))dxdy \\ &- 2\lambda\int_\Omega \zeta_n(x,t)\Psi(\xi_n(x,t))dx.\end{aligned} \tag{3.5.5}$$

$$\begin{aligned}\frac{1}{p}\int_\Omega \frac{\partial}{\partial t}|\zeta_n(x,t)|^p dx &= \int_{\Omega^2} K_n(x,y)\{(v_n(y,t)-v_n(x,t)) - (v(y,t)-v(x,t))\}\Psi(\zeta_n(x,t))dxdy \\ &+ \int_{\Omega^2}(K_n(x,y)-K(x,y))(v(y,t)-v(x,t))\Psi(\zeta_n(x,t))dxdy \\ &- \int_\Omega (f_n-f)(x)\Psi(\zeta_n(x,t))dx + \int_\Omega \xi_n(x,t)\Psi(\zeta_n(x,t))dx.\end{aligned} \tag{3.5.6}$$

We estimate the first term on the right-hand side of (3.5.5) using the fact that K_n is bounded so that there exists a positive constant M independent of n, such that, $\|K_n\|_{L^\infty(\Omega^2)} \le M$,

$$\begin{aligned}&\Big|\int_{\Omega^2} K_n(x,y)\{\Psi(u_n(y,t)-u_n(x,t)) - \Psi(u(y,t)-u(x,t))\}\Psi(\xi_n(x,t))dxdy\Big| \\ &\le M\int_{\Omega^2}\big|\Psi(u_n(y,t)-u_n(x,t)) - \Psi(u(y,t)-u(x,t))\big||\xi_n(x,t)|^{p-1}dxdy.\end{aligned}$$

Now, applying Corollary 2.2.18 with $a = u_n(y,t) - u_n(x,t)$ and $b = u(y,t) - u(x,t)$ (without loss of generality we assume that $b > a$), we get

$$\begin{aligned}&\int_{\Omega^2}\big|\Psi(u_n(y,t)-u_n(x,t)) - \Psi(u(y,t)-u(x,t))\big||\xi_n(x,t)|^{p-1}dxdy \\ &\le (p-1)\int_{\Omega^2}|\xi_n(y,t)-\xi_n(x,t)||\eta(x,y,t)|^{p-2}|\xi_n(x,t)|^{p-1}dxdy,\end{aligned} \tag{3.5.7}$$

where $\eta(x,y,t)$ is an intermediate value between a and b. As we have supposed that $f \in L^\infty(\Omega)$ and $f_n \in L^\infty(\Omega)$, and as Ω is a compact set, so that $L^\infty(\Omega) \subset L^p(\Omega)$, we deduce

$$\begin{aligned}|\eta(x,y,t)|^{p-2} \le |u(y,t)-u(x,t)|^{p-2} &\le \Big(2\|u(t)\|_{L^\infty(\Omega)}\Big)^{p-2} \\ &\le C_1.\end{aligned} \tag{3.5.8}$$

Inserting (3.5.8) into (3.5.7), and then using the Hölder and triangle inequalities, it follows that

$$\begin{aligned}&M\int_{\Omega^2}\big|\Psi(u_n(y,t)-u_n(x,t)) - \Psi(u(y,t)-u(x,t))\big||\xi_n(x,t)|^{p-1}dxdy \\ &\le M(p-1)C_1\int_{\Omega^2}|\xi_n(y,t)-\xi_n(x,t)||\xi_n(x,t)|^{p-1}dxdy \\ &= C_2\int_{\Omega^2}|\xi_n(y,t)-\xi_n(x,t)||\xi_n(x,t)|^{p-1}dxdy \\ &\le C_2\left(\int_{\Omega^2}|\xi_n(y,t)-\xi_n(x,t)|^p dxdy\right)^{\frac{1}{p}} \times \left(\int_\Omega |\xi_n(x,t)|^p dx\right)^{\frac{p-1}{p}} \\ &\le 2C_2\|\xi_n(t)\|^p_{L^p(\Omega)}.\end{aligned} \tag{3.5.9}$$

We bound the second term on the right-hand side of (3.5.5) as follows

$$
\begin{aligned}
&\Big|\int_{\Omega^2}(K_n(x,y)-K(x,y))\Psi(u(y,t)-u(x,t))\Psi(\xi_n(x,t))dxdy\Big| \\
&=\Big|\int_{\Omega^2}(K_n(x,y)-K(x,y))\times \operatorname{sign}(u(y,t)-u(x,t))\big|u(y,t)-u(x,t)\big|^{p-1}\Psi(\xi_n(x,t))dxdy\Big| \\
&\leq 2^{1/p}\big\|u(t)\big\|^{p-1}_{L^\infty(\Omega)}\Big|\int_{\Omega^2}\big|K_n(x,y)-K(x,y)\big|\big|\xi_n(x,t)\big|^{p-1}dxdy\Big| \\
&\leq 2^{1/p}\big\|u(t)\big\|^{p-1}_{L^\infty(\Omega)}\times\left(\int_{\Omega^2}\big|K_n(x,y)-K(x,y)\big|^p dxdy\right)^{\frac{1}{p}}\left(\int_{\Omega}\big|\xi_n(x,t)\big|^p dx\right)^{\frac{p-1}{p}} \\
&\leq 2C_3\big\|K_n-K\big\|_{L^p(\Omega^2)}\big\|\xi_n(t)\big\|^{p-1}_{L^p(\Omega)}.
\end{aligned}
\tag{3.5.10}
$$

Using the hölder inequality, we estimate the last term on the right-hand side of (3.5.5)

$$
\begin{aligned}
2\lambda\Big|\int_\Omega \zeta_n(x,t)\Psi(\xi_n(x,t))dx\Big| &\leq 2\lambda\int_\Omega\big|\zeta_n(x,t)\big|\big|\xi_n(x,t)\big|^{p-1}dx \\
&\leq 2\lambda\big\|\zeta_n(t)\big\|_{L^p(\Omega)}\big\|\xi_n(t)\big\|^{p-1}_{L^p(\Omega)}.
\end{aligned}
\tag{3.5.11}
$$

Similarly to before, we estimate (3.5.6) using the same arguments with the function $\Psi(x)=x$ (for $p=2$), to obtain

$$
\begin{aligned}
&\Big|\int_{\Omega^2}K_n(x,y)\{(v_n(y,t)-v_n(x,t))-(v(y,t)-v(x,t))\}\Psi(\zeta_n(x,t))dxdy\Big| \\
&\leq M\int_{\Omega^2}\big|\zeta_n(y,t)-\zeta_n(x,t)\big|\big|\zeta_n(x,t)\big|^{p-1}dxdy \\
&\leq 2M\big\|\zeta_n(t)\big\|^p_{L^p(\Omega)}.
\end{aligned}
\tag{3.5.12}
$$

$$
\Big|\int_{\Omega^2}(K_n(x,y)-K(x,y))(v(y,t)-v(x,t))\Psi(\zeta_n(x,t))dxdy\Big|\leq 2C_4\big\|K_n-K\big\|_{L^p(\Omega^2)}\big\|\zeta_n(t)\big\|^{p-1}_{L^p(\Omega)}.
\tag{3.5.13}
$$

Applying the Hölder inequality to the last term on the right-hand side of (3.5.6), we get

$$
\begin{aligned}
&\Big|\int_\Omega\xi_n(x,t)\Psi(\zeta_n(x,t))dx-\int_\Omega(f_n-f)(x)\Psi(\zeta_n(x,t))dx\Big| \\
&\leq\big\|f_n-f\big\|_{L^p(\Omega)}\big\|\zeta_n(t)\big\|^{p-1}_{L^p(\Omega)}+\big\|\xi_n(t)\big\|_{L^p(\Omega)}\big\|\zeta_n(t)\big\|^{p-1}_{L^p(\Omega)}.
\end{aligned}
\tag{3.5.14}
$$

So, putting together (3.5.9), (3.5.10) and (3.5.11), we have

$$
\begin{aligned}
\frac{d}{dt}\big\|\xi_n(t)\big\|^p_{L^p(\Omega)} &\leq 2pC_2\big\|\xi_n(t)\big\|^p_{L^p(\Omega)}+2pC_3\big\|K_n-K\big\|_{L^p(\Omega^2)}\big\|\xi_n(t)\big\|^{p-1}_{L^p(\Omega)} \\
&\quad+2p\lambda\big\|\zeta_n(t)\big\|_{L^p(\Omega)}\big\|\xi_n(t)\big\|^{p-1}_{L^p(\Omega)} \\
&=2pC_2\big\|\xi_n(t)\big\|^p_{L^p(\Omega)}+\Big(2pC_3\big\|K_n-K\big\|_{L^p(\Omega^2)}+2p\lambda\big\|\zeta_n(t)\big\|_{L^p(\Omega)}\Big)\big\|\xi_n(t)\big\|^{p-1}_{L^p(\Omega)}.
\end{aligned}
\tag{3.5.15}
$$

Next, combining (3.5.12), (3.5.13) and (3.5.14), we get

$$
\begin{aligned}
\frac{d}{dt}\big\|\zeta_n(t)\big\|^p_{L^p(\Omega)} &\leq 2pM\big\|\zeta_n(t)\big\|^p_{L^p(\Omega)}+2pC_3\big\|K_n-K\big\|_{L^p(\Omega^2)}\big\|\zeta_n(t)\big\|^{p-1}_{L^p(\Omega)} \\
&\quad+p\big\|f_n-f\big\|_{L^p(\Omega)}\big\|\zeta_n(t)\big\|^{p-1}_{L^p(\Omega)}+p\big\|\xi_n(t)\big\|_{L^p(\Omega)}\big\|\zeta_n(t)\big\|^{p-1}_{L^p(\Omega)}.
\end{aligned}
\tag{3.5.16}
$$

Adopting the same strategy [2] as in the proof of Theorem 3.4.4, we obtain

$$
\frac{d}{dt}\big\|\xi_n(t)\big\|_{L^p(\Omega)}\leq 2C_2\big\|\xi_n(t)\big\|_{L^p(\Omega)}+2C_3\big\|K_n-K\big\|_{L^p(\Omega^2)}+2\lambda\big\|\zeta_n(t)\big\|_{L^p(\Omega)}
\tag{3.5.17}
$$

[2] Using the function $\psi_\varepsilon(\cdot)$.

and

$$\frac{d}{dt}\|\zeta_n(t)\|_{L^p(\Omega)} \leq 2M\|\zeta_n(t)\|_{L^p(\Omega)}+2C_3\|K_n-K\|_{L^p(\Omega^2)}+\|\xi_n(t)\|_{L^p(\Omega)}+\|f_n-f\|_{L^p(\Omega)}. \tag{3.5.18}$$

Summing up (3.5.17) and (3.5.18), we have the following inequality

$$\frac{d}{dt}\Big(\|\xi_n(t)\|_{L^p(\Omega)}+\|\zeta_n(t)\|_{L^p(\Omega)}\Big) \leq \underbrace{(2\max(M,C_2)+\max(2\lambda,1))}_{=C(\lambda,p)}\Big(\|\xi_n(t)\|_{L^p(\Omega)}+\|\zeta_n(t)\|_{L^p(\Omega)}\Big)$$
$$+\,4C_3\|K_n-K\|_{L^p(\Omega^2)}+\|f_n-f\|_{L^p(\Omega)}. \tag{3.5.19}$$

We apply the Gronwall's inequality on $[0,T]$ to get

$$\sup_{t\in[0,T]}\Big(\|\xi_n(t)\|_{L^p(\Omega)}+\|\zeta_n(t)\|_{L^p(\Omega)}\Big) \leq \Big(\|f_n-f\|_{L^p(\Omega)}+4C_3T\|K_n-K\|_{L^p(\Omega^2)}\Big)\exp\{2CT\}. \tag{3.5.20}$$

Since we have

$$\|(u-u_n),(v-v_n)\|_{C\left(0,T;(L^p(\Omega))^2\right)} \leq \sup_{t\in[0,T]}\Big(\|u(t)-u_n(t)\|_{L^p(\Omega)}+\|v(t)-v_n(t)\|_{L^p(\Omega)}\Big),$$

then, the desired result holds.

(ii) It follows immediately from the Scheffe-Riesz theorem (see [74, Lemma 2]).

3.5.3 Consistency of the fully discrete scheme

We now consider the following time-discrete approximation of ($\mathcal{S}_{\text{nloc}}$), the forward Euler discretization applied to ($\mathcal{S}^n_{\text{nloc}}$). For that, as we have done before, we take again the partition $\{\tau_h\}_{h=1}^N$ of the time interval $[0,T]$ of maximal size $\tau=\max_{h\in[N]}\tau_h$, i.e; $\tau_{h-1}\stackrel{\text{def}}{=}|t_h-t_{h-1}|$ and let $u_n^h(x)\stackrel{\text{def}}{=}u_n(x,t_h)$, $v_n^h(x)\stackrel{\text{def}}{=}u_n(x,t_h)$. Then, consider

$$\begin{cases}\dfrac{u_n^h(x)-u_n^{h-1}(x)}{\tau_{h-1}}=-\boldsymbol{\Delta}_p^{K_n}u_n^{h-1}(x)-2\lambda v_n^{h-1}(x), & a.e.\ x\in\Omega, h\in[N],\\ \dfrac{v_n^h(x)-v_n^{h-1}(x)}{\tau_{h-1}}=-\boldsymbol{\Delta}_2^{K_n}v_n^{h-1}-(f_n(x)-u_n^{h-1}(x)), & a.a.\ x\in\Omega, h\in[N],\\ u_n^0=f_n(x),\quad v_n^0=0,\quad a.e.\ x\in\Omega. \end{cases} \tag{$\mathcal{S}^f_{\text{nloc},\tau}$}$$

Since the aim is to compare the solutions of problems ($\mathcal{S}_{\text{nloc}}$) and ($\mathcal{S}^f_{\text{nloc},\tau}$), the solution of ($\mathcal{S}^f_{\text{nloc},\tau}$) being discrete, so that it is convenient to introduce an intermediate model which is the continuous extension of the discrete problem using the discrete functions $u_n(x)=(u_n^1(x),\cdots,u_n^N(x))$ and $v_n(x)=(v_n^1(x),\cdots,v_n^N(x))$. Therefore, we consider a time-continuous extensions of u_n^h and v_n^h, respectively, obtained by a linear interpolations as follows

$$\check{u}_n(x,t)=\frac{t_h-t}{\tau_{h-1}}u_n^{h-1}(x)+\frac{t-t_{h-1}}{\tau_{h-1}}u_n^h(x),\quad t\in]t_{h-1},t_h],\quad x\in\Omega, \tag{3.5.21}$$

$$\check{v}_n(x,t)=\frac{t_h-t}{\tau_{h-1}}v_n^{h-1}(x)+\frac{t-t_{h-1}}{\tau_{h-1}}v_n^h(x),\quad t\in]t_{h-1},t_h],\quad x\in\Omega, \tag{3.5.22}$$

and a time piecewise-constant approximations

$$\bar{u}_n(x,t)=\sum_{h=1}^N u_n^{h-1}(x)\chi_{]t_{h-1},t_h]}(t), \tag{3.5.23}$$

$$\bar{v}_n(x,t)=\sum_{h=1}^N v_n^{h-1}(x)\chi_{]t_{h-1},t_h]}(t). \tag{3.5.24}$$

Then, by construction of $\check{u}_n(x,t)$ and $\check{v}_n(x,t)$, we have the following evolution system

$$\begin{cases} \frac{\partial}{\partial t}\check{u}_n(x,t) = -\mathbf{\Delta}_p^{K_n}(\bar{u}_n(x,t)) - 2\lambda\bar{v}_n(x,t), & (x,t) \in \Omega\times]0,T], \\ \frac{\partial}{\partial t}\check{v}_n(x,t) = -\mathbf{\Delta}_2^{K_n}(\bar{v}_n(x,t)) - (f_n(x) - \bar{u}_n(x,t)), & (x,t) \in \Omega\times]0,T], \\ \check{u}_n(x,0) = f_n(x), \check{v}_n(x,0) = 0 \quad x \in \Omega. \end{cases} \tag{3.5.25}$$

We have the following convergence result.

Theorem 3.5.2. *Suppose $p \in]1,+\infty[$, $f, f_n \in L^\infty(\Omega)$ and K, K_n are measurable, symmetric and bounded mappings.*

Let (u,v) be the unique couple of solutions of system ($\mathcal{S}_{\mathrm{nloc}}$), and $\check{u}_n$, $\check{v}_n$ are built as in (3.5.21) and (3.5.22), respectively, from the time-discrete approximations u_n^{h-1} and v_n^{h-1} defined in ($\mathcal{S}^f_{\mathrm{nloc},\tau}$), respectively. Then

$$\left\|(u-u_n, v-v_n)\right\|_{C\left(0,T;(L^p(\Omega))^2\right)} \leq C\left(\left\|f-f_n\right\|_{L^p(\Omega)} + \left\|K-K_n\right\|_{L^p(\Omega^2)}\right) + O(\tau), \tag{3.5.26}$$

where the constant C is independent of n.

Proof : We follow the same lines as in the proof of Theorem 3.3.1. Denote $\check{\xi}_n(x,t) = \check{u}_n(x,t) - u_n(x,t)$ and $\bar{\xi}_n(x,t) = \bar{u}_n(x,t) - u_n(x,t)$ and $\check{\zeta}_n(x,t) = \check{v}_n(x,t) - v_n(x,t)$, $\bar{\zeta}_n(x,t) = \bar{v}_n(x,t) - v_n(x,t)$.

The result of Lemma 3.4.3 remains the same for v, we have

$$\left\|\bar{v}_n(t) - \check{v}_n(t)\right\|_{L^p(\Omega)} = O(\tau), \qquad t \in [0,T]. \tag{3.5.27}$$

Therefrom, we follow the same lines of the proof of Theorem 3.4.4, by fitting it as we have done in that of Theorem 3.3.1 dealing with $\check{\xi}_n(x,t)$ and $\check{\zeta}_n(x,t)$ and applying the Gronwall's lemma separately for each function, combined with (3.5.27) we get the desired result. □

Chapter 4

Convergence Rates for Networks on Convergent Graph Sequences

Main contributions of this chapter

- We apply the error estimate of Chapter 3 to dynamical networks on convergent graph sequences (simple and weighted dense deterministic graphs first and random inhomogeneous ones second).
- We show that the approximation of solutions of the discrete problems on these graph sequences converge to those of the continuum problem.
- We quantify also the rate of convergence for each graph model.
- We reveal the role of the data regularity and the parameter p on the rate of convergence.

The content on deterministic graphs is published in [66]. The case of random graphs is at the heart of [65].

Contents

In this chapter, relying on the general error estimate we obtained in the previous chapter, we deal with networks on convergent graph sequences and quantify the rate of convergence of the discrete solution to the continuous one for two categories of graph sequences.

(i) Deterministic simple and weighted, dense graphs. For weighted graphs, we also investigated the limit as $p \to \infty$ of the discrete model.

(ii) Random inhomogeneous weighted graphs.

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

$$\|u\|_{p,n} = \left(\frac{1}{n}\sum_{i=1}^{n} |u_i|^p\right)^{\frac{1}{p}}. \tag{4.0.1}$$

4.1 Introduction

4.1.1 Problem statement

Recall the discrete form ($\mathcal{P}^d_{\text{nloc}}$) of problem ($\mathcal{P}_{\text{nloc}}$) on a graph sequence $G_n = ([n], E(G_n)))$ from Section 1.1.1. For that, let's redefine the partition (not necessarily uniform) $\{t_h\}_{h=1}^N$ of the time interval $[0, T]$. Let $\tau_{h-1} \stackrel{\text{def}}{=} |t_h - t_{h-1}|$ and the maximal size $\tau = \max_{h\in[N]} \tau_h$

$$\begin{cases} \dfrac{u_i^h - u_i^{h-1}}{\tau_{h-1}} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} K_{nij}\left|u_j^{h-1} - u_i^{h-1}\right|^{p-2}(u_j^{h-1} - u_i^{h-1}), \quad (i,h) \in [n]\times[N], \\ \\ u_i(0) = g_i^0, \quad i \in [n]. \end{cases} \tag{$\mathcal{P}^d_{\text{nloc}}$}$$

As we have explained in the introduction of the manuscript, $(K_{nij})_{1\le i,j\le n}$ is seen as the adjacency matrix of the graph sequence G_n. Its explicit form will be clear later when dealing with each graph model, keeping in mind that the graph sequence converges to a given *graphon*. Thus, ($\mathcal{P}^d_{\text{nloc}}$) induces a discrete diffusion process parametrized by the structure of the graph whose adjacency matrix captures the (nonlocal) interactions. The initial condition $g^0 = (g_1^0, \cdots, g_n^0)^\top$ will also be defined explicitly from the continuous initial data $g(\cdot)$ of ($\mathcal{P}_{\text{nloc}}$).

4.1.2 Related work

Dealing with networks on convergent graph sequences, an important work focusing in this subject is that in [83, 84, 70] which paved the way to study limit phenomena of evolution problems on both deterministic and random (dense and sparse) graphs.

4.1.2.1 Networks on deterministic graphs

In [83], the author focused on a nonlinear (nonlocal) heat evolution equation on graphs, where the operator $\mathbf{\Delta}_p^K$ was replaced by the operator

$$\boldsymbol{D}^W(u(x)) = -\int_\Omega W(x,y)D(u(y)-u(x))dy,$$

with $W(\cdot,\cdot)$ verifying Assumption (**A.2**) and in which the function D was assumed Lipschitz-continuous. This assumption was essential to prove well-posedness (existence and uniqueness follow immediately from the contraction principle), as well as to study the consistency in L^2-norm of the spatial semi-discrete approximation on simple and weighted graph sequences. Though this seminal work was quite inspiring to us, it differs from our work in many crucial aspects. First, the nonlocal p-Laplacian evolution problem at hand is different and cannot be covered by [83] where the function $x \mapsto x|x|^{p-2}$ lacks Lipschitzianity for $p \in]1,+\infty[$, and thus raises several challenges (including for well-posedness and error estimates). We also consider both the semi-discrete and fully-discrete versions with both forward and backward Euler approximations, that we fully characterize which is not the case in [83] where only the semi-discrete scheme was considered (so that no consistency proof was needed when dealing with networks on graphs). In addition to that, for networks on weighted graphs only the uniform convergence of the discrete problem to the continuous nonlocal heat equation (2-Laplacian) was established, we go further by quantifying the rate of convergence of ($\mathcal{P}^d_{\mathrm{nloc}}$) to ($\mathcal{P}_{\mathrm{nloc}}$) and giving a non-asymptotic error bound.

4.1.2.2 Networks on random graphs

In [84] and earlier [83], 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 [84] 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. 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 [83, 84] (see [66]). Second, comparing [84] and our current work, we use completely different paths to prove consistency in the random case. Indeed, while the claim in [84] 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 4.4.10). 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 [84].

In [70], the authors extended the analysis of [84] 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 [82] 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 [70, 82]. 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 ($\mathcal{P}_{\text{nloc}}$) 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 4.4.5(v). Observe finally that the convergence claim of [70] is asymptotic (almost sure convergence), relying on the standard Markov inequality and Borel-Cantelli lemma, while ours are nonasymptotic with a precise probability of success.

4.2 Networks on simple graphs

We begin our study by dealing with the simplest graph model that we defined in Section 2.1.3.1. Remember briefly that this graph model converges to the $\{0,1\}$-graphon.

A fully discrete counterpart of ($\mathcal{P}_{\text{nloc}}$) on $\{G_n\}_n$ is given by

$$\begin{cases} \dfrac{u_i^h - u_i^{h-1}}{\tau_{h-1}} = \dfrac{1}{n} \displaystyle\sum_{j:(i,j)\in E(G_n)} \left|u_j^{h-1} - u_i^{h-1}\right|^{p-2}(u_j^{h-1} - u_i^{h-1}), \quad (i,h) \in [n] \times [N], \\ \\ u_i(0) = g_i^0, \quad i \in [n], \end{cases} \tag{$\mathcal{P}^{s,d}_{\text{nloc}}$}$$

where

$$g_i^0 = n \int_{\Omega_i^{(n)}} g_n^0(x)dx$$

is the average value of $g_n^0(x)$ on $\Omega_i^{(n)}$.

Let us recall that our main goal is to compare the solutions of the discrete and continuous models and establish some consistency results. Since the two solutions do not live on the same spaces, it is practical to represent some intermediate model that is the continuous extension of the discrete problem, using the vector $U^h = (u_1^h, u_2^h, \cdots, u_n^h)^T$ whose components uniquely solve the previous system ($\mathcal{P}^{s,d}_{\text{nloc}}$) (see Lemma 3.4.1) to obtain the following piecewise time linear interpolation on $\Omega \times [0,T]$

$$\check{u}_n(x,t) = \frac{t_h - t}{\tau_{h-1}} u_i^{h-1} + \frac{t - t_{h-1}}{\tau_{h-1}} u_i^h \quad \text{if} \quad x \in \Omega_i^{(n)}, \quad t \in]t_{h-1}, t_h], \tag{4.2.1}$$

and the following piecewise constant approximation

$$\bar{u}_n(x,t) = \sum_{i=1}^{n} \sum_{h=1}^{N} u_i^{h-1} \chi_{]t_{h-1},t_h]}(t) \chi_{\Omega_i^{(n)}}(x). \tag{4.2.2}$$

So that $\check{u}_n$ uniquely solves the following problem

$$\begin{cases} \frac{\partial}{\partial t}\check{u}_n(x,t) = -\mathbf{\Delta}_p^{K_n^s}(\bar{u}_n(x,t)), \quad (x,t) \in \Omega \times]0,T], \\ \check{u}_n^0(x) = g_n^0(x), \quad x \in \Omega, \end{cases} \tag{$\mathcal{P}^{s}_{\text{nloc}}$}$$

where

$$g_n^0(x) = g_i \stackrel{\text{def}}{=} n \int_{\Omega_i^{(n)}} g_n(x)dx \quad \text{if} \quad x \in \Omega_i^{(n)}, i \in [n],$$

g_n being the initial condition taken in problem ($\mathcal{P}^{n}_{\text{nloc}}$) and $K_n^s(x,y)$ is the piecewise constant function such that for $(x,y) \in \Omega_{ij}^{(n)}$, $(i,j) \in [n]^2$

$$\begin{cases} n^2 \int_{\Omega_{ij}^{(n)}} K(x,y)dxdy & \text{if} \quad \Omega_i^{(n)} \times \Omega_j^{(n)} \cap \overline{\mathrm{supp}(K)} \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

As G_n is a simple graph, K_n^s is a $\{0,1\}$-valued mapping.

By analogy of what was done in [83], the rate of convergence of the solution of the discrete problem to the solution of the limiting problem depends on the regularity of the boundary $\mathrm{bd}(\overline{\mathrm{supp}(K)})$ of the support closure. Following [83], we recall the upper box-counting (or Minkowski-Bouligand) dimension of $\mathrm{bd}(\overline{\mathrm{supp}(K)})$ as a subset of $\mathbb{R}^2$:

$$\rho \stackrel{\text{def}}{=} \dim_B(\mathrm{bd}(\overline{\mathrm{supp}(K)})) = \limsup_{\delta \to 0} \frac{\log N_\delta(\mathrm{bd}(\overline{\mathrm{supp}(K)}))}{-\log \delta}, \tag{4.2.3}$$

where $N_\delta(\mathrm{bd}(\overline{\mathrm{supp}(K)}))$ is the number of cells of a $(\delta \times \delta)$-mesh that intersect $\mathrm{bd}(\overline{\mathrm{supp}(K)})$ (see [55]).

Corollary 4.2.1. *Suppose that $p \in]1,+\infty[$, $g \in L^\infty(\Omega)$, and*

$$\rho \in [0,2[.$$

Let u and $\breve{u}_n$ denote the functions corresponding to the solutions of $(\mathcal{P}_{\mathrm{nloc}})$ and $(\mathcal{P}^s_{\mathrm{nloc}})$, respectively.

Then for any $\epsilon > 0$ there exists $N(\epsilon) \in \mathbb{N}$ such that for any $n \geq N(\epsilon)$

$$\|u - \breve{u}_n\|_{C(0,T;L^p(\Omega))} \leq C\left(\|g - g_n\|_{L^p(\Omega)} + n^{-((2-\rho)/p-\epsilon)}\right) + O(\tau), \tag{4.2.4}$$

where the positive constant C is independent of n.

Proof : By Theorem 3.4.4, we have

$$\|u - \breve{u}_n\|_{C(0,T;L^p(\Omega))} \leq C\left(\|g - g_n\|_{L^p(\Omega)} + \|g_n - g_n^0\|_{L^p(\Omega)} + \|K - K_n^s\|_{L^p(\Omega)}\right) + O(\tau). \tag{4.2.5}$$

Since both $(\mathcal{P}^s_{\mathrm{nloc}})$ and $(\mathcal{P}^{s,d}_{\mathrm{nloc}})$ problems share the same initial data, we have that $\|g_n - g_n^0\|_{L^p(\Omega)} = 0$. It remains to estimate $\|K - K_n^s\|_{L^p(\Omega)}$. To do this, we follow the same proof strategy as in [83, Theorem 4.1] . For that, consider the set of discrete cells $\Omega_{ij}^{(n)}$ overlying the boundary of the support of K

$$S(n) = \left\{(i,j) \in [n]^2 : \ \Omega_{ij}^{(n)} \cap \mathrm{bd}(\overline{\mathrm{supp}(K)}) \neq \emptyset\right\}.$$

For any $\epsilon > 0$ and sufficiently large n, we have

$$|S(n)| \leq n^{\rho+\epsilon}.$$

It is easy to see that K and K_n^s coincide almost everywhere on cells $\Omega_{ij}^{(n)}$ for which $(i,j) \notin S(n)$. Thus for any $\epsilon > 0$ and all sufficiently large n, we have

$$\|K - K_n^s\|^p_{L^p(\Omega^2)} = \int_{\Omega^2} |K(x,y) - K_n^s(x,y)|^p dxdy \leq |S(n)| n^{-2} \leq n^{-(2-\rho-\epsilon)}. \tag{4.2.6}$$

Assembling (4.2.5) and (4.2.6), the desired result holds. □

4.3 Networks on weighted graphs

In this section, we deal with the weighted graph models defined in Section 2.1.3.2.

4.3.1 Networks on $K/\mathcal{Q}_n$

We consider the totally discrete counterpart of ($\mathcal{P}_{\text{nloc}}$) on $K/\mathcal{Q}_n$

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

where $\hat{K}_n$ is defined in (2.1.8) and g_i^0 is the average value of $g_n^0(x)$ on $\Omega_i^{(n)}$.

Combining the piecewise constant function $\check{u}_n$ in (4.2.1) with $\bar{u}_n$ in (4.2.2), we rewrite ($\hat{\mathcal{P}}^{w,d}_{\text{nloc}}$) as

$$\begin{cases} \frac{\partial}{\partial t}\check{u}_n(x,t) = -\boldsymbol{\Delta}_p^{\hat{K}_n^w}(\bar{u}_n(x,t)), \quad (x,t) \in \Omega\times]0,T], \\ \check{u}_n^0(x) = g_n^0(x), \quad x \in \Omega, \end{cases} \tag{$\hat{\mathcal{P}}^{w}_{\text{nloc}}$}$$

where $\hat{K}_n^w$ and g_n^0 are the piecewise constant functions such that

$$\hat{K}_n^w(x,y) = (\hat{K}_n)_{ij} \quad \text{for} \quad (x,y) \in \Omega_i^{(n)} \times \Omega_j^{(n)},$$

$$g_n^0(x) = g_i \quad \text{for} \quad x \in \Omega_i^{(n)}, i \in [n].$$

Remark 4.3.1. As already emphasized in [83, Remark 5.1], it is instructive to note that ($\hat{\mathcal{P}}^{w}_{\text{nloc}}$) can be viewed as the time discretized Galerkin approximation of problem ($\mathcal{P}_{\text{nloc}}$). Indeed, let V_n denote a n-dimensional subspace of $L^\infty(\Omega)$

$$V_n = \text{Span}\left\{\chi_{\Omega_i^{(n)}} : i \in [n]\right\}.$$

Replacing $u(x,t)$ in ($\mathcal{P}_{\text{nloc}}$) with

$$\check{u}_n(x,t) = \sum_{k=1}^{n}\check{u}_k(t)\chi_{\Omega_k^{(n)}}(x) \in V_n,$$

where

$$\check{u}_k(t) = \frac{t_h - t}{\tau_{h-1}}u_k^{h-1} + \frac{t - t_{h-1}}{\tau_{h-1}}u_k^h, \quad t \in]t_{h-1}, t_h],$$

and projecting the resulting equation on V_n, we arrive at ($\hat{\mathcal{P}}^{w,d}_{\text{nloc}}$).

Corollary 4.3.2. *Suppose that $p \in]1,+\infty[$, $K : \Omega^2 \to [0,1]$ is a symmetric measurable function, and $g \in L^\infty(\Omega)$. Let u and $\check{u}_n$ be the solutions of ($\mathcal{P}_{\text{nloc}}$) and ($\hat{\mathcal{P}}^{w}_{\text{nloc}}$), respectively. Then*

$$\left\|u - \check{u}_n\right\|_{C(0,T;L^p(\Omega))} \underset{n\to\infty,\tau\to 0}{\longrightarrow} 0. \tag{4.3.1}$$

Proof : This proof strategy was used in [83, Theorem 5.2]. For fixed $(i,j) \in [n]^2$, it is easy to see that $\{\Omega_{ij}^{(n)}\}_n$ is a decreasing sequence, $\bigcap_{n=1}^{\infty}\Omega_{ij}^{(n)} = \{(x,y)\}$, and

$$(\hat{K}_n)_{ij} = \frac{1}{\left|\Omega_{ij}^{(n)}\right|}\int_{\Omega_{ij}^{(n)}} K_n(x,y)dxdy.$$

Then, by the Lebesgue differentiation theorem (see e.g. [89, Theorem 3.4.4]), we have

$$\hat{K}_n^w \underset{n\to\infty}{\longrightarrow} K,$$

almost everywhere on Ω^2, whence, using the same arguments on $\mathbb{R}$, we have also that $g_n \underset{n\to\infty}{\longrightarrow} g$ almost everywhere on Ω. Thus, combining Theorem 3.4.4 and statement (ii) in Theorem 3.3.1, the desired result follows. □

To quantify the rate of convergence in (4.3.1), we need to add some supplementary assumptions on the kernel K and the initial data g. This is where the Lipschitz spaces introduced in Section 2.3 play a prominent role.

We are in position to state the following error bound.

Corollary 4.3.3. *Suppose that $p \in]1, +\infty[$, $K : \Omega^2 \to [0, 1]$ is a symmetric and measurable function in $\mathrm{Lip}(s, L^p(\Omega^2))$, and $g \in \mathrm{Lip}(s, L^p(\Omega)) \cap L^\infty(\Omega)$, $s \in]0, 1]$. Let u and $\check{u}_n$ be the solutions of $(\mathcal{P}_{\mathrm{nloc}})$ and $(\hat{\mathcal{P}}^w_{\mathrm{nloc}})$ respectively. Then*

$$\left\|u - \check{u}_n\right\|_{C(0,T;L^p(\Omega))} \leq O(n^{-s}) + O(\tau). \tag{4.3.2}$$

If $\mathrm{Lip}(s, L^p(\Omega^2))$ is replaced with $\mathrm{BV}(\Omega^2)$, then the rate becomes

$$\left\|u - \check{u}_n\right\|_{C(0,T;L^p(\Omega))} \leq O(n^{-1/p}) + O(\tau). \tag{4.3.3}$$

Proof : By Theorem 3.4.4, we have

$$\left\|u - \check{u}_n\right\|_{C(0,T;L^p(\Omega))} \leq C\left(\left\|g - g_n\right\|_{L^p(\Omega)} + \left\|g_n - g_n^0\right\|_{L^p(\Omega)} + \left\|K - \hat{K}_n^w\right\|_{L^p(\Omega)}\right) + O(\tau).$$

Since the initial conditions for both $(\hat{\mathcal{P}}^{w,d}_{\mathrm{nloc}})$ and $(\hat{\mathcal{P}}^w_{\mathrm{nloc}})$ stem from the same initial data, we have that $\left\|g_n - g_n^0\right\|_{L^p(\Omega)} = 0$. The claimed rates then follow by invoking Lemma 2.3.2 since $\hat{K}_n^w = \mathbf{P}_{V_{n^2}}(K)$ and $g_n = \mathbf{P}_{V_n}(g)$. □

4.3.1.1 The limit as $p \to \infty$

Let us consider the numerical fully discrete approximation of the problem $(\mathcal{P}_{\mathrm{nloc}})$ using the function $\hat{K}_n$ defined in (2.1.8)

$$\begin{cases} \dfrac{U^p_{i,h} - U^p_{i,h-1}}{\tau_{h-1}} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} (\hat{K}_n)_{ij}\left|U^p_{j,h-1} - U^p_{i,h-1}\right|^{p-2}(U^p_{j,h-1} - U^p_{i,h-1}), \quad (i,h) \in [n] \times [N], \\ U^p_{i,0} = g_i^0, \quad i \in [n], \end{cases} \tag{4.3.4}$$

where the vector $U^p \in \mathbb{R}^{nN}$. This problem is associated to the energy functional

$$R_p(V) = \frac{1}{2pn^2}\sum_{i=1}^{n}\sum_{j=1}^{n} (\hat{K}_n)_{ij}\left|V_j - V_i\right|^p,$$

in the Euclidean space $\mathbb{R}^n$.

As before, we consider the linear interpolation of U^p as follows

$$\mathbb{R}^n \ni \check{U}^p(t) = \frac{t_h - t}{\tau_{h-1}}U^p_{h-1} + \frac{t - t_{h-1}}{\tau_{h-1}}U^p_h, \quad t \in]t_{h-1}, t_h], \tag{4.3.5}$$

and a piecewise constant approximation

$$\mathbb{R}^n \ni \bar{U}^p(t) = U^p_h, \quad t \in]t_{h-1}, t_h]. \tag{4.3.6}$$

Consequently, $\check{U}^p$ obeys the following evolution equation

$$\begin{cases} \dfrac{d\check{U}^p(t)}{dt} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} (\hat{K}_n)_{ij}\left|\bar{U}^p_j(t) - \bar{U}^p_i(t)\right|^{p-2}(\bar{U}^p_j(t) - \bar{U}^p_i(t)), \quad (i,t) \in [n]\times]0,T], \\ U^p_i(0) = g_i^0, \quad i \in [n]. \end{cases} \tag{4.3.7}$$

Now we define

$$\begin{cases} \dfrac{dU^p(t)}{dt} = \dfrac{1}{n}\sum_{j=1}^{n}(K_n)_{ij}\left|U_j^p(t) - U_i^p(t)\right|^{p-2}(U_j^p(t) - U_i^p(t)), \quad (i,t) \in [n]\times]0,T], \\ U_i^p(0) = g_i^0, \quad i \in [n]. \end{cases} \tag{4.3.8}$$

To avoid triviality, we suppose that $\text{supp}(\hat{K_n}) \neq \emptyset$, and define the non-empty compact convex set

$$\mathcal{S}_\infty = \left\{ v \in \mathbb{R}^{nN} : \left|v_j - v_i\right| \leq 1, \quad \text{for} \quad (i,j) \in \text{supp}(\hat{K_n}) \right\},$$

where the subscript ∞ will be made clear shortly. Indeed, taking the limit as $p \to \infty$ of R_p, one clearly sees that this limit is $\imath_{\mathcal{S}_\infty}$ (see Definition 2.2.8). Then, the nonlocal time continuous limit problem can be written as

$$\begin{cases} \dfrac{dU^\infty}{dt} + \mathcal{N}_{\mathcal{S}_\infty}(U^\infty(t)) \ni 0, \quad t \in]0,T], \\ U_i^\infty(0) = g_i^0, \quad i \in [n], \end{cases} \tag{$\mathcal{P}^\infty_{\text{nloc}}$}$$

Theorem 4.3.4. *Suppose that* $\text{supp}(\hat{K_n}) \neq \emptyset$ *and* $g^0 \in S_\infty$. *Let* $\check{U}^p$ *be the solution of* (4.3.4)*. If* U^∞ *is the unique solution to* ($\mathcal{P}^\infty_{\text{nloc}}$)*, then*

$$\lim_{p\to\infty}\lim_{\tau\to 0}\sup_{t\in[0,T]}\left|\check{U}^p(t) - U^\infty(t)\right| = 0, \tag{4.3.9}$$

where $\tau = \max_{h\in[N]}\tau_h$ *is is the maximal size of intervals in the partition of* $[0,T]$.

Remark 4.3.5. Note that one cannot interchange the order of limits; the limit as $\tau \to 0$ must be taken before the limit as $p \to \infty$. The reason will be made clear in the proof.

Proof : Using the triangle inequality, we have

$$\left|\check{U}^p(t) - U^\infty(t)\right| \leq \left|\check{U}^p(t) - U^p(t)\right| + \left|U^p(t) - U^\infty(t)\right|.$$

First, proceeding exactly as in the proof of Theorem 3.4.4, and more precisely inequality (3.4.12), we get

$$\left|\check{U}^p(t) - U^p(t)\right| \leq C'\tau \tag{4.3.10}$$

for $C' \geq 0$. Since the constant C' in (4.3.10) depends on p, we first take the limit as $\tau \to 0$, to get

$$\lim_{\tau\to 0}\sup_{t\in[0,T]}\left|\check{U}^p(t) - U^p(t)\right| = 0 \tag{4.3.11}$$

Now, arguing as in [91, Theorem 3.2] (which in turn relies on [30, Theorem 3.1]), we have additionally that

$$\lim_{p\to\infty}\sup_{t\in[0,T]}\left|U^p(t) - U^\infty(t)\right| = 0. \tag{4.3.12}$$

Hence, the combination of (4.3.11) and (4.3.12) yields (4.3.9). □

Remark 4.3.6. Note that we get the same result when dealing with the implicit Euler scheme, following the changes mentioned in Section 3.4.2.

4.3.1.2 Networks on $\mathbb{G}(X_n, K)$

The analysis of the problem ($\mathcal{P}_{\text{nloc}}$) on $\mathbb{G}(X_n, K)$ remains the same modulo the definition of the piecewise constant approximation

$$\check{K}_n^w(x,y) = (\check{K}_n)_{ij} \quad \text{for} \quad (x,y) \in \Omega_{ij}^{(n)},$$

where we recall $\breve{K}_n$ from (2.1.9). The fully discrete counterpart of ($\mathcal{P}_{\text{nloc}}$) on $\mathbb{G}(X_n, K)$ is given by

$$\begin{cases} \dfrac{u_i^h - u_i^{h-1}}{\tau} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}(\breve{K}_n)_{ij}\left|u_i^h - u_i^{h-1}\right|^{p-2}(u_j^{h-1} - u_i^{h-1}), \quad (i,h) \in [n] \times [N], \\ u_i(0) = g_i^0, \quad i \in [n]. \end{cases} \qquad (\breve{\mathcal{P}}_{\text{nloc}}^{w,d})$$

It is worth mentioning that ($\breve{\mathcal{P}}_{\text{nloc}}^{w,d}$) is the time discretized approximation of the problem ($\mathcal{P}_{\text{nloc}}$) using the collocation method. Roughly speaking, it is about the projection of ($\mathcal{P}_{\text{nloc}}$) on X_n (see (2.1.9)) via the interpolation operator $P_n : L^\infty(\Omega) \to X_n$ which to each $u(t_h, .) \in L^\infty(\Omega)$ associates the unique function $f(t_h, .)$ such that for all $i \in [n]$, $u(t_h, \frac{i}{n}) = f(t_h, \frac{i}{n})$. See [93] for more details.

We assume further that the kernel K is almost everywhere continuous on Ω^2. By construction of $\breve{K}_n^w$ (see (2.1.9)),

$$\breve{K}_n^w(x, y) \to K(x, y), \quad \text{as} \quad n \to \infty,$$

at every point of continuity of K, i.e., almost everywhere. Thus, using the Sheffe-Riesz theorem, we have

$$\left\|K - \breve{K}_n^w\right\|_{L^p(\Omega^2)} \to 0 \quad \text{as} \quad n \to \infty.$$

Thereby, the proof of Corollary 4.3.3 applies to the situation at hand. Hence, we have the following result.

Corollary 4.3.7. *Suppose that $p \in]1, +\infty[$, $K : \Omega^2 \to [0, 1]$ is a symmetric measurable function, which is continuous almost everywhere on Ω^2, and $g \in L^\infty(\Omega)$. Let u be the solution of ($\mathcal{P}_{\text{nloc}}$), and $\breve{u}_n$ be the piecewise constant extension as in (4.2.1) using the sequence in ($\breve{\mathcal{P}}_{\text{nloc}}^{w,d}$). Then*

$$\left\|u - \breve{u}_n\right\|_{C(0,T;L^p(\Omega))} \to 0 \quad \textit{as} \quad n \to \infty.$$

Remark 4.3.8. The result of Theorem 4.3.4 remains the same for this graph model taking the kernel $(\breve{K}_n)_{ij}$ instead of $(\hat{K}_n)_{ij}$.

4.4 Networks on random inhomogeneous graphs

4.4.1 Reminders of the random inhomogeneous graph model

In this section, we deal with networks on random inhomogeneous graphs. First, recall the graph model that we perform our analysis with, this model is described in details in Section 2.1.5. The fully discrete counterpart of ($\mathcal{P}_{\text{nloc}}$) on the graph $G_{q_n}(n, K)$ is given by

$$\begin{cases} \dfrac{u_i^h - u_i^{h-1}}{\tau_{h-1}} = \dfrac{1}{n}\displaystyle\sum_{j:(i,j)\in E(G_n)}\left|u_j^{h-1} - u_i^{h-1}\right|^{p-2}(u_j^{h-1} - u_i^{h-1}), \\ u_i^0 = g_i, i \in [n]. \end{cases} \qquad (\mathcal{P}_{\text{nloc}}^{d})$$

Recall the inhomogeneous random graph model introduced in Section 2.1.5.

Remark 4.4.1. 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 [20, 21], i.e. the model constructed replacing (2.1.11) 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 that we will state shortly cover also this graph model. More specifically, the first statements of Theorem 4.4.4 and Theorem 4.4.7 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 intial data g are additionally supposed almost everywhere continuous.

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

$$\Omega^{\mathbf{X}}_{ni} \stackrel{\text{def}}{=}]\mathbf{X}_{(i-1)}, \mathbf{X}_{(i)}], \quad \Omega^{\mathbf{x}}_{ni} \stackrel{\text{def}}{=}]\mathbf{x}_{(i-1)}, \mathbf{x}_{(i)}], \quad \text{and} \quad \Omega^{\mathbf{x}}_{nij} \stackrel{\text{def}}{=}]\mathbf{x}_{(i-1)}, \mathbf{x}_{(i)}] \times]\mathbf{x}_{(j-1)}, \mathbf{x}_{(j)}] \quad i, j \in [n]. \tag{4.4.1}$$

As the realization of the random vector $\mathbf{X}$ is fixed, we define $\hat{K}{}^{\mathbf{x}}_{nij}$ as

$$\hat{K}{}^{\mathbf{x}}_{nij} \stackrel{\text{def}}{=} \min\left(\frac{1}{|\Omega^{\mathbf{x}}_{nij}|}\int_{\Omega^{\mathbf{x}}_{nij}} K(x,y)dxdy, 1/q_n\right). \tag{4.4.2}$$

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

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

Having defined the structure of the network and the discrete counterpart of $(\mathcal{P}_{\text{nloc}})$ on it, we are now in position to state our main error bounds between the discrete dynamics and their continuous ones. First, in Section 4.4.2.1, we assume that X is deterministic. Capitalizing on this result, we will then deal with the totally random model (i.e.; generated by random nodes) in Section 4.4.2.2 by a simple marginalization argument.

4.4.2.1 Networks on graphs generated by deterministic nodes

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

$$\delta(n) = \max_{i\in[n]} \left|\mathbf{x}_{(i)} - \mathbf{x}_{(i-1)}\right|. \tag{4.4.3}$$

Next, we consider the following system of difference equations on $G_{q_n}(n, K)$[1] :

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

where

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

Recall from Section 4.4.1 that λ_{ij} are the i.i.d. random variables such that $q_n\lambda_{ij}$ follows the Bernoulli distribution with parameter $q_n\hat{K}{}^{\mathbf{x}}_{nij}$.

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

Remark 4.4.2. Coming back to Definition 2.1.15, one can easily check that $G_{q_n}(n, K)$ is actually a product probability space[2]

$$\mathbf{\Omega}_n \stackrel{\text{def}}{=} \mathbf{\Omega}^V_n \times \mathbf{\Omega}^E_n \stackrel{\text{def}}{=} \left(\Omega^V_n \stackrel{\text{def}}{=} [0,1]^n, 2^{\Omega^V_n}, \mathbb{P}\right) \times \left(\Omega^E_n \stackrel{\text{def}}{=} \{0,1\}^{n(n+1)/2}, 2^{\Omega^E_n}, \mathbb{P}\right).$$

[1]This is clear by proper normalization by q_n (by dividing and multiplying by q_n). We abuse notation to lighten the system.

[2]To keep notation simple, we allow for loops, in our random graph model. Excluding loops would not lead to any changes in the analysis.

So that, rigorously speaking, if we take a random event $\boldsymbol{\omega}$ from $\boldsymbol{\Omega}_n$, problem ($\mathcal{P}^{d,d}_{\mathrm{nloc}}$) must be written using $\lambda_{ij}(\boldsymbol{\omega})$ instead of λ_{ij}, and likewise for all other random variables. For notational simplicity, we drop $\boldsymbol{\omega}$. But it is important to keep in mind that the evolution equations we write involving random variables must be understood in this sense.

Remark 4.4.3. As the reader may have remarked, the sum in the right-hand side of ($\mathcal{P}^{d,d}_{\mathrm{nloc}}$) 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 chosen 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., [97, Section 1.7], it can be shown that each spacing $\mathbf{X}_{(i)} - \mathbf{X}_{(i-1)}$ concentrates around i/n for $i \in [n]$.

We are now in position to tackle our main goal: comparing the solutions of the discrete and continuous 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 continuous extension of the discrete problem, using the vector $U_h = (u_1^h, u_2^h, \cdots, u_n^h)^\top$ whose components uniquely solve the previous system ($\mathcal{P}^{d,d}_{\mathrm{nloc}}$) (as we have shown in Lemma 3.4.5) to obtain the following piecewise linear interpolation on $\Omega \times [0,T]$

$$\check{u}_n(x,t) = \frac{t_h - t}{\tau_{h-1}} u_i^{h-1} + \frac{t - t_{h-1}}{\tau_{h-1}} u_i^h \quad \text{if} \quad x \in \Omega^{\mathbf{x}}_{ni}, \quad t \in]t_{h-1}, t_h], \tag{4.4.4}$$

and a piecewise approximation

$$\bar{u}_n(x,t) = \sum_{i=1}^{n} \sum_{h=1}^{N} u_i^{h-1} \chi_{]t_{h-1},t_h]}(t) \chi_{\Omega^{\mathbf{x}}_{ni}}(x). \tag{4.4.5}$$

Then, $\check{u}_n$ uniquely solves the following problem

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

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

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

The discrete problem ($\overset{\wedge}{\mathcal{P}}{}^{d}_{\mathrm{nloc}}$) can also be viewed as a discrete p-Laplacian evolution problem over a complete[3] weighted graph on n vertices, where the weight of edge (i,j) is $\hat{K}{}^{\mathbf{x}}_{nij}$.

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

$$\check{v}_n(x,t) = \frac{t_h - t}{\tau_{h-1}} v_i^{h-1} + \frac{t - t_{h-1}}{\tau_{h-1}} v_i^h \quad \text{if} \quad x \in \Omega^{\mathbf{x}}_{ni}, \quad t \in]t_{h-1}, t_h], \tag{4.4.6}$$

[3]Recall that a complete graph is a simple undirected graph in which each pair of vertices is connected by an edge.

and a piecewise-constant approximation

$$\bar{v}_n(x,t) = \sum_{i=1}^{n}\sum_{h=1}^{N} v_i^{h-1}\chi_{]t_{h-1},t_h]}(t)\chi_{\Omega^{\mathbf{x}}_{ni}}(x). \tag{4.4.7}$$

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

$$\hat{K}_n(x,y) = \sum_{(i,j)\in[n]^2} \hat{K}^{\mathbf{x}}_{nij}\chi_{\Omega^{\mathbf{x}}_{nij}}(x,y). \tag{4.4.8}$$

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

$$\begin{cases} \frac{\partial}{\partial t}\check{v}_n(x,t) = -\boldsymbol{\Delta}_p^{\hat{K}_n}(\bar{v}_n(x,t)), & x\in\Omega, t>0,\\ \check{v}_n(x,0) = g_n(x), & x\in\Omega, \end{cases} \tag{$\overset{\wedge}{\mathcal{P}}_{\mathrm{nloc}}$}$$

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 section is the following theorem.

Theorem 4.4.4. *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 ($\mathcal{P}_{\mathrm{nloc}}$) and ($\mathcal{P}^{d,d}_{\mathrm{nloc}}$), respectively. Let $\check{u}_n$ be the continuous extension of U^h given in (4.4.4). 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 $\beta>0$

$$\left\|u-\check{u}_n\right\|_{C(0,T;L^p(\Omega))} \le C_1 T\exp\left(O(T)\right)\left(\left(\beta\frac{\log(n)}{n}+\frac{\max\left(q_n^{-(p-1)},q_n^{-p/2}\right)}{n^{p/2}}\right)^{1/p} + \left\|K-\hat{K}_n\right\|_{L^p(\Omega^2)} + \left\|g-g_n\right\|_{L^p(\Omega)} + \tau\right), \tag{4.4.9}$$

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 $q_n\left\|K\right\|_{L^\infty(\Omega^2)}\le 1$. Then, for $T>0$, there exist positive constants C_1 and C_2, independent of n and T, such that for any $\beta>0$

$$\left\|u-\check{u}_n\right\|_{C(0,T;L^p(\Omega))} \le C_1 T\exp\left(O(T)\right)\left(\left(\beta\frac{\log(n)}{n}+\frac{\max\left(q_n^{-(p-1)},q_n^{-p/2}\right)}{n^{p/2}}\right)^{1/p} + \delta(n)^{\min(s,s')\min(1,q/p)} + \tau\right), \tag{4.4.10}$$

with probability at least $1-n^{-C_2 q_n^{2p-1}\beta}$, where $\delta(n)$ is the spacing parameter defined in (4.4.3).

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

Remark 4.4.5.

(i) The constant in (4.4.9) depends on p and the data via $\left\|g\right\|_{L^\infty(\Omega)}$ and $\left\|K\right\|_{L^\infty(\Omega^2)}$. For the bound (4.4.10), it also depends on (q,s,s').

(ii) By Lemma 2.2.16, it is clear that the first term in the bounds (4.4.9)-(4.4.10) 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

$$n^{-1/2}\max\left(q_n^{-(1-1/p)},q_n^{-1/2}\right) = \begin{cases} (q_n n)^{-1/2} & \text{if } p\in]1,2],\\ q_n^{1/p}(q_n^2 n)^{-1/2} & \text{if } p>2. \end{cases} \tag{4.4.11}$$

Thus, if $\inf_{n\geq 1} q_n > 0$, as is the case when the graph is dense (see discussion after Proposition 2.1.17), then the term (4.4.11) 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 2.1.17), then (4.4.11) 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 2.1.17 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 4.4.4 hold almost surely. The same reasoning carries over for the bounds of Theorem 4.4.7.

(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 4.4.8(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 [78], one can show that the graphon corresponding to the nearest neighbour graphs, which are very popular in practice (e.g. in image processing [49, 46]), 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 4.4.4, we first show the following key lemma.

Lemma 4.4.6. *Under the assumptions of Theorem 4.4.4, 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 4.4.6: 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 [66, Lemma 5.1], we have that both $(\mathcal{P}^{d,d}_{\mathrm{nloc}})$ and $(\hat{\mathcal{P}}^{d}_{\mathrm{nloc}})$ are well posed. In turn U^h and V^h are bounded and V^h uniquely solves $(\hat{\mathcal{P}}^{d}_{\mathrm{nloc}})$, and similarly for $\check{u}_n$ and $\check{v}_n$ as solutions to $(\mathcal{P}^{\Lambda_n}_{\mathrm{nloc}})$ and $(\hat{\mathcal{P}}_{\mathrm{nloc}})$. 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 $(\mathcal{P}^{\Lambda_n}_{\mathrm{nloc}})$ from those of $(\hat{\mathcal{P}}_{\mathrm{nloc}})$, evaluated at the cell $\Omega^{\mathbf{x}}_{ni}$, we obtain

$$\begin{aligned}\frac{d}{dt}\check{\xi}_{ni}(t) &= \frac{1}{n}\sum_{j=1}^{n}\Big(\lambda_{ij}\Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)) - \hat{K}{}^{\mathbf{x}}_{nij}\Psi(\bar{\mathbf{v}}_{nj}(t)-\bar{\mathbf{v}}_{ni}(t))\Big)\\ &= Z_{ni}(t) + \frac{1}{n}\sum_{j=1}^{n}\hat{K}{}^{\mathbf{x}}_{nij}\big(\Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)) - \Psi(\bar{\mathbf{v}}_{nj}(t)-\bar{\mathbf{v}}_{ni}(t))\big),\end{aligned} \tag{4.4.12}$$

where

$$Z_{ni}(t) = \frac{1}{n}\sum_{j=1}^{n}(\lambda_{ij} - \hat{K}{}^{\mathbf{x}}_{nij})\alpha_{ij}(t) \text{ and } \alpha_{ij}(t) = \Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)), \forall (i,j)\in[n]^2, t\in[0,T]. \tag{4.4.13}$$

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 (4.4.12) by $\frac{1}{n}\Psi(\check{\xi}_{ni}(t))$ and sum over i to obtain

$$\frac{1}{p}\frac{d}{dt}\|\check{\xi}_n(t)\|^p_{p,n} = \frac{1}{n}\sum_{i=1}^{n}Z_{ni}(t)\Psi(\check{\xi}_{ni}(t)) + \frac{1}{n^2}\sum_{i,j=1}^{n}\hat{K}{}^{\mathbf{x}}_{nij}\big(\Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)) - \Psi(\bar{\mathbf{v}}_{nj}(t)-\bar{\mathbf{v}}_{ni}(t))\big)\Psi(\check{\xi}_{ni}(t)). \tag{4.4.14}$$

We estimate the first term on the right-hand side of (4.4.14) using the Hölder inequality, to get

$$\frac{1}{n}\left|\sum_{i=1}^{n}Z_{ni}(t)\Psi(\check{\xi}_{ni}(t))\right| \le \frac{1}{n}\left(\sum_{i=1}^{n}\big|Z_{ni}(t)\big|^p\right)^{\frac{1}{p}} \times \left(\sum_{i=1}^{n}\big|\check{\xi}_{ni}(t)\big|^p\right)^{\frac{p-1}{p}} \le \big\|Z_n(t)\big\|_{p,n}\big\|\check{\xi}_n(t)\big\|^{p-1}_{p,n}. \tag{4.4.15}$$

Now, using the fact that $\hat{K}{}^{\mathbf{x}}_{nij} \le \big\|K\big\|_{L^\infty(\Omega^2)}$ (see (2.1.12)), $\forall (i,j)\in[n]^2$, and applying [66, Corollary B.1] to the function Ψ between $a = \bar{\mathbf{v}}_{nj}(t) - \bar{\mathbf{v}}_{ni}(t)$ and $b = \bar{\mathbf{u}}_{nj}(t) - \bar{\mathbf{u}}_{ni}(t)$ (without loss of generality, we suppose that $b > a$), we get

$$\begin{aligned}&\left|\frac{1}{n^2}\sum_{i,j=1}^{n}\hat{K}{}^{\mathbf{x}}_{nij}\big(\Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)) - \Psi(\bar{\mathbf{v}}_{nj}(t)-\bar{\mathbf{v}}_{ni}(t)\big)\Psi(\xi_{ni}(t))\right|\\ &\le \frac{(p-1)\big\|K\big\|_{L^\infty(\Omega^2)}}{n^2}\sum_{i,j=1}^{n}\big|\bar{\xi}_{nj}-\bar{\xi}_{ni}\big|\big|\eta_n(t)\big|^{p-2}\big|\check{\xi}_{ni}\big|^{p-1},\end{aligned} \tag{4.4.16}$$

where $\eta_n(t)$ is an intermediate value between a and b. Using that fact that $g \in L^\infty(\Omega)$ and the construction of $\bar{\mathbf{u}}_n(\cdot)$, we deduce from [66, Theorem 3.1(ii)] that for $t \in [0,T]$

$$\big|\eta_n(t)\big|^{p-2} \le \big|\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t)\big|^{p-2} \le \big(2\big\|u(\cdot,t)\big\|_{L^\infty(\Omega)}\big)^{p-2} \le \big(2\big\|g\big\|_{L^\infty(\Omega)}\big)^{p-2}. \tag{4.4.17}$$

Let $C_2 = \big(2\big\|g\big\|_{L^\infty(\Omega)}\big)^{p-2}\big\|K\big\|_{L^\infty(\Omega^2)}$. Inserting (4.4.17) into (4.4.16), and then using the Hölder and triangle inequalities, it follows that

$$
\begin{aligned}
&\left|\frac{1}{n^2}\sum_{i,j=1}^{n}\hat{K}{}^{\mathbf{x}}_{nij}\big(\Psi(\bar{\mathbf{u}}_{nj}(t)-\bar{\mathbf{u}}_{ni}(t))-\Psi(\bar{\mathbf{v}}_{nj}(t)-\bar{\mathbf{v}}_{ni}(t)\big)\Psi(\check{\xi}_{ni}(t))\right|\\
&\le C_2\frac{p-1}{n^2}\sum_{i,j=1}^{n}\left|\bar{\xi}_{nj}(t)-\bar{\xi}_{ni}(t)\right|\left|\check{\xi}_{ni}\right|^{p-1}\\
&\le C_2\frac{p-1}{n^2}\left(\left(\sum_{i,j=1}^{n}\left|\bar{\xi}_{nj}(t)-\bar{\xi}_{ni}(t)\right|^{p}\right)^{\frac{1}{p}}\left(\sum_{i,j}\left|\check{\xi}_{ni}(t)\right|^{p}\right)^{\frac{p-1}{p}}\right)\\
&\le C_2\frac{p-1}{n^2}\left(\left(\sum_{i,j=1}^{n}\left|\bar{\xi}_{nj}(t)\right|^{p}\right)^{\frac{1}{p}}+\left(\sum_{i,j=1}^{n}\left|\bar{\xi}_{ni}(t)\right|^{p}\right)^{\frac{1}{p}}\right)\left(n^{\frac{2(p-1)}{p}}\left(\frac{1}{n}\sum_{i=1}^{n}\left|\check{\xi}_{ni}(t)\right|^{p}\right)^{\frac{p-1}{p}}\right)\\
&\le C_2\frac{p-1}{n^2}\left(2n^{\frac{2}{p}}\left\|\bar{\xi}_{n}(t)\right\|_{p,n}\right)\left(n^{\frac{2(p-1)}{p}}\left\|\check{\xi}_{n}(t)\right\|_{p,n}^{p-1}\right)\\
&\le 2C_2(p-1)\left\|\bar{\xi}_{n}(t)\right\|_{p,n}\left\|\check{\xi}_{n}(t)\right\|_{p,n}^{p-1}.
\end{aligned}
\tag{4.4.18}
$$

Using the triangle inequality combined with [66, Lemma 5.2], we have

$$
\begin{aligned}
\left\|\bar{\xi}_{n}(t)\right\|_{p,n}&=\left\|\bar{\mathbf{v}}_{n}(t)-\bar{\mathbf{u}}_{n}(t)\right\|_{p,n}\\
&\le\left\|\bar{\mathbf{v}}_{n}(t)-\check{\mathbf{v}}_{n}(t)\right\|_{p,n}+\left\|\check{\mathbf{v}}_{n}(t)-\check{\mathbf{u}}_{n}(t)\right\|_{p,n}+\left\|\check{\mathbf{u}}_{n}(t)-\bar{\mathbf{u}}_{n}(t)\right\|_{p,n}\\
&\le C\tau+\left\|\check{\xi}_{n}(t)\right\|_{p,n}+C'\tau\\
&\le C''\tau+\left\|\check{\xi}_{n}(t)\right\|_{p,n}.
\end{aligned}
\tag{4.4.19}
$$

Putting together (4.4.14), (4.4.15), (4.4.18) and (4.4.19), we have

$$
\begin{aligned}
\frac{d}{dt}\left\|\check{\xi}_{n}(t)\right\|_{p,n}^{p}&\le\left\|Z_n(t)\right\|_{p,n}\left\|\check{\xi}_{n}(t)\right\|_{p,n}^{p-1}+2C_2(p-1)\left(C''\tau+\left\|\check{\xi}_{n}(t)\right\|_{p,n}\right)\left\|\check{\xi}_{n}(t)\right\|_{p,n}^{p-1}\\
&\le\left(2C_3(p-1)\tau+\left\|Z_n(t)\right\|_{p,n}\right)\left\|\check{\xi}_{n}(t)\right\|_{p,n}^{p-1}+2C_2(p-1)\left\|\check{\xi}_{n}(t)\right\|_{p,n}^{p}.
\end{aligned}
\tag{4.4.20}
$$

Then, from (4.4.20) via the Gronwall's inequality in its differential form (see, e.g., [51, Appendix B]), we obtain

$$
\left\|\check{u}_n-\check{v}_n\right\|_{C(0,T;L^p(\Omega))}=\sup_{t\in[0,T]}\left\|\check{\xi}_{n}(t)\right\|_{p,n}\le\left(2C_3T\tau+\int_0^T\left\|Z_n(t)\right\|_{p,n}dt\right)\exp\left(2C_2T\right). \tag{4.4.21}
$$

It remains to bound $\int_0^T\left\|Z_n(t)\right\|_{p,n}dt$. For this purpose, we use Lemma 4.4.10 (see Section 4.4.4)[4]. Thus, plugging the bound of Lemma 4.4.10(i) into inequality (4.4.21), we get the desired conclusion. □

We are now ready to prove our main result.

Proof of Theorem 4.4.4:

(i) Using the triangle inequality, we have

$$
\left\|u-\check{u}_n\right\|_{C(0,T;L^p(\Omega))}\le\left\|u-\check{v}_n\right\|_{C(0,T;L^p(\Omega))}+\left\|\check{v}_n-\check{u}_n\right\|_{C(0,T;L^p(\Omega))}. \tag{4.4.22}
$$

Since by construction $\hat{K}_n$ is a bounded mapping, we bound the first term on the right-hand side of (4.4.22) using [66, Theorem 5.1][5] to get

$$
\left\|u-\check{v}_n\right\|_{C(0,T;L^p(\Omega))}=O\left(T\exp(O(T))\big(\left\|K-\hat{K}_n\right\|_{L^p(\Omega^2)}+\left\|g-g_n\right\|_{L^p(\Omega)}+\tau\big)\right), \tag{4.4.23}
$$

[4]This inequality is sharp as can be seen for instance from assertion (ii) of Lemma 4.4.10, at least for $p\ge 2$.
[5]Here, we have made the constant explicit in T compared to the statement in Theorem 3.4.4.

Claim (4.4.9) then follows by plugging (4.4.23) and Lemma 4.4.6 into (4.4.22).

(ii) Our assumption on q_n together with (4.4.2) and (4.4.8) entail that

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

Since $g \in \mathrm{Lip}(s, L^q(\Omega))$ and $K \in \mathrm{Lip}(s', L^q(\Omega^2))$, we can invoke Lemma 2.3.3 to get

$$\left\|K - \hat{K}_n\right\|_{L^p(\Omega^2)} \leq C(p,q,s')\delta(n)^{s'\min(1,q/p)} \text{ and } \left\|g - g_n\right\|_{L^p(\Omega)} \leq C(p,q,s)\delta(n)^{s\min(1,q/p)}. \tag{4.4.24}$$

Inserting the bound (4.4.24) into (4.4.9), and using the fact that $\delta(n) < 1$, yields (4.4.10).

□

4.4.2.2 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 $G_{q_n}(n,K)$[6] :

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

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

$$\check{u}_n(x,t) = \frac{t_h - t}{\tau_{h-1}}u_i^{h-1} + \frac{t - t_{h-1}}{\tau_{h-1}}u_i^h \quad \text{if} \quad x \in \Omega^{\mathbf{X}}_{ni}, \quad t \in]t_{h-1}, t_h], \tag{4.4.25}$$

and a piecewise approximation

$$\bar{u}_n(x,t) = \sum_{i=1}^n\sum_{h=1}^N u_i^{h-1}\chi_{]t_{h-1},t_h]}(t)\chi_{\Omega^{\mathbf{X}}_{ni}}(x). \tag{4.4.26}$$

Then, we have

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

where

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

and the random variable Γ_n is such that

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

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

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

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.

[6] Recall again from Remark 4.4.2, that rigorously speaking, each random variable involved in the problems and equations of this section should be understood as a function of an event $\boldsymbol{\omega}$ from $\boldsymbol{\Omega}_n$. This dependence is dropped only to lighten notation.

Theorem 4.4.7. *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* ($\mathcal{P}_{\text{nloc}}$) *and* ($\mathcal{P}^{r,d}_{\text{nloc}}$), *respectively. Let* $\check{u}_n$ *be the continuous extension of* U_h *given in* (4.4.25). *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* $\beta > 0$

$$\|u-\check{u}_n\|_{C(0,T;L^p(\Omega))} \leq C_1 T \exp\left(O(T)\right)\left(\left(\beta\frac{\log(n)}{n} + \frac{\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)}{n^{p/2}}\right)^{1/p} + \|K-\hat{K}_n\|_{L^p(\Omega^2)} + \|g-g_n\|_{L^p(\Omega)} + \tau\right), \tag{4.4.27}$$

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

(ii) Suppose furthermore that $g \in \text{Lip}(s, L^q(\Omega))$ *and* $K \in \text{Lip}(s', L^q(\Omega^2))$, $s, s' \in]0,1]$, *and* $q_n\|K\|_{L^\infty(\Omega^2)} \leq 1$. *Let* $\theta \stackrel{\text{def}}{=} \min(s,s')\min(1,q/p)$. *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* $t \in]0,e[$

$$\|u-\check{u}_n\|_{C(0,T;L^p(\Omega))} \leq C_1 T \exp\left(O(T)\right)\left(\left(\beta\frac{\log(n)}{n} + \frac{\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)}{n^{p/2}}\right)^{1/p} + \left(\frac{t\log(n)}{n}\right)^{\theta} + \tau\right), \tag{4.4.28}$$

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

The dependence of the constant C in the parameters is similar to Remark 4.4.5(ii).

Remark 4.4.8.

(i) The dependence of the constant C in the parameters is similar to Remark 4.4.5(i).

(ii) As observed in Remark 4.4.5(v), one can take $T = c_2\log(n)$, in which case $T\exp(c_1 T) = c_2 n^{c_1c_2}\log(n)$, with $c_1, c_2 > 0$. Consequently, if one sets $q_n = \log(n)^{-\delta/(2p-1)}$, for $\delta \in]0,1[$ (see Remark 4.4.5(iii)), then the bound in (4.4.28) 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 $c_1c_2 < \min(1/p, 1/2, \theta)$.

As a preparatory step to prove Theorem 4.4.7, 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 4.4.9. *Consider the sequence of random spacings* $(\mathbf{X}_{(1)}, \mathbf{X}_{(2)}-\mathbf{X}_{(1)}, \cdots, 1-\mathbf{X}_{(n)})$, *where we recall* $\{\mathbf{X}_{(i)}\}_{i=1}^n$ *are the order statistics of* $\mathbf{X}$. *Let* $t \in]0,e[$. *Then, for any* $i \in [n]$

$$\delta_i \stackrel{\text{def}}{=} \mathbf{X}_{(i)} - \mathbf{X}_{(i-1)} \leq t\frac{\log(n)}{n}, \tag{4.4.29}$$

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

Proof of Lemma 4.4.9: Since $\mathbf{X}_i$ are i.i.d. uniform random variables on Ω, we have, by virtue of [97, 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 [97, 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 $\varepsilon \in [0,1[$

$$\begin{aligned}\mathbb{P}(\delta_i \geq \varepsilon) = \mathbb{P}(Z_i \geq \varepsilon S_{n+1}) &= \mathbb{P}((1-\varepsilon)Z_i \geq \varepsilon(S_{n+1}-Z_i)) \\ &= \mathbb{P}\left(Z_{n+1} \geq \frac{\varepsilon}{1-\varepsilon}S_n\right) \\ &= \int_0^{+\infty} \mathbb{P}\left(Z_{n+1} \geq \frac{\varepsilon}{1-\varepsilon}s\right) f_{S_n}(s)ds \\ &= \int_0^{+\infty} e^{-\frac{\varepsilon}{1-\varepsilon}s}e^{-s}\frac{s^{n-1}}{(n-1)!}ds \\ &= (1-\varepsilon)^n.\end{aligned} \tag{4.4.30}$$

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 4.4.7: 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 4.4.4, which is independent of $\mathbf{x}$, and finally integrate with respect to the uniform density on Ω^n.

(i) We have

$$\begin{aligned}\mathbb{P}\left(\|u-\check{u}_n\|_{C(0,T;L^p(\Omega))} \geq \varepsilon'\right) &= \frac{1}{|\Omega|^n}\int_{\Omega^n} \mathbb{P}\left(\|u-\check{u}_n\|_{C(0,T;L^p(\Omega))} \geq \varepsilon' | \mathbf{X}=\mathbf{x}\right) d\mathbf{x} \\ &\leq \frac{1}{|\Omega|^n}\int_{\Omega^n} n^{-C_2 q_n^{2p-1}\beta} d\mathbf{x} \\ &= n^{-C_2 q_n^{2p-1}\beta},\end{aligned} \tag{4.4.31}$$

with

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

Thus, (4.4.27) follows from the fact that the obtained bound in (4.4.9) is independent of the random choice of $\mathbf{x}$.

(ii) In view of (4.4.24), we can argue that

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

Taking $\kappa = (C(p,q,s)+C(p,q,s'))\left(t\frac{\log(n)}{n}\right)^\theta$, for $t \in]0,e[$, and applying Lemma 4.4.9, 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

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

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

□

4.4.3 Asymptotic regimes

A close inspection of the error bound in (4.4.28) (Theorem 4.4.7) reveals three contributions:

- Spatial discretization: the first contribution is materialized in the first term which scales as (see Remark 4.4.5(i))
$$O\left(\left(\frac{\log(n)}{n}\right)^{1/p}+\frac{\max\left(q_n^{-(1-1/p)},q_n^{-1/2}\right)}{n^{1/2}}\right).$$
This term represents the spatial discretization error when approximating the continuous evolution equation ($\mathcal{P}_{\mathrm{nloc}}$) on the random inhomogeneous graph model $G_{q_n}(n,K)$ generated according to Definition 2.1.15 with the graphon K.
- Data approximation: the second term is $O\left(\left(\frac{\log(n)}{n}\right)^{\theta}\right)$ which captures the error of discretizting the initial data g and the graphon K. 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 (4.4.28) independent of p and we choose to quantify the error in the classical $L^2(\Omega)$ norm. Consequently, thanks to Lemma 2.2.16 and Lemma 2.2.19, as well as boundedness of the solutions, it is not difficult to see that

$$\|u-\check{u}_n\|_{C(0,T;L^2(\Omega))} = \begin{cases} O\left(\left(\beta\frac{\log(n)}{n}\right)^{1/p}+\frac{\max\left(q_n^{-(1-1/p)},q_n^{-1/2}\right)}{n^{1/2}}+\left(\frac{t\log(n)}{n}\right)^{\theta}+\tau\right), & p\in[2,+\infty[\\ O\left(\left(\beta\frac{\log(n)}{n}\right)^{1/2}+\frac{\max\left(q_n^{-(p-1)/2},q_n^{-p/4}\right)}{n^{p/4}}+\left(\frac{t\log(n)}{n}\right)^{p\theta/2}+\tau^{p/2}\right) & p\in]1,2], \end{cases} \tag{4.4.32}$$

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 4.4.5(iii)) and Section 2.1.5). Thus, we have

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

In turn, the second term in (4.4.32) is bounded by

$$\left(\frac{\log(n)}{n}\right)^{\min(p/4,1/2)}, \forall p \in]1,+\infty[. \tag{4.4.33}$$

Without loss of generality[7], 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, (4.4.32) reads

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

The term depending on n then exhibits four different regimes as a function of p, s and q (see Figure 4.1). Indeed, it is straightforward to see that it scales as

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

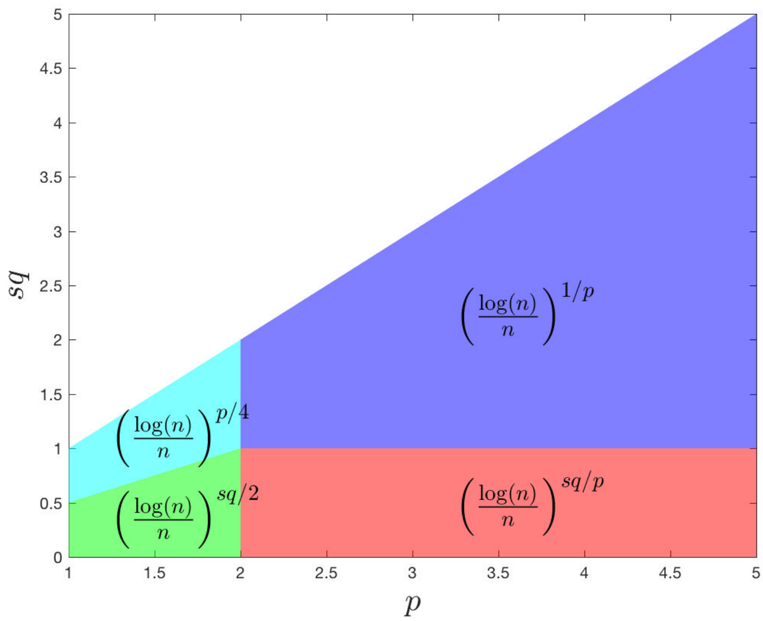


Figure 4.1: Different regimes according to the values of p and s, and q.

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.

4.4.4 A key deviation result

The following lemma establishes a key deviation inequality for $\sup_{t\in[0,T]} \|Z_n(t)\|_{p,n}$ where $Z_n(\cdot)$ is a random process defined as

$$Z_{ni}(t) = \frac{1}{n}\alpha_{ij}(t)\sum_{j=1}^{n}(\lambda_{ij} - \gamma_{ij}), \tag{4.4.34}$$

[7]This setting is true for many graphons, see, e.g., Remark 4.4.5(vi).

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 $q_n\gamma_{ij}$. It is obvious that this process covers that in (4.4.13) as a special case.

Lemma 4.4.10. *Let $Z_n(\cdot)$ be the random process defined in (4.4.34). 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,

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

Proof of Lemma 4.4.10:

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

$$\begin{aligned}
\mathbb{E}(Y_{ni}) &= \frac{1}{n^p}\int_0^T \mathbb{E}\left(\left|\sum_{j=1}^n U_{nij}(t)\right|^p\right) dt \\
&\leq \frac{C_1 T}{n^p}\sup_{t\in[0,T]} \max\left(\sum_{j=1}^n \mathbb{E}(|U_{nij}(t)|^p), \left(\sum_{j=1}^n \mathbb{E}(U_{nij}(t)^2)\right)^{p/2}\right).
\end{aligned} \tag{4.4.35}$$

We have

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

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,

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

Therefore

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(Y_{ni}\right) \leq C_1C_2T\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)n^{-p/2}. \tag{4.4.36}$$

- $p \in [1, 2[$. 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

$$\mathbb{E}\left(\left(\sum_{j=1}^{n}U_{nij}(t)\right)^2\right) = \mathrm{Var}\left(\sum_{j=1}^{n}U_{nij}(t)\right) = \sum_{j=1}^{n}\mathbb{E}\left(U_{nij}(t)^2\right). \tag{4.4.37}$$

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

$$\begin{aligned}\mathbb{E}\left(Y_{ni}\right) \leq \frac{T}{n^p}\sup_{t\in[0,T]}\mathbb{E}\left(\left|\sum_{j=1}^{n}U_{nij}(t)\right|^p\right) &\leq \frac{T}{n^p}\sup_{t\in[0,T]}\left(\mathbb{E}\left(\left(\sum_{j=1}^{n}U_{nij}(t)\right)^2\right)\right)^{p/2}\\ &= \frac{T}{n^p}\sup_{t\in[0,T]}\left(\sum_{j=1}^{n}\mathbb{E}\left(U_{nij}(t)^2\right)\right)^{p/2}\\ &= \frac{T}{n^p}\sup_{t\in[0,T]}\left(\sum_{j=1}^{n}\frac{\alpha_{ij}(t)^2}{q_n}\gamma_{ij}(1-q_n\gamma_{ij})\right)^{p/2}\\ &\leq \frac{C_2T}{q_n^{p/2}}n^{-p/2} \leq C_2T\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)n^{-p/2}.\end{aligned} \tag{4.4.38}$$

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

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(Y_{ni}\right) \leq C_3T\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)n^{-p/2}, \tag{4.4.39}$$

where $C_3 = C_2\max(1, C_1)$.
Hence, setting $W_{ni} = Y_{ni} - \mathbb{E}\left(Y_{ni}\right)$ and $\kappa = T^{1-p}\varepsilon^p - C_3T\max\left(q_n^{-(p-1)}, q_n^{-p/2}\right)n^{-p/2}$, we have

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

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

▷ $\sup_{i\in[n]}\left|W_{ni}\right| \leq 2\sup_{i\in[n]}\left|Y_{ni}\right| \leq C_4T$, since α_{ij} and $q_n\gamma_{ij}$ are both uniformly bounded.

▷ $\sum_{i=1}^{n}\mathbb{E}\left(W_{ni}^2\right) = \sum_{i=1}^{n}\mathrm{Var}\left(Y_{ni}\right) \leq \sum_{i=1}^{n}\mathbb{E}\left(Y_{ni}^2\right)$. Using the Jensen inequality with the function $x \mapsto x^2$, and replacing the exponent "p" in inequality (4.4.35), by "$2p$" which is greater than 2, we obtain

$$\sum_{i=1}^{n}\mathbb{E}\left(W_{ni}^2\right) \leq \sum_{i=1}^{n}\mathbb{E}\left(Y_{ni}^2\right) \leq C_5T^2\max\left(q_n^{-(2p-1)}, q_n^{-p}\right)\frac{1}{n^{p-1}}.$$

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

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

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

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

Thus

$$\mathbb{P}\left(\int_0^T \left\|Z_n(t)\right\|_{p,n} dt \geq \varepsilon\right) \leq n^{-C\min\left(q_n^{2p-1}, q_n^p\right)\beta}. \tag{4.4.40}$$

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

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

Thus, for $p \in [2, +\infty[$, applying the Jensen inequality and using (4.4.37), we have

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

Combining this lower-bound with the upper-bounded (4.4.39), we get the claimed equivalence.

□

Chapter 5

The Normalized p-Laplacian Evolution Problem on Graphs

Main contributions of this chapter

- We deal with the discrete (in space) normalized p-Laplacian evolution problem on graphs. We establish the well-posedness of this problem.
- We illustrate the use of this problem on filtering images and 3D point clouds.

Contents

By convention, the Hilbert space of vectors associating a real value to each vertex $i \in V$ of a weighted graph $G_n = (V, E, K_n)$ is denoted by $\mathcal{H}(V)$. Each $u_n : V \to \mathbb{R}$ in $\mathcal{H}(V)$ associates a real value $u_n(i)$ to each vertex $i \in V$. One can see u_n as a column vector $U_n = [u_n(1), \cdots, u_n(n)]^\top$ of $\mathbb{R}^n$, where $n = |V|$ and in which each component corresponds to a vertex $i \in V$. The space $\mathcal{H}(V)$ is endowed with the inner product defined for two vectors $u_n, v_n \in \mathcal{H}(V)$ by

$$\langle u_n, v_n \rangle = \sum_{i \in V} u_n(i) v_n(i).$$

5.1 Introduction

The normalized p-Laplacian recently introduced in its infinite form in connection with a stochastic game called the Tug-of-War game [115] and Tug-of-War with noise [90] is a normalized version of the p-Laplacian. The interest of this class of operators derives from the fact that it contains particular cases for the p-Laplacian depending on the value of p. One can find the mean curvature operator for $p = 1$, a multiple of the ordinary Laplace operator for $p = 2$. In the homogeneous case, the game p-Laplacian equation coincides with the variational p-Laplacian equation for which many approximations have been proposed. Some of these approximations are based on finite elements [13]. Some other approximations using finite difference were also proposed for the normalized p-Laplacian for $p = 1$, $p = \infty$ and $p \geq 2$ [87]. One can also cite the approximations of the normalized p-Laplacian for $1 \leq p \leq \infty$ by statistical operators [98]. Nevertheless, all of these proposed methods deal with regular domains. However, potential existing and future applications require to tackle this problem in general domains or graphs with arbitrary topology.

Motivated by the desire to extend this operator on all kinds of discrete domains, the authors of [1] have proposed an adaptation and generalization of the normalized p-Laplacian on weighted graphs using the frame of EdPs [105, 47]. This adaptation can be considered as a new class of p-Laplacian on graphs as an interpolation between the nonlocal 1-Laplacian, the nonlocal infinity Laplacian and the nonlocal 2-Laplacian on graphs.

In this chapter, motivated by this recent work dealing with the normalized p-Laplacian on graphs, we study the Cauchy problem associated to this operator on graphs and show the existence and uniqueness of a solution to this diffusion problem. First, we begin by recalling the definition of the normalized p-Laplacian as given in [90]. Then, we recall the main definitions related to the normalized p-Laplacian on graphs using the 'so-called' statistical operators proposed in [1]. Finally, we show some applications in image and data processing such as filtering to illustrate the use of this class of operators on graphs.

5.2 The normalized p-Laplacian

We recall from Section 1.1.1 that the local p-Laplacian operator of a function $u : \Omega \subset \mathbb{R}^N \to \mathbb{R}$ is given for $1 \leq p < \infty$ as

$$\boldsymbol{\Delta}_p u = \mathrm{div}(\left|\nabla u\right|^{p-2}\nabla u).$$

In the case $p = \infty$, it is traditionally given by $\boldsymbol{\Delta}_\infty u = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_i}\frac{\partial^2 u}{\partial x_i \partial x_j}$.

The *game* or *normalized* p-Laplacian recently introduced in [90] is written as for $1 \leq p < \infty$

$$\boldsymbol{\Delta}_p^{\mathrm{Nor}} u = \frac{1}{p}\left|\nabla u\right|^{2-p}\mathrm{div}(\left|\nabla u\right|^{p-2}\nabla u). \tag{5.2.1}$$

When $p = \infty$

$$\boldsymbol{\Delta}_\infty^{\mathrm{Nor}} u = \left|\nabla u\right|^{-2}\boldsymbol{\Delta}_\infty u.$$

$\boldsymbol{\Delta}_p^{\mathrm{Nor}}$ is called normalized since it is homogeneous of degree 1, i.e. $\boldsymbol{\Delta}_p^{\mathrm{Nor}}(s\, u) = s\,.\,\boldsymbol{\Delta}_p^{\mathrm{Nor}} u$ for $s \in \mathbb{R}$ in contrast to the p-Laplacian which is homogeneous of degree $p-1$ (see Proposition 3.2.1). Thus, the parabolic problems involving the normalized p-Laplacian are scale invariant. This is a useful property in the context of image processing. If u is a smooth function, equation (5.2.1) can be rewritten as (see [107]):

$$\begin{aligned} \boldsymbol{\Delta}_p^{\mathrm{Nor}} u &= \frac{(p-2)}{p}\boldsymbol{\Delta}_\infty^{\mathrm{Nor}} u + \frac{1}{p}\Delta u \\ &= \frac{(p-2)}{p}\boldsymbol{\Delta}_\infty^{\mathrm{Nor}} u + \frac{2}{p}\boldsymbol{\Delta}_2^{\mathrm{Nor}} u \\ &= \alpha(p)\boldsymbol{\Delta}_\infty^{\mathrm{Nor}} u + \beta(p)\boldsymbol{\Delta}_2^{\mathrm{Nor}} u \end{aligned} \tag{5.2.2}$$

with $\alpha(p) = (p-2)/p$ and $\beta(p) = 2/p$. In plain words, $\boldsymbol{\Delta}_p^{\mathrm{Nor}}$ is a convex combination of $\boldsymbol{\Delta}_\infty^{\mathrm{Nor}}$ and $\boldsymbol{\Delta}_2^{\mathrm{Nor}}$ for $p \geq 2$.

The game p-Laplacian for $p = 1$ can be written as:

$$\boldsymbol{\Delta}^{\mathrm{Nor}} u = \mathrm{div}\left(\frac{\nabla(u)}{|\nabla(u)|}\right)\left|\nabla(u)\right|. \tag{5.2.3}$$

As $\boldsymbol{\Delta}_1^{\mathrm{Nor}} u = \boldsymbol{\Delta} u - \boldsymbol{\Delta}_\infty^{\mathrm{Nor}} u$, (5.2.2) can be rewritten as:

$$\boldsymbol{\Delta}_p^{\mathrm{Nor}} u = \alpha'(p)_2\boldsymbol{\Delta}_2^{\mathrm{Nor}} u + \beta'(p)\boldsymbol{\Delta}_1^{\mathrm{Nor}} u, \tag{5.2.4}$$

with $\alpha'(p) = \frac{2(p-1)}{p}$ and $\beta'(p) = \frac{2-p}{p}$, which is again a convex sum for $p \in [1, 2]$.

In view of (5.2.2) and (5.2.4), the game p-Laplacian for $1 \leq p \leq \infty$ can be rewritten in the form of a convex sum as:

$$\boldsymbol{\Delta}_p^{\mathrm{Nor}} u = \begin{cases} \frac{2}{p}\boldsymbol{\Delta}_2^{\mathrm{Nor}} u + \frac{p-2}{p}\boldsymbol{\Delta}_\infty^{\mathrm{Nor}} u & \text{for } 2 \leq p \leq \infty; \\ \frac{2(p-1)}{p}\boldsymbol{\Delta}_2^{\mathrm{Nor}} u + \frac{2-p}{p}\boldsymbol{\Delta}_1^{\mathrm{Nor}} u & \text{for } 1 \leq p \leq 2. \end{cases} \tag{5.2.5}$$

5.3 The normalized p-Laplacian on graphs

In [1] and earlier [105], the authors have proposed an extension of the game (normalized) p-Laplacian on weighted graphs. For this they introduced *statistical operators* needed to define the normalized p-Laplacian they propose. In this section we recall these definitions as well as the new definition of the game p-Laplacian on weighted graphs.

5.3.1 Nonlocal statistical operators on weighted graphs

All the definitions below are borrowed from [1] and slightly modified/adjusted to be adapted to our setting and notations.

We first define the following difference operators on graphs needed to define the normalized p-Laplacian on graphs and we give some classical definitions of the nonlocal p-Laplacian on graphs resulting from these defintions.

Let us fix a weighted graph $G_n = (V, E, K_n)$. The *directional derivative* (or edge derivative) of a function u_n at a vertex i along an edge $e = (i, j) \in E(G_n)$, is defined as

$$\partial_j u_n(i) \stackrel{\text{def}}{=} K_{nij}(u_n(j) - u_n(i)).$$

The *difference operator* $\mathcal{G}_{K_n} : \mathcal{H}(V) \to \mathcal{H}(E)$ is given for all $u_n \in \mathcal{H}(V)$ and $(i, j) \in E(G_n)$ by

$$(\mathcal{G}_{K_n} u_n)(i, j) \stackrel{\text{def}}{=} \partial_j u_n(i).$$

The *weighted gradient* of a function $u_n \in \mathcal{H}(V)$ at vertex i is the vector of all edge derivatives

$$(\nabla_{K_n} u_n)(i) \stackrel{\text{def}}{=} ((\partial_j u_n)(i))^\top_{j:(i,j)\in E(G)}.$$

The discrete nonlocal p-Laplacian operator of $u_n \in \mathcal{H}(V)$ (see ($\mathcal{P}^d_{\text{nloc}}$)) evaluated at a vertex $i \in V$ for $1 \le p < \infty$ reads

$$\mathbf{\Delta}_p^{K_n}(u_n)(i) \stackrel{\text{def}}{=} \sum_{j:(i,j)\in E(G)} K_{nij} \left|u_n(j) - u_n(i)\right|^{p-2} (u_n(j) - u_n(i)).$$

For $p = 2$, we obtain the 2-Laplacian as follows:

$$\mathbf{\Delta}_2^{K_n}(u_n)(i) = \sum_{j:(i,j)\in E(G)} K_{nij}(u_n(j) - u_n(i)).$$

For $p = 1$, we obtain the following 1-Laplacian on graphs:

$$\mathbf{\Delta}_1^{K_n}(u_n)(i) \stackrel{\text{def}}{=} \sum_{j:(i,j)\in E(G)} K_{nij} \operatorname{sign}(u_n(j) - u_n(i)),$$

with

$$\operatorname{sign}(x) = \begin{cases} 1 & \text{if} \quad x \ge 0, \\ -1 & \text{otherwise.} \end{cases}$$

The ∞-Laplacian on graphs is defined in [45] by

$$\mathbf{\Delta}_\infty^{K_n}(u_n)(i) \stackrel{\text{def}}{=} \frac{1}{2}\left[\max_{j:(i,j)\in E(G)} (K_{nij} \max((u_n(j) - u_n(i)), 0)) + \min_{j:(i,j)\in E(G)} (K_{nij} \min((u_n(j) - u_n(i)), 0)) \right].$$

Now, we define the following nonlocal statistical operators, which are extensions of the classical local operators (Mean, Max, Min, Midrange, Median):

$$\begin{aligned} \text{NLMean}(u_n)(i) &= \frac{\sum_{j:(i,j)\in E(G)} K_{nij} u_n(j)}{\sum_{j:(i,j)\in E(G)} K_{nij}}, \\ \text{NLMax}(u_n)(i) &= \max_{j:(i,j)\in E(G)} (K_{nij} \max(u_n(j) - u_n(i), 0)) + u_n(i), \\ \text{NLMin}(u_n)(i) &= \max_{j:(i,j)\in E(G)} (K_{nij} \max(u_n(i) - u_n(j), 0)) + u_n(i), \\ \text{NLMidrange}(u_n)(i) &= \frac{1}{2}(\text{NLMin}(u_n)(i) + \text{NLMax}(u_n)(i)), \\ \text{NLMedian}(u_n)(i) &= \text{median}((\nabla_{K_n} u_n)(i)) + u_n(i), \end{aligned} \tag{5.3.1}$$

where median is the classical discrete median operator defined as follows. For $x_i \in \mathbb{R}$ and $i = 1..., m$

$$\text{median}_{1\le j\le m}\{x_j\} = \begin{cases} x_{\left(\frac{m+1}{2}\right)} & \text{if } m \text{ is odd,} \\ \frac{x_{\left(\frac{m}{2}\right)}+x_{\left(\frac{m}{2}+1\right)}}{2} & \text{if } m \text{ is even.} \end{cases} \tag{5.3.2}$$

with $\{x_{(1)}, \cdots, x_{(m)}\}$ a nondecreasing arrangement of $\{x_1, ..., x_m\}$.

One can see that by setting $K_{nij} = 1$, for $(i, j) \in [n]^2$, we recover the classical statistical Mean, Midrange, and Median filters.

Definition 5.3.1 ([1, Section 3.1]). The normalized version of the nonlocal 1-Laplacian, 2-Laplacian and ∞-Laplacian are defined as:

$$\begin{aligned} \boldsymbol{\Delta}^{\text{Nor}}_{K_n,2}(u_n)(i) =& \text{NLMean}(u_n)(i) - u_n(i), \\ \boldsymbol{\Delta}^{\text{Nor}}_{K_n,1}(u_n)(i) =& \text{NLMedian}(u_n)(i) - u_n(i), \\ \boldsymbol{\Delta}^{\text{Nor}}_{K_n,\infty}(u_n)(i) =& \text{NLMidrange}(u_n)(i) - u_n(i). \end{aligned} \tag{5.3.3}$$

An important observation is that these operators are related to partial operators on graphs in the following way.

$$\begin{aligned} \boldsymbol{\Delta}^{\text{Nor}}_{K_n,2}(u_n)(i) &= \frac{1}{\mu(i)}\boldsymbol{\Delta}^{K_n}_{2}(u_n)(i), \\ \boldsymbol{\Delta}^{\text{Nor}}_{K_n,\infty}(u_n)(i) &= \boldsymbol{\Delta}^{K_n}_{\infty}(u_n)(i), \\ \boldsymbol{\Delta}^{\text{Nor}}_{K_n,1}(u_n)(i) &= \text{median}(\nabla_{K_n}(u_n)(i)), \end{aligned} \tag{5.3.4}$$

where $\mu(i)$ is the degree of the vertex $i \in V$.

5.3.2 Game p-Laplacian on graphs

Using the discrete version (5.3.3) of game p-Laplacian with $p = 1$, $p = 2$ and $p = \infty$, we propose the game p-Laplacian on graphs, which can be seen as a nonlocal version of (5.2.5). This is given by the following equations.

$$\boldsymbol{\Delta}^{\text{Nor}}_{K_n,p}(u_n)(i) = \begin{cases} \frac{2}{p}\boldsymbol{\Delta}^{\text{Nor}}_{K_n,2}(u_n)(i) + \frac{p-2}{p}\boldsymbol{\Delta}^{\text{Nor}}_{K_n,\infty}(u_n)(i) & \text{for } 2 \le p \le \infty, \\ \frac{2(p-1)}{p}\boldsymbol{\Delta}^{\text{Nor}}_{K_n,2}(u_n)(i) + \frac{2-p}{p}\boldsymbol{\Delta}^{\text{Nor}}_{K_n,1}(u_n)(i) & \text{for } 1 \le p \le 2. \end{cases} \tag{5.3.5}$$

Using (5.3.3) and (5.3.5), the game p-Laplacian formulation on weighted graphs can be rewritten as

$$\boldsymbol{\Delta}^{\text{Nor}}_{K_n,p}(u_n)(i) = \text{NLA}(u_n)(i) - u_n(i), \tag{5.3.6}$$

where $\text{NLA}(u_n)(i)$ is a nonlocal average operator as

$$\text{NLA}(u_n)(i) = \begin{cases} \frac{2}{p}\text{NLMean}(u_n)(i) + \frac{p-2}{p}\text{NLMidrange}(u_n)(i), \\ \qquad \text{for } 2 \le p \le \infty, \\ \frac{2(p-1)}{p}\text{NLMean}(u_n)(i) + \frac{2-p}{p}\text{NLMedian}(u_n)(i), \\ \qquad \text{for } 1 \le p \le 2. \end{cases} \tag{5.3.7}$$

5.4 The discrete evolution problem

5.4.1 Problem statement

Given a function $g : V \to \mathbb{R}$, we study the following discrete Cauchy problem

$$\begin{cases} \frac{\partial}{\partial t} u_n(i,t) = \boldsymbol{\Delta}^{\text{Nor}}_{K_n,p} u_n(i,t) & \text{in} \quad V \times [0,T], \\ u_n(i,0) = g_{ni} & \text{in} \quad V. \end{cases} \tag{$\mathcal{P}^{\text{Nor},d}_{\text{nloc}}$}$$

Our goal in this section is to study well-posedness of this problem.

5.4.2 Existence and uniqueness

The operator $\boldsymbol{\Delta}^{\text{Nor}}_{K_n,p}$ can be rewritten in the following form

$$\begin{aligned} \boldsymbol{\Delta}^{\text{Nor}}_{K_n,p} u_n(i,t) = \boldsymbol{\Delta}_F u_n(i,t) &= F\big(u_n((j_1 \sim i),t), \cdots, u_n((j_m \sim i),t)\big) - u_n(i,t) \\ &= F\big(u_n\big((j \sim i),t\big)\big) - u_n(i,t), \quad y \in V_m(x),\ m \leq n, \end{aligned}$$

where the operator F will be described later on. We can see that f is a solution of ($\mathcal{P}^{\text{Nor},d}_{\text{nloc}}$) if and only if it is a solution of the integral equation

$$u_n(i,t) = K_g u_n(i,t), \tag{5.4.1}$$

where

$$K_g u_n(i,t) \overset{\text{def}}{=} \int_0^t e^{s-t} F(u_n((j_1 \sim i),t), \cdots, u_n((j_m \sim i),t)) ds + e^{-t} g(i).$$

We verify that the operators defined in (5.3.1) are all averaging operators according to the following definition taken from [75].

Definition 5.4.1. Let $F : \mathbb{R}^m \to \mathbb{R}$ be a continuous function. We call F an averaging operator if it satisfies the following set of conditions :

(i) $F(0, \cdots, 0) = 0$ and $F(1, \cdots, 1) = 1$;

(ii) $F(tx_1, \cdots, tx_m) = tF(x_1, \cdots, x_m)$ for all $t \in \mathbb{R}$;

(iii) $F(t + x_1, \cdots, t + x_m) = t + F(x_1, \cdots, x_m)$ for all $t \in \mathbb{R}$;

(iv) F is nondecreasing with respect to each variable.

Based on this definition, we have the following lemma.

Lemma 5.4.2. *It holds that, if $(x_1, \cdots, x_m)$, $(y_1, \cdots, y_m) \in \mathbb{R}^m$, then*

$$x_j \leq y_j + \max_{1 \leq j \leq m} \{x_j - y_j\} \quad \textit{for all} \quad j \in \{1, \cdots, m\}.$$

Let F be an averaging operator. As a result of combining (iii) and (iv) in definition 5.4.1, we have

$$F(x_1, \cdots, x_m) \leq F(y_1, \cdots, y_m) + \max_{1 \leq j \leq m} \{x_j - y_j\}.$$

Therefore

$$F(x_1, \cdots, x_m) - F(y_1, \cdots, y_m) \leq \max_{1 \leq j \leq m} \{x_j - y_j\},$$

and moreover

$$\big|F(x_1, \cdots, x_m) - F(y_1, \cdots, y_m)\big| \leq \big\|x - y\big\|_\infty. \tag{5.4.2}$$

Lemma 5.4.3. *The operators defined in (5.3.1) are averaging operators.*

PROOF :

- NLMean $u_n(i,t) = F_1(u_n((j_1 \sim i),t),\cdots,u_n((j_m \sim i),t)) = \frac{1}{\sum\limits_{j:(i,j)\in E(G)} K_{nij}} \sum\limits_{j:(i,j)\in E(G)} K_{nij} u_n(j)$.
  (i) $F_1(0,\cdots,0) = 0$ and $F_1(1,\cdots,1) = \frac{1}{\sum\limits_{j:(i,j)\in E(G)} K_{nij}} \sum\limits_{j:(i,j)\in E(G)} K_{nij} = 1$
  (ii) For all $s \in \mathbb{R}$,
$$\begin{aligned} F_1(su_n((j_1 \sim i),t),\cdots,su_n((j_m \sim i),t)) &= \frac{1}{\sum\limits_{j:(i,j)\in E(G)} K_{nij}} \sum_{j:(i,j)\in E(G)} K_{nij} s u_n(j,t) \\ &= \frac{s}{\sum\limits_{j:(i,j)\in E(G)} K_{nij}} \sum_{j:(i,j)\in E(G)} K_{nij} u_n(j,t) \\ &= sF_1(u_n((j_1 \sim i),t),\cdots,u_n((j_m \sim i),t)). \end{aligned}$$
  (iii) For all $s \in \mathbb{R}$
$$\begin{aligned} &F_1(s + u_n((j_1 \sim i),t),\cdots,s + u_n((j_m \sim i),t)) \\ &= \frac{1}{\sum\limits_{j:(i,j)\in E(G)} K_{nij}} \sum_{j:(i,j)\in E(G)} K_{nij}(s + u_n(j,t)) \\ &= \frac{1}{\sum\limits_{j:(i,j)\in E(G)} K_{nij}} \left(s \sum_{j:(i,j)\in E(G)} K_{nij} + \sum_{j:(i,j)\in E(G)} K_{nij} u_n(j,t) \right) \\ &= s + \frac{1}{\sum\limits_{j:(i,j)\in E(G)} K_{nij}} \sum_{j:(i,j)\in E(G)} K_{nij} u_n(j,t) \\ &= s + F_1(u_n((j_1 \sim i),t),\cdots,u_n((j_m \sim i),t)). \end{aligned}$$
  (iv) F_1 is nondecreasing with respect to each variable. Indeed, for $i \in \{1,\cdots,m\}$, taking $g : V \times [0,T] \to \mathbb{R}$, such that $g((j \sim i),t) \geq u_n((j \sim i),t)$. Since the weight function K_{nij} is positive, we have
$$F_1(g(j,t)) \geq F_1(u_n(j,t)).$$
- NLMax $u_n(i,t) = F_2(u_n((j_1 \sim i),t),\cdots,u_n((j_m \sim i),t))$
$$= \max_{j:(i,j)\in E(G)} \big(K_{nij} \max(u_n(j) - u_n(i),0)\big) + u_n(i,t).$$
  (i) $F_2(0,\cdots,0) = 0$ and $F_2(1,\cdots,1) = 1$.
  (ii) For all $s \in \mathbb{R}$,
$$\begin{aligned} &F_2(su_n((j_1 \sim i),t),\cdots,su_n((j_m \sim i),t)) \\ &= \max_{j:(i,j)\in E(G)} (K_{nij} \max(su_n(j,t) - su_n(i,t),0)) + su_n(i,t) \\ &= \max_{j:(i,j)\in E(G)} (sK_{nij} \max((u_n(j,t) - u_n(i,t)),0)) + su_n(i,t) \\ &= s \max_{j:(i,j)\in E(G)} (sK_{nij} \max((u_n(j,t) - u_n(i,t)),0)) + su_n(i,t) \\ &= sF_2(u_n((j_1 \sim i),t),\cdots,u_n((j_m \sim i),t)). \end{aligned}$$
  (iii) For all $s \in \mathbb{R}$,
$$\begin{aligned} &F_2(s + u_n((j_1 \sim i),t),\cdots,s + u_n((j_m \sim i),t)) \\ &= \max_{j:(i,j)\in E(G)} (K_{nij} \max((s + u_n(j,t)) - (s + u_n(i,t)),0)) + (s + u_n(i,t)) \\ &= \max_{j:(i,j)\in E(G)} (K_{nij} \max(u_n(j,t) - u_n(i,t),0)) + u_n(i,t) + s \\ &= s + F_2(u_n((j_1 \sim i),t),\cdots,u_n((j_m \sim i),t)). \end{aligned}$$

(iv) F_2 is nondecreasing with respect to each variable. Indeed, for $i \in \{1, \cdots, m\}$, taking $g : V \times [0, T] \to \mathbb{R}$, such that $g((j \sim i), t) \geq u_n((j \sim i), t)$. Since the weight function K_{nij} is positive, we have

$$\begin{aligned}
&\max(g(j,t) - u_n(i,t), 0) \geq \max(u_n(j,t) - u_n(i,t), 0) \\
&\Rightarrow K_{nij} \max(g(j,t) - u_n(i,t), 0) \geq K_{nij} \max(u_n(j,t) - u_n(i,t), 0) \\
&\Rightarrow F_2(g((j \sim i), t)) \geq F_2(u_n((j \sim i), t)).
\end{aligned}$$

- NLMin $u_n(i,t) = F_3(u_n((j_1 \sim i), t), \cdots, u_n((j_m \sim i), t))$
$= \max_{j:(i,j)\in E(G)} (K_{nij} \max(u_n(i) - u_n(j), 0)) + u_n(i,t)$. By symmetry, using the same arguments as before, we have that NLMin$(\cdot)$ is also an average operator.
- NLMidrange $u_n(i,t) = \frac{1}{2}($NLMin$u_n(i,t)$+NLMax$u_n(i,t)$. By construction, the operator NLMidrange$(\cdot)$ is a linear combination between two averaging operators.
- NLMedian$u_n(i,t) = F_4(u_n((j_1 \sim i), t), \cdots, u_n((j_m \sim i), t))$

$$\begin{aligned}
&= \operatorname*{median}_{j:(i,j)\in E(G)} (K_{nij}(u_n(j,t) - u_n(i,t)) + u_n(i,t)) \\
&= \operatorname*{median}_{j:(i,j)\in E(G)} (K_{nij} u_n(j,t) + (1 - K_{nij}) u_n(i,t)) \\
&= \operatorname*{median}_{1\leq j\leq m} (K_{nij} u_n((j \sim i), t) + (1 - K_{nij}) u_n(i,t)).
\end{aligned}$$

If we call $X_i \stackrel{\text{def}}{=} K_{nij} u_n((j \sim i), t) + (1 - K_{nij})\, u_n(i,t)$, then the median operator is defined as follows

$$\text{median}_{1\leq i\leq m}\{X_i\} \stackrel{\text{def}}{=} \begin{cases} X_{\left(\frac{m+1}{2}\right)} & \text{if } m \text{ is odd,} \\ \frac{X_{\left(\frac{m}{2}\right)} + X_{\left(\frac{m}{2}+1\right)}}{2} & \text{if } m \text{ is even.} \end{cases}$$

with $\{X_{(1)}, \cdots, X_{(m)}\}$ is a nondecreasing rearrangement of $\{X_1, \cdots, X_m\}$.

(i) $F_4(0, \cdots, 0) = 0$, $F_4(1, \cdots, 1) = 1$.

(ii) For all $s \in \mathbb{R}$,

$$\begin{aligned}
F_4(s u_n((j_1 \sim i), t), \cdots, s u_n((j_m \sim i), t)) &= \text{median}_{1\leq i\leq m}\{s X_i\} \\
&= s\ \text{median}_{1\leq i\leq m}\{X_i\}.
\end{aligned}$$

(iii) For all $s \in \mathbb{R}$,

$$\begin{aligned}
F_4(s + u_n((j_1 \sim i), t), \cdots, s + u_n((j_m \sim i), t)) &= \text{median}_{1\leq i\leq m}\{s + X_i\} \\
&= s +\ \text{median}_{1\leq i\leq m}\{X_i\}.
\end{aligned}$$

(iv) F_4 is nondecreasing with respect to each variable. It follows immediately from the definition of the median operator.

□

Let $\ell^\infty(V)$ be the space of bounded vectors in $\mathcal{H}(V)$.

Theorem 5.4.4. *Assume $g \in \ell^\infty(V)$. Then there exists a unique solution in $C(0, T; \ell^\infty(V))$ of $(\mathcal{P}^{\text{Nor},d}_{\text{nloc}})$.*

Proof : By construction, the operator $\boldsymbol{\Delta}^{\text{Nor}}_{K_n,p}$ defined in (5.3.1) is a linear combination of averaging operators, and is in turn itself an averaging operator. It then follows from Lemmas 5.4.3 and 5.4.2 that $\boldsymbol{\Delta}^{\text{Nor}}_{K_n,p}$ is Lipschitz continuous on $\ell^\infty(V)$. This allows us to conclude immediately applying the Cauchy Lipschitz theorem. □

5.5 Numerical experiments

In this section, we illustrate the behavior of the normalized p-Laplacian operator presented in this chapter, through the associated discrete Cauchy problem on graphs. The experiments provided are not here to solve a particular problem but rather to highlight the potentialities of this operator. For this, we solve the discrete evolution Cauchy problem ($\mathcal{P}^{\text{Nor},d}_{\text{nloc}}$) for which the initial function g is application-dependent.

To solve ($\mathcal{P}^{\text{Nor},d}_{\text{nloc}}$) iteratively we use an explicit forward Euler time discretization:

$$\frac{\partial}{\partial t}u_n(i,t) = \frac{u_n^{h+1}(i) - u_n^h(i)}{\Delta t}, \tag{5.5.1}$$

with $u_n^h(i) = u_n(i, h\Delta t)$, $h \in [N]$.

Hence, we can try to solve ($\mathcal{P}^{\text{Nor},d}_{\text{nloc}}$) by the following iteration scheme:

$$\begin{cases} u_n^{h+1}(i) = u_n^h(i) + \Delta t\, \boldsymbol{\Delta}^{\text{Nor}}_{K_n,p} u_n^h(i) & \text{in} \quad V, \\ u_n^0(i) = g_{ni} \quad \text{in} \quad V. \end{cases} \tag{5.5.2}$$

Using $\boldsymbol{\Delta}^{\text{Nor}}_{K_n,p} = \text{NLA}\ (u_n) - u_n$ and setting $\Delta t = 1$, we get the nonlocal average filter

$$\begin{cases} u_n^{h+1}(i) = \begin{cases} \frac{2}{p}\ \text{NLMean}\ u_n^h(i) + \frac{p-2}{p}\ \text{NLMidrange}\ u_n^h(i) & \text{for} \quad 2 \le p \le \infty, \\ \frac{2(p-1)}{p}\ \text{NLMean}\ u_n^h(i) + \frac{2-p}{p}\ \text{NLMedian}\ u_n^h(i) & \text{for} \quad 1 \le p \le 2. \end{cases} \end{cases} \tag{5.5.3}$$

5.5.1 Weighted graph construction

There exists several popular methods to transform discrete data $\{u_1, \cdots, u_n\}$ into a weighted graph G. Considering a set of vertices $V(G)$, the construction of such graphs consists in modeling the neighborhood relationships between the data through the definition of a set of edges E and using a pairwise distance measure $d : V(G) \times V(G) \to \mathbb{R}^+$. In the particular case of images, the ones based on geometric neighborhoods are particularly well-adapted to represent the geometry of the space, as well as the geometry of the function defined on that space. One can quote:

- *Grid graphs* which are most natural structures to describe an image with a graph. Each pixel is connected by an edge to its adjacent pixels. Classical grid graphs are 4-adjacency grid graphs and 8-adjacency grid graphs. Larger adjacency can be used to model nonlocal neighborhoods.
- *k-nearest neighborhood (nn) graphs* where each vertex is connected with its k-nearest neighbors according to d. Such construction implies to build a directed graph, as the neighborhood relationship is not symmetric. Nevertheless, an undirected graph can be obtained while adding an edge between two vertices i and js if i is among the k-nearest neighbors of j or if j is among the k-nearest neighbors of i (see Example 2.1.10).

The weights of the edges will capture the similarity between vertices such that

$$K_{nij} = \begin{cases} s(i,j) & \text{if} \ (i,j) \in E(G), \\ 0, & \text{otherwise}, \end{cases} \tag{5.5.4}$$

where is $: E(G) \to \mathbb{R}^+$ is a similarity function. Typically, one can choose:

- $s_0(i,j) = 1$;
- $s_1(i,j) = e^{-\frac{d(i,j)}{\sigma}}$ with $\sigma > 0$, where d is a metric controlling the similarity between edges, and σ is a scale parameter.
- For patch-based methods, the similarity function is

$$s_2(i,j) = e^{-\frac{d(i,j)^2}{\sigma^2}},$$

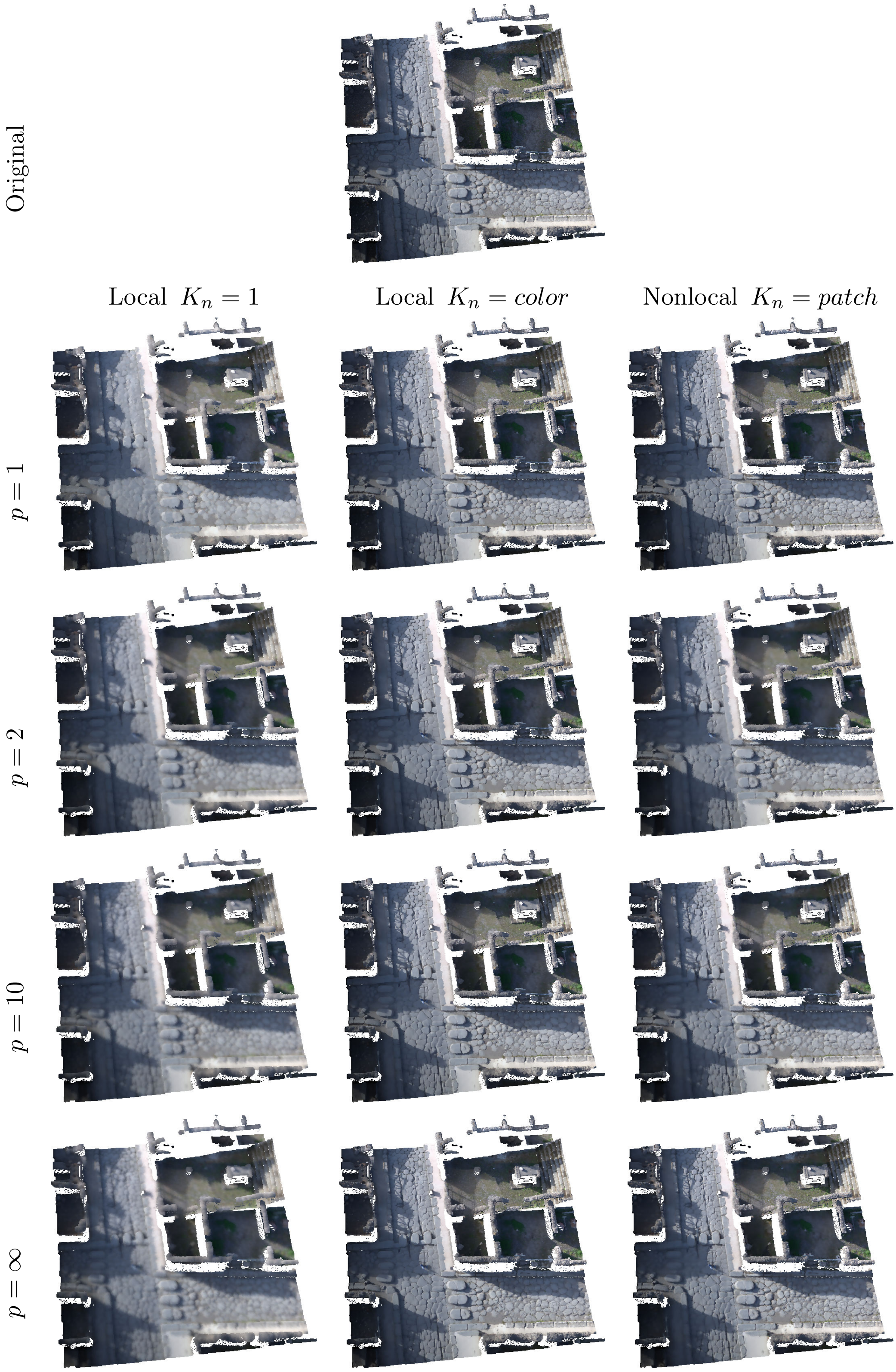


Figure 5.1: Colored point cloud filtering with the Normalized p-Laplacian $\mathbf{\Delta}^{\text{Nor}}_{K_n,p}$. First column presents results with a local knn-graph (with $k = 5$ and $K_n = 1$). Second column presents results under the same configuration but with different similarity function (K_n =color, which depends on the color similarity between two different 3D-points). The last column presents nonlocal results obtained with a larger neighborhood (with K_n depending on patches). In all cases results are provided for $p = 1$, $p = 2$, $p = 10$ and $p = \infty$.

where now $d(i,j) = \left\|\mathcal{P}(j) - \mathcal{P}(i)\right\|_2$, and $\mathcal{P} : i \in V \mapsto \mathcal{P}(i) \in \mathbb{R}^m$ is the patch extraction operator at i. For each node/vertex i, $\mathcal{P}(i)$ is an m-dimensional real vector containing, e.g., spatial coordinates, intensities, etc., of the neighbours of i. This definition of patches is valid only for grid-graphs and cannot be considered for arbitrary graphs. To compute the patch on a 3D point cloud, the reader is referred to [81].

5.5.2 Results

Figure 5.1 and Figure 5.2 show respectively filtering effects on an image and a colored point cloud by implementing (5.5.3) with several parameters. The weight functions K_n are computed from the colors of images or the point cloud. Results are shown with $K_n = s_0$ (constant weight), $K_n = s_1$ (color-based) and $K_n = s_2$ (patch-based).

When $K_n \neq 1$, an adaptive filtering processing (taking into account the difference of the colors in the image/ point cloud) is obtained that can better preserve some features of the graph signal, depending on the graph weights. When patch-based weights ($K_n = s_2$) are considered, repetitive (or texture) patterns are better preserved while providing the usual expected simplification effects.

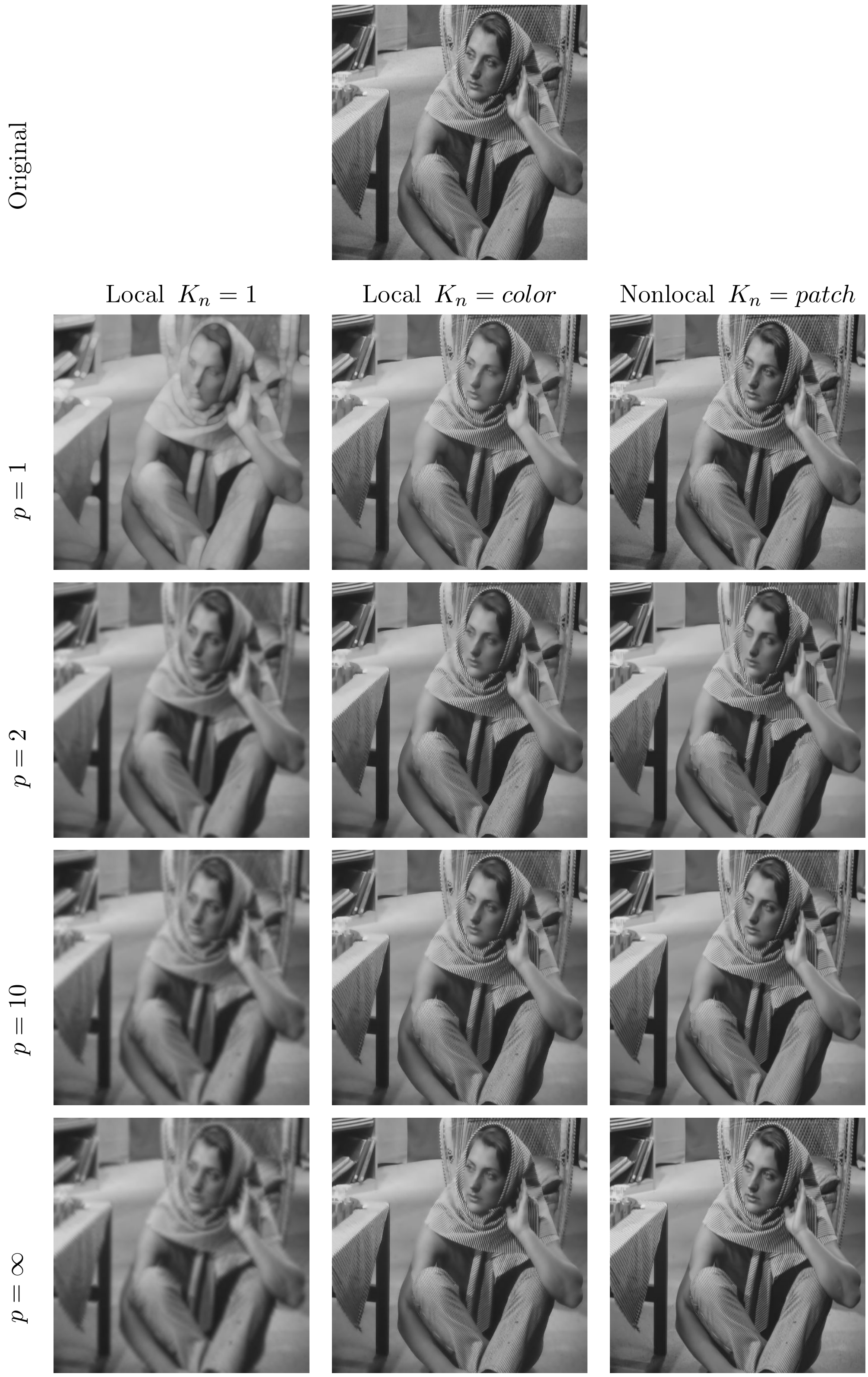


Figure 5.2: Colored image filtering with the Normalized p-Laplacian $\mathbf{\Delta}^{\text{Nor}}_{K_n,p}$. First column presents results with a local 8-adjacency grid graph where each pixel is characterized by it's grayscale value (with $K_n = 1$). Second column presents results under the same configuration but with different similarity function (K_n =color, which depends on the color similarity between two different pixels). The last column presents nonlocal results obtained with a larger neighborhood (with K_n depending on patches). In all cases results are provided for $p = 1$, $p = 2$, $p = 10$ and $p = \infty$.

Part II

The nonlocal p-Laplacian Variational Problem

Chapter 6

General Error Bound

Main contributions of this chapter

- We establish well-posedness of ($\mathcal{VP}_{\mathrm{nloc}}$).
- We give a general error estimate in $L^2(\Omega)$ controlling the error of between the continuous extension of the numerical solution to the discrete variational problem ($\mathcal{VP}^d_{\mathrm{nloc}}$) and its continuum analogue of ($\mathcal{VP}_{\mathrm{nloc}}$) (Theorem 6.3.2).
- The dependence of the error bound on the error induced by discretizing the kernel and the initial data is made explicit.

These results are part of [67].

Contents

6.1 Introduction

6.1.1 Problem statement

Let us recall the variational problem we introduced in Section 1.1.1

$$\min_{u \in L^2(\Omega)} \left\{ E_\lambda(u,g,K) \stackrel{\text{def}}{=} \frac{1}{2\lambda}\left\|u-g\right\|^2_{L^2(\Omega)} + R_p(u,K) \right\}, \tag{$\mathcal{VP}_{\text{nloc}}$}$$

$$R_p(u,K) \stackrel{\text{def}}{=} \frac{1}{2p}\int_{\Omega^2} K(x,y)\left|u(y)-u(x)\right|^p dxdy, \tag{6.1.1}$$

where $p \in [1,+\infty[$ and K and Ω satisfy assumptions (**A.1**)-(**A.2**).

Here λ is a positive regularization parameter that balances the relative importance of the smoothness of the minimizer and fidelity to the initial data. The chief goal of this chapter is to study numerical approximations of the nonlocal variational problem ($\mathcal{VP}_{\text{nloc}}$), which in turn, will allow us to establish consistency estimates of the discrete counterpart of this problem on graphs in Chapter 7.

In the context of image processing, smoothing and denoising are key processing tasks. Among the existing methods, the variational ones, based on nonlocal regularization such as ($\mathcal{VP}_{\text{nloc}}$), provide a popular and versatile framework to achieve these goals. In image processing, such variational problems are in general formulated and studied on the continuum and then discretized on sampled images. On the other hand, many data sources, such as point clouds or meshes, are discrete by nature. Thus, handling such data necessitates a discrete counterpart of ($\mathcal{VP}_{\text{nloc}}$), which reads

$$\min_{u_n \in \mathbb{R}^n} \left\{ E_{n,\lambda} \stackrel{\text{def}}{=} \frac{1}{2\lambda n}\left\|u_n-g_n\right\|^2_2 + R_{n,p}(u_n,K_n) \right\}, \tag{$\mathcal{VP}^d_{\text{nloc}}$}$$

where

$$R_{n,p}(u_n,K_n) \stackrel{\text{def}}{=} \frac{1}{2n^2p}\sum_{i,j=1}^n K_{nij}\left|u_{nj}-u_{ni}\right|^p. \tag{6.1.2}$$

Our aim is to study the relationship between the variational problems ($\mathcal{VP}_{\text{nloc}}$) and ($\mathcal{VP}^d_{\text{nloc}}$). More specifically we aim at deriving error estimates between the corresponding minimizers, respectively $u^\star$ and $u^\star_n$.

6.1.2 Relation to prior work

Nonlocal regularization in machine learning The authors in [59] studied the consistency of *rescaled* total variation minimization on random point clouds in $\mathbb{R}^d$ with a clustering application. They considered the total variation on graphs with a radially symmetric and rescaled kernel $K(x,y) = \varepsilon^{-d}J((x-y)/\varepsilon)$, $\varepsilon > 0$. This corresponds to an instance of $R_{n,p}$ for $d=1$ and $p=1$. For an appropriate scaling of ε with respect to n and under some assumptions on J, those authors they proved that the discrete total variation on graphs Γ-converges in an appropriate topology, as $n \to \infty$, to weighted local total variation, where the weight function is the density of the point cloud distribution. This work were extended in [99] to the graph p-Laplacian for semisupervised learning in $\mathbb{R}^d$. More precisely, the authors considered a constrained and penalized minimization of $R_{n,p}$ with a radially symmetric and rescaled kernel as explained before. They investigated asymptotic behavior when the number of unlabeled points increases, with a fixed number of training points. They uncovered ranges on the scaling of ε with respect to n for the asymptotic consistency (in Γ-convergence sense) to hold. For the same problem, the authors of [2] obtained iterated pointwise convergence of graph p-Laplacians to the continuum p-Laplacian; see [99] for a thorough review in the context of machine learning. Note however that all these results on asymptotic behavior of minimizers do not provide any error estimates for finite n and do not provide precise guidance on what ε would lead to best approximation.

Nonlocal regularization in imaging Several edge-aware filtering schemes have been proposed in the literature [113, 100, 106, 101]. The nonlocal means filter [8] averages pixels that can be arbitrary far away, using a similarity measure based on distance between patches. As shown in [103, 92], these filters can also be interpreted within the variational framework with nonlocal regularization functionals. They correspond to one step of gradient descent on ($\mathcal{VP}^d_{\text{nloc}}$) with $p = 2$, where $K_{nij} = J(x_i - x_j)$ is computed from the input noisy image g using either a distance between the pixels x_i and x_j [113, 106, 101] or a distance between the patches around x_i and x_j [8, 104]. This nonlocal variational denoising can be related to sparsity in an adapted basis of eigenvector of the nonlocal diffusion operator [39, 104, 92]. This nonlocal variational framework was also extended to handle several linear inverse problems [103, 62, 32, 63]. In [94, 52, 112], the authors proposed a variational framework with nonlocal regularizers on graphs to solve linear inverse problems in imaging where both the image to recover and the graph structure are inferred.

Consistency of the ROF model For local variational problems, the only work on consistency that we are aware of is the one of [109] who studied the numerical approximation of the Rudin-Osher-Fatemi (ROF) model, which amounts to minimizing in $L^2(\Omega^2)$ the well-known energy functional

$$E(v) \stackrel{\text{def}}{=} \frac{1}{2\lambda}\|u - g\|^2_{L^2(\Omega^2)} + \|v\|_{\text{TV}(\Omega^2)},$$

where $g \in L^2(\Omega^2)$, and $\|\cdot\|_{\text{TV}(\Omega^2)}$ denotes the total variation seminorm. They bound the difference between the continuous solution and the solutions to various finite-difference approximations to this model. They gave an error estimate in $L^2(\Omega^2)$ of the difference between these two solutions and showed that it scales as $n^{-\frac{s}{2(s+1)}}$, where $s \in]0, 1]$ is the smoothness parameter of the Lipschitz space containing g.

However, to the best of our knowledge, there is no such consistency result in the nonlocal variational setting. In particular, the problem of the continuum limit and consistency of ($\mathcal{VP}^d_{\text{nloc}}$) with error estimates is still open in the literature. It is our aim in this work to rigorously settle this question.

6.2 Well-posedness

Before carrying out the consistency of ($\mathcal{VP}_{\text{nloc}}$), we need to ensure the existence and uniqueness of a solution, that is, the absolute minimizer of problem ($\mathcal{VP}_{\text{nloc}}$). We have the following result:

Theorem 6.2.1. *Suppose that $p \in [1, +\infty[$, K is a nonnegative measurable and bounded mapping, and $g \in L^2(\Omega)$. Then, $E_\lambda(\cdot, g, K)$ has a unique minimizer in $\left\{u \in L^2(\Omega) : R_p(u, K) \le (2\lambda)^{-1}\|g\|^2_{L^2(\Omega)}\right\}$, and $E_{n,\lambda}(\cdot, g_n, K_n)$ has a unique minimizer.*

PROOF : The arguments are standard (coercivity, lower semicontinuity and strict convexity) but we provide a self-contained proof (only for $E_\lambda(\cdot, g, K)$). Let $\{u^\star_k\}_{k\in\mathbb{N}}$ be a minimizing sequence in $L^2(\Omega)$.

By optimality and Jensen's inequality, we have

$$\|u_k^\star\|_{L^2(\Omega)}^2 \le 2\left(2\lambda E_\lambda(u_k^\star, g, K) + \|g\|_{L^2(\Omega)}^2\right) \le 2\left(2\lambda E_\lambda(0, g, K) + \|g\|_{L^2(\Omega)}^2\right) = 4\|g\|_{L^2(\Omega)}^2 < +\infty. \tag{6.2.1}$$

Moreover

$$R_p(u_k^\star, K) \le E_\lambda(u_k^\star, g, K) \le E_\lambda(0, g, K) = \frac{1}{2\lambda}\|g\|_{L^2(\Omega)}^2 < +\infty. \tag{6.2.2}$$

Thus $\|u_k^\star\|_{L^2(\Omega)}$ is bounded uniformly in k so that the Banach-Alaoglu theorem for $L^2(\Omega)$ and compactness provide a weakly convergent subsequence (not relabelled) with a limit $\bar{u} \in L^2(\Omega)$. By lower semicontinuity of the $L^2(\Omega)$ norm with respect to weak convergence and that of $R_p(\cdot, K)$, $\bar{u}$ must be a minimizer. The uniqueness follows from strict convexity of $\|\cdot\|_{L^2(\Omega)}^2$ and convexity of $R_p(\cdot, K)$. ☐

Remark 6.2.2. Theorem 6.2.1 can be extended to linear inverse problems where the data fidelity in $E_\lambda(0, g, K)$ is replaced by $\|g - \mathrm{A}u\|_{L^2(\Sigma)}^2$, and where A is a continuous linear operator. The case where $\mathrm{A} : L^2(\Omega) \to L^2(\Sigma)$ is injective is immediate. The general case is more intricate and would necessitate appropriate assumptions on A and a Poincaré-type inequality. For instance, if $\mathrm{A} : L^p(\Omega) \to L^2(\Sigma)$, and the kernel of A intersects constant functions trivially, then using the Poincaré inequality in [7, Proposition 6.19], one can show existence and uniqueness in $L^p(\Omega)$, and thus in $L^2(\Omega)$ if $p \ge 2$. We omit the details here as this is beyond the scope of the manuscript.

We now turn to provide useful characterization of the minimizers $u^\star$ and $u_n^\star$. We stress that the minimization problem ($\mathcal{VP}_{\text{nloc}}$) that we deal with is considered over $L^2(\Omega)$ ($L^2(\Omega) \subset L^p(\Omega)$ only for $p \in [1, 2]$) over which the function $R_p(\cdot, K)$ may not be finite. In correspondence, we will consider the subdifferential of the proper lower semicontinuous convex function $R_p(\cdot, K)$ on $L^2(\Omega)$ defined as

$$\partial R_p(u, K) \stackrel{\text{def}}{=} \left\{\eta \in L^2(\Omega) : \ R_p(v, K) \ge R_p(u, K) + \langle \eta, v - u\rangle_{L^2(\Omega)}, \ \forall v \in L^2(\Omega)\right\},$$

and $\partial R_p(u, K) = \emptyset$ if $R_p(u, K) = +\infty$.

Lemma 6.2.3. *Suppose that the assumptions of Theorem 6.2.1 hold. Then $u^\star$ is the unique solution to ($\mathcal{VP}_{\text{nloc}}$) if and only if*

$$u^\star = \operatorname{prox}_{\lambda R_p(\cdot, K)}(g) \stackrel{\text{def}}{=} \left(\mathbf{I} + \lambda \partial R_p(\cdot, K)\right)^{-1}(g). \tag{6.2.3}$$

Moreover, the proximal mapping $\operatorname{prox}_{\lambda R_p(\cdot, K)}$ is non-expansive on $L^2(\Omega)$, i.e., for $g_1, g_2 \in L^2(\Omega)$, the corresponding minimizers $u_1^\star, u_2^\star \in L^2(\Omega)$ obey

$$\|u_1^\star - u_2^\star\|_{L^2(\Omega)} \le \|g_1 - g_2\|_{L^2(\Omega)}. \tag{6.2.4}$$

A similar claim is easily obtained for ($\mathcal{VP}_{\text{nloc}}^d$) as well.

Proof : The proof is again classical. By the first order optimality condition and since the squared $L^2(\Omega)$-norm is Fréchet differentiable, $u^\star$ is the unique solution to ($\mathcal{VP}_{\text{nloc}}$) if, and only if,

$$0 \in \frac{1}{2\lambda}(u^\star - g) + \partial R_p(u^\star, K),$$

and the first claim follows. Writing the subgradient inequality for $u_1^\star$ and $u_2^\star$ we have

$$R_p(u_2^\star, K) \ge R_p(u_1^\star, K) + \langle g_1 - u_1^\star, u_2^\star - u_1^\star\rangle_{L^2(\Omega)}$$
$$R_p(u_1^\star, K) \ge R_p(u_2^\star, K) + \langle g_2 - u_2^\star, u_1^\star - u_2^\star\rangle_{L^2(\Omega)}.$$

Adding these two inequalities we get

$$\|u_2^\star - u_1^\star\|_{L^2(\Omega)}^2 \le \langle u_2^\star - u_1^\star, g_2 - g_1\rangle_{L^2(\Omega)},$$

and we conclude upon applying Cauchy-Schwartz inequality. □

We now formally derive the directional derivative of $R_p(\cdot, K)$ when $p \in]1, +\infty[$. For this the symmetry assumption on K is needed as well. Let $h \in L^2(\Omega)$. Then the following derivative exists

$$\frac{d}{dt} R_p(u + th, K)|_{t=0} = \frac{1}{2} \int_{\Omega^2} K(x, y) |u(y) - u(x)|^{p-2} (u(y) - u(x))(v(y) - v(x)) dx dy.$$

Since K is symmetric, we apply the integration by parts formula in [66, Lemma A.1] (or split the integral in two terms and apply a change of variable $(x, y) \mapsto (y, x)$), to conclude that

$$\frac{d}{dt} R_p(u + th, K)|_{t=0} = -\int_{\Omega^2} K(x, y) |u(y) - u(x)|^{p-2} (u(y) - u(x)) v(x) dx dy = \langle \mathbf{\Delta}_p^K, v \rangle_{L^2(\Omega)},$$

where

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

is precisely the nonlocal p-Laplacian operator, see [7, 66]. This shows that under the above assumptions, $R_p(\cdot, K)$ is Fréchet differentiable (hence Gâteaux differentiable) on $L^2(\Omega)$ with Fréchet gradient $\mathbf{\Delta}_p^K$.

6.3 Error estimate for the discrete variational problem

6.3.1 Projector and injector

Let us recall the subdevision of Ω into n intervals

$$\Omega_1^{(n)} = \left[0, \frac{1}{n}\right[, \Omega_2^{(n)} = \left[\frac{1}{n}, \frac{2}{n}\right[, \ldots, \Omega_j^{(n)} = \left[\frac{j-1}{n}, \frac{j}{n}\right[, \ldots, \Omega_n^{(n)} = \left[\frac{n-1}{n}, 1\right[,$$

and recall $\mathcal{Q}_n = \{\Omega_i^{(n)}, i \in [n]\}$ and $\Omega_{ij}^{(n)} \stackrel{\text{def}}{=} \Omega_i^{(n)} \times \Omega_j^{(n)}$. Without loss of generality, we assume that the points are equispaced so that $|\Omega_i^{(n)}| = 1/n$, where $|\Omega_i^{(n)}|$ is the measure of $\Omega_i^{(n)}$. The discussion can be easily extended to non-equispaced points by appropriate normalization; see Section 7.3.

We also consider the operator $P_n : L^1(\Omega) \to \mathbb{R}^n$

$$(P_n v)_i \stackrel{\text{def}}{=} \frac{1}{|\Omega_i^{(n)}|} \int_{\Omega_i^{(n)}} v(x) dx.$$

This operator can be also seen as a piecewise constant projector of u on the space of discrete functions. For simplicity, and with a slight abuse of notation, we keep the same notation for the projector $P_n : L^1(\Omega^2) \to \mathbb{R}^{n \times n}$.

We assume that the discrete initial data g_n and the discrete kernel K_n are constructed as

$$g_n = P_n g \stackrel{\text{def}}{=} (g_{n1}, \cdots, g_{nn})^\top \text{ and } K_n = P_n K \stackrel{\text{def}}{=} (K_{nij})_{1 \le i,j \le n}, \tag{6.3.1}$$

where

$$g_{ni} = (P_n g)_i = \frac{1}{|\Omega_i^{(n)}|} \int_{\Omega_i^{(n)}} g(x) dx \text{ and } K_{nij} = (P_n K)_{ij} = \frac{1}{|\Omega_{ij}^{(n)}|} \int_{\Omega_{ij}^{(n)}} K(x, y) dx dy. \tag{6.3.2}$$

As we mentioned previously, our aim is to study the relationship between the minimizer $u^\star$ of $E_\lambda(\cdot, g, K)$ and the discrete minimizer $u_n^\star$ of $E_{n,\lambda}(\cdot, g_n, K_n)$ and estimate the error between solutions of discrete approximations and the solution of the continuous model. But the solution of problem ($\mathcal{VP}_{\text{nloc}}^d$) being discrete, it is convenient to introduce an intermediate model which is the continuous extension of the discrete solution. Towards this goal, we consider the piecewise constant injector I_n of the discrete functions $u_n^\star$ and g_n into $L^2(\Omega)$, and of K_n into $L^\infty(\Omega^2)$, respectively. This injector I_n is defined as

$$\begin{aligned} I_n u_n(x) &\stackrel{\text{def}}{=} \sum_{i=1}^{n} u_{ni}\chi_{\Omega_i^{(n)}}(x), \\ I_n g_n(x) &\stackrel{\text{def}}{=} \sum_{i=1}^{n} g_{ni}\chi_{\Omega_i^{(n)}}(x), \\ I_n K_n(x,y) &\stackrel{\text{def}}{=} \sum_{i=1}^{n}\sum_{j=1}^{n} K_{nij}\chi_{\Omega_i^{(n)}\times\Omega_j^{(n)}}(x,y), \end{aligned} \tag{6.3.3}$$

where we recall that $\chi_{\mathcal{C}}$ is the characteristic function of the set $\mathcal{C}$, i.e., takes 0 on $\mathcal{C}$ and 1 otherwise.

With these definitions, we have the following well-known properties whose proofs are immediate using the $\left\|\cdot\right\|_{q,n}$ norm defined in (4.0.1) with the usual adaptation for $q=+\infty$.

Lemma 6.3.1. *For a function $v \in L^q(\Omega)$, $q \in [1,+\infty]$, we have*

$$\left\|P_n v\right\|_{q,n} \le \left\|v\right\|_{L^q(\Omega)}; \tag{6.3.4}$$

and for $v_n \in \mathbb{R}^n$

$$\left\|I_n v_n\right\|_{L^q(\Omega)} = \left\|v_n\right\|_{q,n}. \tag{6.3.5}$$

In turn

$$\left\|I_n P_n v\right\|_{L^q(\Omega)} \le \left\|v\right\|_{L^q(\Omega)}. \tag{6.3.6}$$

It is immediate to see that the composition of the operators I_n and P_n yields the operator $\mathbf{P}_{V_n} = I_n P_n$ which is the orthogonal projector on the subspace $V_n \stackrel{\text{def}}{=} \text{Span}\left\{\chi_{\Omega_i^{(n)}} : i \in [n]\right\}$ of $L^1(\Omega)$.

6.3.2 Main result

Our goal is to bound the difference between the unique minimizer of the continuous functional $E_\lambda(\cdot, g, K)$ defined on $L^2(\Omega)$ and the continuous extension by I_n of that of $E_{n,\lambda}(\cdot, g_n, K_n)$. We are now ready to state the main result of this section.

Theorem 6.3.2. *Suppose that $g \in L^2(\Omega)$ and K is a nonnegative measurable, symmetric and bounded mapping. Let $u^\star$ and $u_n^\star$ be the unique minimizers of ($\mathcal{VP}_{\text{nloc}}$) and ($\mathcal{VP}^d_{\text{nloc}}$), respectively. Then, we have the following error bounds.*

(i) If $p \in [1,2]$, then

$$\begin{aligned} \left\|I_n u_n^\star - u^\star\right\|_{L^2(\Omega)}^2 \le C \Bigg(&\left\|g - I_n g_n\right\|_{L^2(\Omega)}^2 + \left\|g - I_n g_n\right\|_{L^2(\Omega)} + \left\|K - I_n K_n\right\|_{L^{\frac{2}{2-p}}(\Omega^2)} \\ &+ \left\|u^\star - I_n P_n u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)} \Bigg), \end{aligned} \tag{6.3.7}$$

where C is a positive constant independent of n.

(ii) If $\inf_{(x,y)\in\Omega^2} K(x,y) \ge \kappa > 0$, then for any $p \in [1,+\infty[$,

$$\begin{aligned} \left\|I_n u_n^\star - u^\star\right\|_{L^2(\Omega)}^2 \le C \Big(&\left\|g - I_n g_n\right\|_{L^2(\Omega)}^2 + \left\|g - I_n g_n\right\|_{L^2(\Omega)} + \left\|K - I_n K_n\right\|_{L^\infty(\Omega^2)} \\ &+ \left\|u^\star - I_n P_n u^\star\right\|_{L^p(\Omega)} \Big), \end{aligned} \tag{6.3.8}$$

where C is a positive constant independent of n.

Observe that $2/(3-p) \le p$ for $p \in [1,2]$. Thus by standard embeddings of $L^q(\Omega)$ spaces for Ω bounded, we have for $p \in [1,2]$

$$\left\|K - I_n K_n\right\|_{L^{\frac{2}{2-p}}(\Omega^2)} \le \left\|K - I_n K_n\right\|_{L^\infty(\Omega^2)} \text{ and } \left\|u^\star - I_n P_n u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)} \le \left\|u^\star - I_n P_n u^\star\right\|_{L^p(\Omega)},$$

which means that our bound in (6.3.7) not only does not require an extra-assumption on K but is also sharper than (6.3.8). The assumption on K in the second statement seems difficult to remove or weaken. Whether this is possible or not is an open question that we leave to a future work.

PROOF :

(i) Since $E_\lambda(\cdot, g, K)$ is a strongly convex function, we have

$$\begin{aligned}\frac{1}{2\lambda}\left\|I_n u_n^\star - u^\star\right\|^2_{L^2(\Omega)} &\leq E_\lambda(I_n u_n^\star, g, K) - E_\lambda(u^\star, g, K) \\ &\leq \left(E_\lambda(I_n u_n^\star, g, K) - E_{n,\lambda}(u_n^\star, g_n, K_n)\right) - \left(E_\lambda(u^\star, g, K) - E_{n,\lambda}(u_n^\star, g_n, K_n)\right).\end{aligned} \tag{6.3.9}$$

A closer inspection of E_λ and $E_{n,\lambda}$ and equality (6.3.5) allows to assert that

$$E_\lambda(I_n u_n^\star, I_n g_n, I_n K_n) = E_{n,\lambda}(u_n^\star, g_n, K_n). \tag{6.3.10}$$

Now, applying the Cauchy-Schwarz inequality and using (6.3.10), we have

$$\begin{aligned}E_\lambda(I_n u_n^\star, g, K) &= \frac{1}{2\lambda}\left\|I_n u_n^\star - g\right\|^2_{L^2(\Omega)} + R_p(I_n u_n^\star, K) \\ &= \frac{1}{2\lambda}\left\|I_n u_n^\star - I_n g_n\right\|^2_{L^2(\Omega)} + \frac{1}{\lambda}\left\langle I_n u_n^\star - I_n g_n, I_n g_n - g\right\rangle_{L^2(\Omega)} \\ &+ \frac{1}{2\lambda}\left\|I_n g_n - g\right\|^2_{L^2(\Omega)} + R_p(I_n u_n^\star, K) \\ &\leq \frac{1}{2\lambda}\left\|I_n u_n^\star - I_n g_n\right\|^2_{L^2(\Omega)} + \frac{1}{\lambda}\left\|I_n u_n^\star - I_n g_n\right\|_{L^2(\Omega)}\left\|I_n g_n - g\right\|_{L^2(\Omega)} \\ &+ \frac{1}{2\lambda}\left\|I_n g_n - g\right\|^2_{L^2(\Omega)} + R_p(I_n u_n^\star, K) \\ &\leq E_{n,\lambda}(u_n^\star, g_n, K_n) + \frac{1}{2\lambda}\left\|I_n g_n - g\right\|^2_{L^2(\Omega)} + \frac{1}{\lambda}\left\|I_n u_n^\star - I_n g_n\right\|_{L^2(\Omega)}\left\|I_n g_n - g\right\|_{L^2(\Omega)} \\ &+ \left(R_p(I_n u_n^\star, K) - R_p(I_n u_n^\star, I_n K_n)\right) \\ &\leq E_{n,\lambda}(u_n^\star, g_n, K_n) + \frac{1}{2\lambda}\left\|I_n g_n - g\right\|^2_{L^2(\Omega)} + \frac{1}{\lambda}\left\|I_n u_n^\star - I_n g_n\right\|_{L^2(\Omega)}\left\|I_n g_n - g\right\|_{L^2(\Omega)} \\ &+ \frac{1}{2p}\left|\int_{\Omega^2}\left(K(x,y) - I_n K_n(x,y)\right)\left|I_n u_n^\star(y) - I_n u_n^\star(x)\right|^p dxdy\right|.\end{aligned} \tag{6.3.11}$$

As we suppose that $g \in L^2(\Omega)$ and since $I_n u_n^\star$ is the (unique) minimizer of $E_\lambda(\cdot, I_n g_n, I_n K_n)$ (by virtue of (6.3.10)), it is immediate to see, using (6.3.6), that

$$\begin{aligned}\frac{1}{2\lambda}\left\|I_n u_n^\star - I_n g_n\right\|^2_{L^2(\Omega)} &\leq \frac{1}{2\lambda}\left\|I_n u_n^\star - I_n g_n\right\|^2_{L^2(\Omega)} + R_p(I_n u_n^\star, I_n K_n) \\ &\leq E_\lambda(0, I_n g_n, I_n K_n) \\ &= \frac{1}{2\lambda}\left\|I_n g_n\right\|^2_{L^2(\Omega)} \\ &= \frac{1}{2\lambda}\left\|I_n P_n g\right\|^2_{L^2(\Omega)} \\ &\leq \frac{1}{2\lambda}\left\|g\right\|^2_{L^2(\Omega)} < +\infty,\end{aligned}$$

and thus

$$\left\|I_n u_n^\star - I_n g_n\right\|_{L^2(\Omega)} \leq \left\|g\right\|_{L^2(\Omega)} \stackrel{\text{def}}{=} C_1. \tag{6.3.12}$$

Since $p \in [1,2]$, by Hölder and triangle inequalities, and (6.2.1) applied to $I_n u_n^\star$, we have that

$$\begin{aligned}
&\left|\int_{\Omega^2} \left(K(x,y) - I_n K_n(x,y)\right)\left|I_n u_n^\star(y) - I_n u_n^\star(x)\right|^p dxdy\right| \\
&\leq \left\|K - I_n K_n\right\|_{L^{\frac{2}{2-p}}(\Omega^2)} \left(\int_{\Omega^2} \left|I_n u_n^\star(y) - I_n u_n^\star(x)\right|^2 dxdy\right)^{p/2} \\
&\leq 2^p \left\|I_n u_n^\star\right\|^p_{L^2(\Omega)} \left\|K - I_n K_n\right\|_{L^{\frac{2}{2-p}}(\Omega^2)} \\
&\leq 2^{2p} \left\|I_n P_n g\right\|^p_{L^2(\Omega)} \left\|K - I_n K_n\right\|_{L^{\frac{2}{2-p}}(\Omega^2)} \\
&\leq 2^{2p} \left\|g\right\|^p_{L^2(\Omega)} \left\|K - I_n K_n\right\|_{L^{\frac{2}{2-p}}(\Omega^2)} = C_2 \left\|K - I_n K_n\right\|_{L^{\frac{2}{2-p}}(\Omega^2)},
\end{aligned} \tag{6.3.13}$$

where $C_2 \stackrel{\text{def}}{=} 2^{2p} C_1^p$.

We now turn to bounding the second term on the right-hand side of (6.3.9). Using (6.3.6) and the fact that $u_n^\star$ is the (unique) minimizer of ($\mathcal{VP}^d_{\text{nloc}}$), we have

$$\begin{aligned}
E_\lambda(I_n u_n^\star, I_n g_n, I_n K_n) &\leq E_\lambda(I_n P_n u^\star, I_n g_n, I_n K_n) \\
&= \frac{1}{2\lambda}\left\|I_n P_n u^\star - I_n P_n g\right\|^2_{L^2(\Omega)} + R_p(I_n P_n u^\star, I_n K_n) \\
&\leq \frac{1}{2\lambda}\left\|u^\star - g\right\|^2_{L^2(\Omega)} + R_p(u^\star, K) + R_p(I_n P_n u^\star, I_n K_n) - R_p(u^\star, K) \\
&\leq E_\lambda(u^\star, g, K) + (R_p(I_n P_n u^\star, K) - R_p(u^\star, K)) \\
&+ (R_p(I_n P_n u^\star, I_n K_n) - R_p(I_n P_n u^\star, K)).
\end{aligned} \tag{6.3.14}$$

We bound the second term on the right-hand side of (6.3.14) by applying the mean value theorem on $[a(x,y), b(x,y)]$ to the function $t \in \mathbb{R}^+ \mapsto t^p$ with $a(x,y) = |u^\star(y) - u^\star(x)|$ and $b(x,y) = |I_n P_n u^\star(y) - I_n P_n u^\star(x)|$. Let $\eta(x,y) \stackrel{\text{def}}{=} \rho a(x,y) + (1-\rho) b(x,y)$, $\rho \in [0,1]$, be an intermediate value between $a(x,y)$ and $b(x,y)$. We then get

$$\begin{aligned}
&\left|R_p(I_n P_n u^\star, K) - R_p(u^\star, K)\right| \\
&= \Big|\int_{\Omega^2} K(x,y)\left(\left|I_n P_n u^\star(y) - I_n P_n u^\star(x)\right|^p - \left|u^\star(y) - u^\star(x)\right|^p\right) dxdy\Big| \\
&= p\Big|\int_{\Omega^2} K(x,y)\eta(x,y)^{p-1}\left(\left|I_n P_n u^\star(y) - I_n P_n u^\star(x)\right| - \left|u^\star(y) - u^\star(x)\right|\right) dxdy\Big| \\
&\leq pC_3 \int_{\Omega^2} \eta(x,y)^{p-1}\left|\left(I_n P_n u^\star(y) - u^\star(y)\right) - \left(I_n P_n u^\star(x) - u^\star(x)\right)\right| dxdy \\
&\leq 2pC_3 \int_{\Omega^2} \eta(x,y)^{p-1}\left|I_n P_n u^\star(x) - u^\star(x)\right| dxdy,
\end{aligned} \tag{6.3.15}$$

where we used the triangle inequality, symmetry after the change of variable $(x,y) \mapsto (y,x)$, and boundedness of K, say $\left\|K\right\|_{L^\infty(\Omega^2)} \stackrel{\text{def}}{=} C_3$. Thus using Hölder and Jensen inequalities as well as (6.3.6), and arguing as in (6.3.13), leads to

$$\begin{aligned}
&\left|R_p(I_n P_n u^\star, K) - R_p(u^\star, K)\right| \\
&\leq 2pC_3 \left\|\eta\right\|^{p-1}_{L^2(\Omega^2)} \left\|u^\star - I_n P_n u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)} \\
&\leq 2pC_3 \left(\rho\left\|a\right\|_{L^2(\Omega^2)} + (1-\rho)\left\|b\right\|_{L^2(\Omega^2)}\right)^{p-1} \left\|u^\star - I_n P_n u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)} \\
&\leq 2pC_3 \left\|a\right\|^{p-1}_{L^2(\Omega^2)} \left\|u^\star - I_n P_n u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)} \\
&\leq 2^{2p-1} pC_3 \left\|g\right\|^{p-1}_{L^2(\Omega)} \left\|u^\star - I_n P_n u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)} = C_4 \left\|u^\star - I_n P_n u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)}
\end{aligned} \tag{6.3.16}$$

where $C_4 \stackrel{\text{def}}{=} 2^{2p-1} p C_1^{p-1}$.

To bound the last term on the right-hand side of (6.3.14), we follow the same steps as for establishing (6.3.13) and get

$$\begin{aligned}&|R_p(I_nP_nu^\star, I_nK_n) - R_p(I_nP_nu^\star, K)|\\&\leq \int_{\Omega^2} \big|K(x,y) - I_nK_n(x,y)\big|\big|I_nP_nu^\star(y) - I_nP_nu^\star(x)\big|^p dxdy\\&\leq C_2\big\|K - I_nK_n\big\|_{L^{\frac{2}{2-p}}(\Omega^2)}.\end{aligned} \tag{6.3.17}$$

Finally, plugging (6.3.11), (6.3.12), (6.3.13), (6.3.14), (6.3.16) and (6.3.17) into (6.3.9), we get the desired result.

(ii) The case $p \geq 2$ follows the same proof steps, except that now, we need to modify inequalities (6.3.13), (6.3.16) and (6.3.17) which do not hold anymore.
Under our assumption on K, and using (6.2.2), (6.3.13) now reads

$$\begin{aligned}&\int_{\Omega^2} \big|K(x,y) - I_nK_n(x,y)\big|\big|I_nu_n^\star(y) - I_nu_n^\star(x)\big|^p dxdy\\&\leq \kappa^{-1}\big\|K - I_nK_n\big\|_{L^\infty(\Omega^2)} \int_{\Omega^2} I_nK_n(x,y)\big|I_nu_n^\star(y) - I_nu_n^\star(x)\big|^p dxdy\\&= \kappa^{-1}\big\|K - I_nK_n\big\|_{L^\infty(\Omega^2)} R_p(I_nu_n^\star, I_nK_n)\\&\leq (2\lambda\kappa)^{-1}C_1^2\big\|K - I_nK_n\big\|_{L^\infty(\Omega^2)},\end{aligned} \tag{6.3.18}$$

where $C_1 = \|g\|_{L^2(\Omega)}$ as in the proof of (i).
Applying Hölder inequality in (6.3.15) and using again (6.2.2) and the assumption on K, we obtain

$$\begin{aligned}&\big|R_p(I_nP_nu^\star, K) - R_p(u^\star, K)\big|\\&\leq 2pC_3\left(\int_{\Omega^2}\big|I_nP_nu^\star(y) - I_nP_nu^\star(x)\big|^p dxdy\right)^{(p-1)/p}\big\|u^\star - I_nP_nu^\star\big\|_{L^p(\Omega)}\\&\leq 2\kappa^{(1-p)/p}pC_3\left(\int_{\Omega^2} I_nK_n(x,y)\big|I_nP_nu^\star(y) - I_nP_nu^\star(x)\big|^p dxdy\right)^{(p-1)/p}\big\|u^\star - I_nP_nu^\star\big\|_{L^p(\Omega)}\\&= 2\kappa^{(1-p)/p}pC_3\left(R_p(I_nu_n^\star, I_nK_n)\right)^{(p-1)/p}\big\|u^\star - I_nP_nu^\star\big\|_{L^p(\Omega)}\\&\leq 2(2\lambda\kappa)^{(1-p)/p}pC_3C_1^{2(p-1)/p}\big\|u^\star - I_nP_nu^\star\big\|_{L^p(\Omega)}.\end{aligned} \tag{6.3.19}$$

To get the new form of (6.3.17), we use (6.3.6), (6.2.2) and the assumption on K to arrive at

$$\begin{aligned}&|R_p(I_nP_nu^\star, I_nK_n) - R_p(I_nP_nu^\star, K)|\\&\leq \int_{\Omega^2} \big|K(x,y) - I_nK_n(x,y)\big|\big|I_nP_nu^\star(y) - I_nP_nu^\star(x)\big|^p dxdy\\&\leq \big\|K - I_nK_n\big\|_{L^\infty(\Omega^2)} \int_{\Omega^2} \big|u^\star(y) - u^\star(x)\big|^p dxdy\\&\leq \kappa^{-1}\big\|K - I_nK_n\big\|_{L^\infty(\Omega^2)} \int_{\Omega^2} K(x,y)\big|u^\star(y) - u^\star(x)\big|^p dxdy\\&= \kappa^{-1}\big\|K - I_nK_n\big\|_{L^\infty(\Omega^2)} R_p(u^\star, K)\\&\leq (2\lambda\kappa)^{-1}C_1^2\big\|K - I_nK_n\big\|_{L^\infty(\Omega^2)}.\end{aligned} \tag{6.3.20}$$

Plugging now (6.3.11), (6.3.12), (6.3.14), (6.3.18), (6.3.19) and (6.3.20) into (6.3.9), we conclude the proof.

□

6.3.3 Regularity of the minimizer

Thee error bound of Theorem 6.3.2 contain three terms: one which corresponds to the error in discretizing g, the second is the discretization error of the kernel K, and the last term reflects the discretization error of the minimizer $u^\star$ of the continuous problem ($\mathcal{VP}_{\text{nloc}}$). Thus, this form is not convenient to transfer our bounds to networks on graph and establish convergence rates. Clearly, we need a control on the term $\left\|I_n P_n u^\star - u^\star\right\|_{L^q(\Omega)}$ on the right-hand side of (6.3.7)-(6.3.8). This is what we are about to do in the following key regularity lemma. In a nutshell, it states that if the kernel K only depends on $|x-y|$ (as is the case for many kernels used in data processing), then as soon as the initial data g belongs to some Lipschitz space, so does the minimizer $u^\star$.

Lemma 6.3.3. *Suppose $g \in L^\infty(\Omega) \cap \mathrm{Lip}(s, L^q(\Omega))$ with $s \in]0,1]$ and $q \in [1,+\infty]$. Suppose furthermore that $K(x,y) = J(|x-y|)$, where J is a nonnegative bounded measurable mapping on Ω.*

(i) If $q \in [1,2]$, then $u^\star \in \mathrm{Lip}(sq/2, L^q(\Omega))$.

(ii) If $q \in [2,+\infty]$, then $u^\star \in \mathrm{Lip}(sq/2, L^2(\Omega))$.

The boundedness assumption on g can be removed for $q = 2$.

Proof : We denote the torus $\mathbb{T} \overset{\text{def}}{=} \mathbb{R}/2\mathbb{Z}$. For any function $u \in L^2(\Omega)$, we denote by $\bar{u} \in L^2(\mathbb{T})$ its periodic extension such that

$$\bar{u}(x) = \begin{cases} u(x) & \text{if} \quad x \in [0,1], \\ u(2-x) & \text{if} \quad x \in]1,2], \end{cases} \tag{6.3.21}$$

In the rest of the proof, we use letters with bars to indicate functions defined on $\mathbb{T}$.

Let us define

$$\bar{E}_{\lambda/2}(\bar{v}, \bar{g}, \bar{J}) \overset{\text{def}}{=} \frac{1}{\lambda}\left\|\bar{v} - \bar{g}\right\|^2_{L^2(\mathbb{T})} + \bar{R}_p(\bar{v}, \bar{J})$$

where

$$\bar{R}_p(\bar{v}, \bar{J}) \overset{\text{def}}{=} \frac{1}{2p}\int_{\mathbb{T}^2} \bar{J}(|x-y|)\left|\bar{v}(y) - \bar{v}(x)\right|^p dxdy.$$

Consider the following minimization problem

$$\min_{\bar{v} \in L^2(\mathbb{T})} \bar{E}_{\lambda/2}(\bar{v}, \bar{g}, \bar{J}), \tag{6.3.22}$$

which also has a unique minimizer by arguments similar to those of Theorem 6.2.1. Since $u^\star$ is the unique minimizer of ($\mathcal{VP}_{\text{nloc}}$), we have, using (6.3.21),

$$\begin{aligned} \bar{E}_{\lambda/2}(\bar{u^\star}, \bar{g}, \bar{J}) &= \frac{2}{\lambda}\left\|u^\star - g\right\|^2_{L^2(\Omega)} + 4R_p(u^\star, J) \\ &= 4E_\lambda(u^\star, g, J) \\ &< 4E_\lambda(v, g, J), \forall v \neq u^\star \\ &= \bar{E}_{\lambda/2}(\bar{v}, \bar{g}, \bar{J}), \forall \bar{v} \neq \bar{u^\star}, \end{aligned} \tag{6.3.23}$$

which shows that $\bar{u^\star}$ is the unique minimizer of (6.3.22). Then, we have via Lemma 6.2.3

$$\bar{u^\star} = \mathrm{prox}_{\lambda/2\bar{R}_p(\cdot,\bar{J})}(\bar{g}). \tag{6.3.24}$$

We define the translation operator

$$(T_h v)(x) = v(x+h), \forall h \in \mathbb{R}.$$

Now, using our assumption on the kernel K, that is $K(x,y) = J(|x-y|)$ (then invariant by translation), and periodicity of the functions on $\mathbb{T}$, we have

$$\begin{aligned}
\bar{E}_{\lambda/2}(\bar{v}, T_h\bar{g}, \bar{J}) &= \frac{1}{\lambda}\big\|\bar{v} - T_h\bar{g}\big\|^2_{L^2(\mathbb{T})} + \bar{R}_p(\bar{v}, \bar{J}) \\
&= \frac{1}{\lambda}\big\|T_h(T_{-h}\bar{v} - \bar{g})\big\|^2_{L^2(\mathbb{T})} \\
&\quad + \int_{\mathbb{T}^2} \bar{J}(|x-y|)\big|\bar{v}((y+h)-h) - \bar{v}((x+h)-h)\big|^p dxdy \\
&= \frac{1}{\lambda}\big\|T_{-h}\bar{v} - \bar{g}\big\|^2_{L^2(\mathbb{T})} + \int_{\mathbb{T}^2} \bar{J}(|x-y|)\big|T_{-h}\bar{v}(y) - T_{-h}\bar{v}(x)\big|^p dxdy \\
&= \frac{1}{\lambda}\big\|T_{-h}\bar{v} - \bar{g}\big\|^2_{L^2(\mathbb{T})} + \int_{\mathbb{T}^2} \bar{J}(|x-y|)\big|T_{-h}\bar{v}(y) - T_{-h}\bar{v}(x)\big|^p dxdy \\
&= \bar{E}_{\lambda/2}(T_{-h}\bar{v}, \bar{g}, \bar{J}).
\end{aligned}$$

This implies that the unique minimizer $\bar{v}^\star$ of $\bar{E}_{\lambda/2}(\cdot, T_h\bar{g}, \bar{J})$ given by (see Lemma 6.2.3)

$$\bar{v}^\star = \mathrm{prox}_{\lambda/2\bar{R}_p(\cdot,\bar{J})}(T_h\bar{g}), \tag{6.3.25}$$

is also the unique minimizer of $\bar{E}_{\lambda/2}(T_{-h}\cdot, \bar{g}, \bar{J})$. But since $\bar{E}_{\lambda/2}(\cdot, \bar{g}, \bar{J})$ has a unique minimizer $\bar{u^\star}$, we deduce from (6.3.24) and (6.3.25) that

$$T_h\mathrm{prox}_{\lambda/2\bar{R}_p(\cdot,\bar{J})}(\bar{g}) = \mathrm{prox}_{\lambda/2\bar{R}_p(\cdot,\bar{J})}(T_h\bar{g}). \tag{6.3.26}$$

That is, the proximal mapping of $\lambda/2\bar{R}_p(\cdot, \bar{J})$ commutes with translation.

We now split the two cases of q.

(i) For $q \in [1, 2]$: combining (6.3.24), (6.3.26), (6.2.4), [66, Lemma C.1] and that $L^2(\Omega) \subset L^q(\Omega)$, we have

$$\begin{aligned}
\big\|T_h\bar{u^\star} - \bar{u^\star}\big\|_{L^q(\mathbb{T})} &= \big\|\mathrm{prox}_{\lambda/2\bar{R}_p(\cdot,\bar{J})}(T_h\bar{g}) - \mathrm{prox}_{\lambda/2\bar{R}_p(\cdot,\bar{J})}(\bar{g})\big\|_{L^q(\mathbb{T})} \\
&\le \big\|\mathrm{prox}_{\lambda/2\bar{R}_p(\cdot,\bar{J})}(T_h\bar{g}) - \mathrm{prox}_{\lambda/2\bar{R}_p(\cdot,\bar{J})}(\bar{g})\big\|_{L^2(\mathbb{T})} \\
&\le \big\|T_h\bar{g} - \bar{g}\big\|_{L^2(\mathbb{T})} \\
&\le \big\|g\big\|^{1-q/2}_{L^\infty(\Omega)}\big\|T_h\bar{g} - \bar{g}\big\|^{q/2}_{L^q(\mathbb{T})} \le C_1\big\|T_h\bar{g} - \bar{g}\big\|^{q/2}_{L^q(\mathbb{T})}.
\end{aligned} \tag{6.3.27}$$

Let $\Omega_h \stackrel{\text{def}}{=} \{x \in \Omega : \ x + h \in \Omega\}$. Recalling the modulus of smoothness in (2.3.1), we have

$$\begin{aligned}
w(u^\star, t)_q \stackrel{\text{def}}{=} \sup_{|h|<t} \big\|T_h u^\star - u^\star\big\|_{L^q(\Omega_h)} &\le C_2 \sup_{|h|<t} \big\|T_h\bar{u^\star} - \bar{u^\star}\big\|_{L^q(\mathbb{T})} \\
&\le C_1C_2\left(\sup_{|h|<t}\big\|T_h\bar{g} - \bar{g}\big\|_{L^q(\mathbb{T})}\right)^{q/2} \\
&= C_1C_2 w(\bar{g}, t)_q^{q/2} \\
&\le C_1C_2(C_3 w(g,t)_q)^{q/2}.
\end{aligned} \tag{6.3.28}$$

We get the last inequality by applying the Whitney extension theorem [41, Ch. 6, Theorem 4.1]. Invoking Definition 2.3.1, there exists a constant $C > 0$ such that

$$|u^\star|_{\mathrm{Lip}(sq/2, L^q(\Omega))} \stackrel{\text{def}}{=} \sup_{t>0} t^{-sq/2} w(u^\star, t)_q \le C\left(\sup_{t>0} t^{-s} w(u^\star, t)_q\right)^{q/2} \le C\,|g|^{q/2}_{\mathrm{Lip}(s, L^q(\Omega))}, \tag{6.3.29}$$

whence the claim follows after observing that $u^\star \in L^2(\Omega) \subset L^q(\Omega)$.

(ii) For $q \in [2, +\infty]$, we argue as in (6.3.27) to show that

$$\big\|T_h\bar{u^\star} - \bar{u^\star}\big\|_{L^2(\mathbb{T})} \le C_1\big\|T_h\bar{g} - \bar{g}\big\|^{q/2}_{L^q(\mathbb{T})}.$$

The rest of the proof is similar to that of (i).

□

In view of the regularity Lemma 6.3.3 and Theorem 6.3.2, one can derive convergence rates but only for $p \in [1, 2]$. Indeed, the approximation bounds of Lemma 2.3.2 cannot be applied to $u^\star - I_n P_n u^\star$ for $p \geq 2$ since the bound in Theorem 6.3.2(ii) is in the $L^p(\Omega)$ norm while Lemma 6.3.3 proves that $u^\star$ is only in $\mathrm{Lip}(sq/2, L^2(\Omega))$. In particular, one cannot invoke (2.3.3) since there is no guarantee that $u^\star$ is bounded. This is the reason why in Chapter 7, we will only focus on the case $p \in [1, 2]$.

Chapter 7

Convergence Rates for Networks on Convergent Graph Sequences

Main contributions of this chapter

- We apply the error estimate of Chapter 6 to networks on simple and weighted dense graphs and we show that the approximation of minimizers of the discrete problems on simple and weighted graph sequences converge to those of the continuous problem.
- Under very mild conditions on the kernel and the initial data, typically belonging to Lipschitz functional spaces, precise convergence rates are exhibited.
- We study networks on random inhomogeneous graphs. We establish nonasymptotic convergence claims and give the rate of convergence of the discrete solution to its continuous limit with high probability under the same assumptions on the kernel and the initial data.
- We reveal the role of the data regularity/geometry of the graph models and the parameter p on the rate of convergence.

These results are part of [67] .

Contents

In this chapter, we present an analysis of networks on convergent graph sequences for the variational p-Laplacian problem. Our results include three main parts: We show that the approximation of minimizers of the discrete problems on simple and weighted graph sequences converge to those of the continuous problem.This sets the question that solving a discrete variational problem on graphs has indeed a continuum limit. Under very mild conditions on K and g, typically belonging to Lipschitz functional spaces, precise convergence rates can be exhibited. These functional spaces allow to cover a large class of graphs (through K) and initial data g, including those functions of bounded variation. For simple graph sequences, we also show how the accuracy of the approximation depends on the regularity of the boundary of the support of the graph limit. Finally, building upon these error estimates, we study networks on random inhomogeneous graphs. We combine them with sharp deviation inequalities to establish nonasymptotic convergence claims and give the rate of convergence of the discrete solution to its continuous limit with high probability under the same assumptions on the kernel K and the initial data g.

7.1 Networks on simple graphs

Recall the construction of the simple graph model $\{G_n\}_{n\in\mathbb{N}^*}$ described in Section 2.1.3.1. The discrete counterpart of ($\mathcal{VP}_{\text{nloc}}$) on the graph G_n is then given by

$$\min_{u_n\in\mathbb{R}^n}\left\{E_{n,\lambda}(u_n,g_n,K_n)\stackrel{\text{def}}{=}\frac{1}{2\lambda n}\left\|u_n-g_n\right\|_2^2+\frac{1}{n^2}\sum_{i,j:(i,j)\in E(G_n)}\left|u_{nj}-u_{ni}\right|^p\right\}, \qquad (\mathcal{VP}^d_{s,\text{nloc}})$$

where the initial data g_n is given by (6.3.2). For this model, $I_nK_n(x,y)$ is the piecewise constant function such that for $(x,y)\in\Omega_{ij}^{(n)}$, $(i,j)\in[n]^2$

$$I_nK_n(x,y)=\begin{cases}\frac{1}{|\Omega_{ij}^{(n)}|}\displaystyle\int_{\Omega_{ij}^{(n)}}K(x,y)dxdy & \text{if} \quad \Omega_{ij}^{(n)}\cap\overline{\text{supp}(K)}\neq\emptyset,\\ 0 & \text{otherwise.}\end{cases} \tag{7.1.1}$$

Relying on what we did in Sectionsimplegraphs, the rate of convergence of the solution of the discrete problem to the solution of the limiting problem depends on the regularity of the boundary $\text{bd}(\overline{\text{supp}(K)})$ of the support closure. Recall the upper box-counting (or Minkowski-Bouligand) dimension ρ defined in (4.2.3).

Theorem 7.1.1. *Assume that $p\in[1,2]$, $g\in L^2(\Omega)$. Let $u^\star$ and $u_n^\star$ be the unique minimizers of* ($\mathcal{VP}_{\text{nloc}}$) *and* ($\mathcal{VP}^d_{s,\text{nloc}}$)*, respectively. Then, the following hold.*

(i) We have

$$\left\|I_nu_n^\star-u^\star\right\|_{L^2(\Omega)}\underset{n\to+\infty}{\longrightarrow}0.$$

(ii) For $p \in [1,2[$: assume moreover $g \in L^\infty(\Omega) \cap \mathrm{Lip}(s, L^q(\Omega))$, with $s \in]0,1]$ and $q \in [2/(3-p), 2]$, that $\rho \in [0,2[$ and that $K(x,y) = J(|x-y|)$, $\forall (x,y) \in \Omega^2$, with J a nonnegative bounded measurable mapping on Ω. Then for any $\epsilon > 0$ there exists $N(\epsilon) \in \mathbb{N}$ such that for any $n \geq N(\epsilon)$

$$\left\|I_n u_n^\star - u^\star\right\|^2_{L^2(\Omega))} \leq C n^{-\min\left\{sq/2,(2-p)(1-\frac{\rho+\epsilon}{2})\right\}},$$

where C is a positive constant independent of n.

(iii) For $p = 2$: under the same assumptions as (ii), we have

$$\left\|I_n u_n^\star - u^\star\right\|^2_{L^2(\Omega))} \leq C n^{-\min\{sq/2,2\}},$$

where C is a positive constant independent of n.

Proof :

(i) In view of (6.3.2), by the Lebesgue differentiation theorem (see e.g. [89, Theorem 3.4.4]), we have

$$I_n g_n(x) \underset{n\to\infty}{\longrightarrow} g(x), \quad I_n P_n u^\star(x) \underset{n\to\infty}{\longrightarrow} u^\star(x) \text{ and } I_n K_n(x,y) \underset{n\to\infty}{\longrightarrow} K(x,y)$$

almost everywhere on Ω and Ω^2, respectively. Combining this with Fatou's lemma and (6.3.6), we have

$$\begin{aligned}\left\|g\right\|^2_{L^2(\Omega)} = \int_\Omega \left|\lim_{n\to\infty} I_n g_n(x)\right|^2 dx &= \int_\Omega \liminf_{n\to\infty} |I_n g_n(x)|^2 dx \\ &\leq \liminf_{n\to\infty} \left\|I_n g_n\right\|^2_{L^2(\Omega)} \\ &\leq \limsup_{n\to\infty} \left\|I_n P_n g\right\|^2_{L^2(\Omega)} \leq \left\|g\right\|^2_{L^2(\Omega)},\end{aligned}$$

which entails that $\lim_{n\to\infty} \left\|I_n g_n\right\|_{L^2(\Omega)} = \left\|g\right\|_{L^2(\Omega)}$. Similarly, we have $\lim_{n\to\infty} \left\|I_n P_n u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)} = \left\|u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)}$. Since $g \in L^2(\Omega)$, $u^\star \in L^2(\Omega) \subset L^{\frac{2}{3-p}}(\Omega)$ (Theorem 6.2.1), we are in position to apply the Riesz-Scheffé lemma [74, Lemma 2] to deduce that

$$\left\|I_n g_n - g\right\|_{L^2(\Omega)} \underset{n\to\infty}{\longrightarrow} 0 \text{ and } \left\|I_n P_n u^\star - u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)} \underset{n\to\infty}{\longrightarrow} 0.$$

Observe that for simple graphs, $I_n K_n$ is not an orthogonal projection of K (see (7.1.1)) and thus, the above argument proof used for g and $u^\star$ does not hold. We argue however using the fact that K is bounded, $|\Omega| < \infty$, and that $\forall n$ and $(x,y) \in \Omega^2$, $|I_n K_n(x,y)| \leq \left\|K\right\|_{L^\infty(\Omega)}$. We can thus invoke the dominated convergence theorem to get that

$$\left\|I_n K_n - K\right\|_{L^{\frac{2}{2-p}}(\Omega^2)} \underset{n\to\infty}{\longrightarrow} 0.$$

Passing to the limit in (6.3.7), we get the claim.

(ii) In the following C is any positive constant independent of n. Since $g \in L^\infty(\Omega) \cap \mathrm{Lip}(s, L^q(\Omega))$, $q \leq 2$, and we are dealing with a uniform partition of Ω ($|\Omega_i^{(n)}| = 1/n$, $\forall i \in [n]$), we get using inequality (2.3.3) that

$$\left\|I_n g_n - g\right\|_{L^2(\Omega)} \leq C n^{-s\min\{1,q/2\}} = C n^{-sq/2}. \tag{7.1.2}$$

By Lemma 6.3.3(i), we have $u^\star \in \mathrm{Lip}(sq/2, L^q(\Omega))$, and it follows from (2.3.2) and the fact that $q \geq 2/(3-p)$ that

$$\left\|I_n P_n u^\star - u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)} \leq \left\|I_n P_n u^\star - u^\star\right\|_{L^q(\Omega)} \leq C n^{-sq/2}. \tag{7.1.3}$$

Combining (7.1.2) and (7.1.3), we get

$$\left\|I_n g_n - g\right\|^2_{L^2(\Omega)} + \left\|I_n g_n - g\right\|_{L^2(\Omega)} + \left\|I_n P_n u^\star - u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)} \leq C\left(n^{-sq} + n^{-sq/2}\right) \leq C n^{-sq/2}. \tag{7.1.4}$$

It remains to bound $\|K - I_n K_n\|_{L^{\frac{2}{2-p}}(\Omega^2)}$. For that, consider the set of discrete cells $\Omega_{ij}^{(n)}$ overlying the boundary of the support of K

$$S(n) = \left\{(i,j) \in [n]^2 : \Omega_{ij}^{(n)} \cap \mathrm{bd}(\overline{\mathrm{supp}(K)}) \neq \emptyset\right\} \text{ and } C(n) = |S(n)|.$$

For any $\epsilon > 0$ and sufficiently large n, we have

$$C(n) \leq n^{\rho+\epsilon}.$$

It is easy to see that K and $I_n K_n$ coincide almost everywhere on cells $\Omega_{ij}^{(n)}$ such that $(i,j) \notin S(n)$. Thus, for any $\epsilon > 0$ and all sufficiently large n, we have

$$\|K - I_n K_n\|_{L^{\frac{2}{2-p}}(\Omega^2)}^{\frac{2}{2-p}} \leq C(n) n^{-2} \leq n^{-2(1-\frac{\rho+\epsilon}{2})}. \tag{7.1.5}$$

Inserting (7.1.4) and (7.1.5) into (6.3.7), the desired result follows.

(iii) For $p = 2$, let $\Omega_{S(n)} = \bigcup_{(i,j)\in S(n)} \Omega_{ij}^{(n)}$. We then have

$$\begin{aligned}
\|K - I_n K_n\|_{L^\infty(\Omega^2)} &\leq \|K - I_n K_n\|_{L^\infty(\Omega^2 \setminus \Omega_{S(n)})} + \|K - I_n K_n\|_{L^\infty(\Omega_{S(n)})} \\
&= \|K - I_n K_n\|_{L^\infty(\Omega_{S(n)})} \\
&\leq \max_{(i,j)\in S(n)} \sup_{(x,y)\in\Omega_{ij}^{(n)}} |K(x,y) - I_n K_n(x,y)| \leq n^{-2}.
\end{aligned}$$

□

7.2 Networks on weighted graphs

We now turn to the more general class of deterministic weighted graph sequences. The kernel K is used to assign weights to the edges of the graphs considered bellow, we allow only positive weights. These weights K_{nij} are obtained by averaging K over the cells in the partition $\mathcal{Q}_n$ following (6.3.2), and $I_n K_n$ is given by (6.3.3).

Proceeding similarly to the proof of statement (i) of Theorem 7.1.1, we conclude immediately that

$$\|I_n u_n^\star - u\|_{L^2(\Omega)} \underset{n\to+\infty}{\longrightarrow} 0.$$

We are rather interested now in quantifying the rate of convergence in (6.3.7). To do so, we need to add some regularity assumptions on the kernel K.

Theorem 7.2.1. *Let $p \in [1,2[$, and assume that $g \in L^\infty(\Omega) \cap \mathrm{Lip}(s, L^q(\Omega))$, with $s \in]0,1]$ and $q \in [2/(3-p), 2]$. Suppose moreover that $K(x,y) = J(|x-y|)$, $\forall (x,y) \in \Omega^2$, with J a nonnegative bounded measurable mapping on Ω. Let $u^\star$ and $u_n^\star$ be the unique minimizers of ($\mathcal{VP}_{\mathrm{nloc}}$) and ($\mathcal{VP}^d_{\mathrm{nloc}}$), respectively. Then, the following error bounds hold.*

(i) If $p \in [1,2[$ $K \in \mathrm{Lip}(s', L^{q'}(\Omega^2))$, $(s',q') \in]0,1] \times [1,+\infty[$, then

$$\|I_n u_n^\star - u^\star\|_{L^2(\Omega)}^2 \leq C n^{-\min\{sq/2, s', s'q'(1-p/2)\}}. \tag{7.2.1}$$

where C is a positive constant independent of n.
In particular, if $g \in L^\infty(\Omega) \cap \mathrm{BV}(\Omega)$ and $K \in L^\infty(\Omega^2) \cap \mathrm{BV}(\Omega^2)$, then

$$\|I_n u_n^\star - u\|_{L^2(\Omega)}^2 = O\left(n^{p/2-1}\right). \tag{7.2.2}$$

(ii) If $p \in [1,2]$ and $K \in \mathrm{Lip}(s', L^{q'}(\Omega^2))$, $(s',q') \in]0,1] \times [2/(2-p), +\infty]$, then

$$\|I_n u_n^\star - u^\star\|_{L^2(\Omega)}^2 \leq C n^{-\min\{sq/2, s'\}}. \tag{7.2.3}$$

where C is a positive constant independent of n.

In particular, if $g \in L^\infty(\Omega) \cap \mathrm{BV}(\Omega)$ then

$$\left\|I_n u_n^\star - u\right\|_{L^2(\Omega)}^2 = O\left(n^{-\min\{1/2, s'\}}\right). \tag{7.2.4}$$

Proof : In the following C is any positive constant independent of n. Under the setting of the theorem, for all cases, (7.1.4) still holds. It remains to bound $\left\|K - I_n K_n\right\|_{L^{\frac{2}{2-p}}(\Omega^2)}$. This is achieved using (2.3.3) for case (i) and (2.3.2) for case (ii), which yields

$$\begin{cases} \left\|K - I_n K_n\right\|_{L^{\frac{2}{2-p}}(\Omega^2)} \leq C n^{-s' \min\{1, q'(1-p/2)\}} & \text{for case (i)}, \\ \left\|K - I_n K_n\right\|_{L^{\frac{2}{2-p}}(\Omega^2)} \leq \left\|K - I_n K_n\right\|_{L^{q'}(\Omega^2)} \leq C n^{-s'} & \text{for case (ii)}. \end{cases} \tag{7.2.5}$$

Plugging (7.1.4) and (7.2.5) into (6.3.7), the bounds (7.2.1) and (7.2.3) follow.

We know that $\mathrm{BV}(\Omega) \subset \mathrm{Lip}(1/2, L^2(\Omega))$. Thus setting $s = s' = 1/2$ and $q = q' = 2$ in (7.2.1), and observing that $1 - p/2 \in [0, 1/2]$, the bound (7.2.2) follows. That of (7.2.4) is immediate. □

When $p = 1$ (i.e., nonlocal total variation), $g \in L^\infty(\Omega) \cap \mathrm{Lip}(s, L^2(\Omega))$ and K is a sufficiently smooth function, one can infer from Theorem 7.2.1 that the solution to the discrete problem ($\mathcal{VP}^d_{\mathrm{nloc}}$) converges to that of the continuous problem ($\mathcal{VP}_{\mathrm{nloc}}$) at the rate $O(n^{-s})$. This is to be compared to the slower convergence rate $O(n^{-s/(s+1)})$ established in [109, Theorem 4.1 and 5.1] for the discretization of the local ROF model.

7.3 Networks on random inhomogeneous graphs

We now turn to applying our bounds of Theorem 6.3.2 of Chapter 6 to networks on random inhomogeneous graphs. Recall the random inhomogeneous graph model defined in Section 2.1.5.

Following the same reasoning as that done for networks on random graphs for the evolution problem in Section 4.4.2.1, we assume first that the sequence $\mathbf{X}$ is deterministic. Capitalizing on this result, we will then deal with the totally random model (i.e.; generated by random nodes) in Section 7.3.2 by a simple marginalization argument combined with additional assumptions to get the convergence and quantify the corresponding rate.

7.3.1 Networks on graphs generated by deterministic nodes

As we have mentioned before, we shall denote $\mathbf{x} = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$ as we assume that the sequence of nodes is deterministic. Recall the parameter $\delta(n)$ defined in (4.4.3).

Next, we consider the discrete counterpart of ($\mathcal{VP}_{\mathrm{nloc}}$) on the graph G_n

$$\min_{u_n \in \mathbb{R}^n} \left\{ E_{n,\lambda}(u_n, g_n, K_n) \stackrel{\text{def}}{=} \frac{1}{2\lambda n}\left\|u_n - g_n\right\|_2^2 + \frac{1}{2pn^2} \sum_{i,j=1}^{n} \lambda_{ij}\left|u_{nj} - u_{ni}\right|^p \right\}, \tag{$\mathcal{VP}^d_{d,\mathrm{nloc}}$}$$

where

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

Theorem 7.3.1. *Suppose that $p \in [1, 2[$, $g \in L^2(\Omega)$ and K is a nonnegative measurable, symmetric and bounded mapping. Let $u^\star$ and $u_n^\star$ be the unique minimizers of ($\mathcal{VP}_{\mathrm{nloc}}$) and ($\mathcal{VP}^d_{d,\mathrm{nloc}}$), respectively. Let $p' = \frac{2}{2-p}$.*

(i) There exist positive constants C and C_1 that do not depend on n, such that for any $\beta > 0$

$$\left\|I_n u_n^\star - u^\star\right\|_{L^2(\Omega)}^2 \leq C\left(\left(\beta\frac{\log(n)}{n} + \frac{\max\left(q_n^{-(p'-1)}, q_n^{-p'/2}\right)}{n^{p'/2}}\right)^{1/p'} + \left\|g - I_n g_n\right\|_{L^2(\Omega)}^2 \right.$$
$$\left. + \left\|g - I_n g_n\right\|_{L^2(\Omega)} + \left\|K - I_n \hat{K}{}_n^{\mathbf{x}}\right\|_{L^{p'}(\Omega^2)} + \left\|u^\star - I_n P_n u^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)}\right), \tag{7.3.1}$$

with probability at least $1 - 2n^{-C_1 \min\left(q_n^{(2p'-1)}, q_n^{p'}\right)\beta}$.

(ii) Assume moreover that $g \in L^\infty(\Omega) \cap \mathrm{Lip}(s, L^q(\Omega))$, with $s \in]0,1]$ and $q \in [2/(3-p), 2]$, that $K(x,y) = J(|x-y|)$, $\forall (x,y) \in \Omega^2$, with J a nonnegative bounded measurable mapping on Ω, and $K \in \mathrm{Lip}(s', L^{q'}(\Omega^2))$, $(s', q') \in]0,1] \times [p', +\infty]$. Then there exist positive constants C and C_1 that do not depend on n, such that for any $\beta > 0$

$$\left\|I_n u_n^\star - u^\star\right\|_{L^2(\Omega)}^2 \leq C\left(\left(\beta\frac{\log(n)}{n} + \frac{\max\left(q_n^{-(p'-1)}, q_n^{-p'/2}\right)}{n^{p'/2}}\right)^{1/p'} + \delta(n)^{-\min(sq/2, s')}\right), \tag{7.3.2}$$

with probability at least $1 - 2n^{-C_1 \min\left(q_n^{(2p'-1)}, q_n^{p'}\right)\beta}$.

PROOF : In the following C is any positive constant independent of n.

(i) We start by arguing as in the proof of Theorem 6.3.2. Similarly to (6.3.9), we now have

$$\frac{1}{2\lambda}\left\|I_n u_n^\star - u^\star\right\|_{L^2(\Omega)}^2 \leq \left(E_\lambda(I_n u_n^\star, g, K) - E_{n,\lambda}(u_n^\star, g_n, \Lambda_n)\right) - \left(E_\lambda(u^\star, g, K) - E_{n,\lambda}(u_n^\star, g_n, \Lambda_n)\right). \tag{7.3.3}$$

The first term can be bounded similarly to (6.3.11)-(6.3.12) to get

$$\begin{aligned}
E_\lambda(I_n u_n^\star, g, K) - E_{n,\lambda}(u_n^\star, g_n, \Lambda_n) &\leq C\Bigg(\left\|I_n g_n - g\right\|_{L^2(\Omega)}^2 + \left\|I_n g_n - g\right\|_{L^2(\Omega)} \\
&\quad + \left|\int_{\Omega^2}\left(K(x,y) - I_n\Lambda_n(x,y)\right)\left|I_n u_n^\star(y) - I_n u_n^\star(x)\right|^p dxdy\right|\Bigg) \\
&\leq C\Bigg(\left\|I_n g_n - g\right\|_{L^2(\Omega)}^2 + \left\|I_n g_n - g\right\|_{L^2(\Omega)} \\
&\quad + \left|\int_{\Omega^2}\left(K(x,y) - I_n \hat{K}{}_n^{\mathbf{x}}(x,y)\right)\left|I_n u_n^\star(y) - I_n u_n^\star(x)\right|^p dxdy\right| \\
&\quad + \left|\int_{\Omega^2}\left(I_n \hat{K}{}_n^{\mathbf{x}}(x,y) - I_n\Lambda_n(x,y)\right)\left|I_n u_n^\star(y) - I_n u_n^\star(x)\right|^p dxdy\right|\Bigg).
\end{aligned} \tag{7.3.4}$$

The second term in (7.3.4) is $O\left(\left\|K - I_n \hat{K}{}_n^{\mathbf{x}}\right\|_{L^{p'}(\Omega^2)}\right)$, see (6.3.13). For the last term, we have using Jensen and Hölder inequalities,

$$\begin{aligned}
&\left|\int_{\Omega^2}\left(I_n\hat{K}{}^{\mathbf{x}}_n(x,y)-I_n\Lambda_n(x,y)\right)\left|I_nu^\star_n(y)-I_nu^\star_n(x)\right|^p dxdy\right| \\
&\le 2^{p-1}\left(\int_\Omega\left|\int_\Omega\left(I_n\hat{K}{}^{\mathbf{x}}_n(x,y)-I_n\Lambda_n(x,y)\right)dy\right|\left|I_nu^\star_n(x)\right|^p dx\right.\\
&\quad\left.+\int_\Omega\left|\int_\Omega\left(I_n\hat{K}{}^{\mathbf{x}}_n(x,y)-I_n\Lambda_n(x,y)\right)dx\right|\left|I_nu^\star_n(y)\right|^p dy\right) \\
&\le C\left(\left(\int_\Omega\left|\int_\Omega\left(I_n\hat{K}{}^{\mathbf{x}}_n(x,y)-I_n\Lambda_n(x,y)\right)dy\right|^{p'}dx\right)^{1/p'}\right.\\
&\quad\left.+\left(\int_\Omega\left|\int_\Omega\left(I_n\hat{K}{}^{\mathbf{x}}_n(x,y)-I_n\Lambda_n(x,y)\right)dx\right|^{p')}dy\right)^{1/p'}\right)\\
&=C\left(\left\|Z_n\right\|_{p',n}+\left\|W_n\right\|_{p',n}\right),
\end{aligned} \tag{7.3.5}$$

where

$$Z_{ni}\overset{\text{def}}{=}\frac{1}{n}\sum_{j=1}^n\left(\hat{K}{}^{\mathbf{x}}_{nij}-\lambda_{ij}\right)\ \text{ and }\ W_{nj}\overset{\text{def}}{=}\frac{1}{n}\sum_{i=1}^n\left(\hat{K}{}^{\mathbf{x}}_{nij}-\lambda_{ij}\right).$$

By virtue of Lemma 4.4.10, which is valid since $p'\in[2,+\infty[$, there exists a positive constant C_1, such that for any $\beta>0$

$$\mathbb{P}\left(\left\|Z_n\right\|_{p',n}\ge\varepsilon\right)\le n^{-C_1\min\left(q_n^{(2p'-1)},q_n^{p'}\right)\beta},$$

with

$$\varepsilon=\left(\beta\frac{\log(n)}{n}+\frac{\max\left(q_n^{-(p'-1)},q_n^{-p'/2}\right)}{n^{p'/2}}\right)^{1/p'}. \tag{7.3.6}$$

The same bound also holds for $\left\|W_n\right\|_{p',n}$. A union bound then leads to

$$\left\|Z_n\right\|_{p',n}+\left\|W_n\right\|_{p',n}\le 2\varepsilon \tag{7.3.7}$$

with probability at least $1-2n^{-C_1\min\left(q_n^{(2p'-1)},q_n^{p'}\right)\beta}$.

Let us now turn to the second term in (7.3.3). Using (6.3.6) and the fact that $u^\star_n$ is the unique minimizer of ($\mathcal{VP}^d_{d,\mathrm{nloc}}$), we have

$$\begin{aligned}
E_\lambda(I_nu^\star_n,I_ng_n,I_n\Lambda_n)-E_\lambda(u^\star,g,K)\le{}&\left(R_p(I_nP_nu^\star,K)-R_p(u^\star,K)\right)\\
&+\left(R_p(I_nP_nu^\star,I_nK_n)-R_p(I_nP_nu^\star,K)\right)\\
&+\left(R_p(I_nP_nu^\star,I_n\Lambda_n)-R_p(I_nP_nu^\star,I_nK_n)\right).
\end{aligned} \tag{7.3.8}$$

The first term is bounded as in (6.3.16), which yields

$$\left|R_p(I_nP_nu^\star,K)-R_p(u^\star,K)\right|\le C\left\|u^\star-I_nP_nu^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)}. \tag{7.3.9}$$

The second term follows from (6.3.17)

$$\left|R_p(I_nP_nu^\star,I_nK_n)-R_p(I_nP_nu^\star,K)\right|\le C\left\|K-I_nK_n\right\|_{L^{p'}(\Omega^2)}. \tag{7.3.10}$$

The last term is upper-bounded exactly as in (7.3.5) and (7.3.7).

Inserting (7.3.4), (7.3.5), (7.3.7), (7.3.8), (7.3.9) and (7.3.10) into (7.3.3), we get the claimed bound.

(ii) Insert (7.1.4) and (7.2.5) into (7.3.1) after replacing $1/n$ by $\delta(n)$.

□

7.3.2 Networks on graphs generated by random nodes

Let us turn now to the totally random model. The discrete counterpart of ($\mathcal{VP}_{\text{nloc}}$) on the totally random sequence of graphs $\{G_{q_n}\}_{n\in\mathbb{N}^*}$ is given by

$$\min_{u_n\in\mathbb{R}^n}\left\{E_{n,\lambda}(u_n,g_n,K_n)\stackrel{\text{def}}{=}\frac{1}{2\lambda n}\left\|u_n-g_n\right\|_2^2+\frac{1}{n^2}\sum_{i,j=1}^{n}\Upsilon_{ij}\left|u_{nj}-u_{ni}\right|^p\right\}, \qquad (\mathcal{VP}^d_{r,\text{nloc}})$$

where we recall that the random variables Υ_{ij} are the independent with $q_n\Upsilon_{ij}$ following the Bernoulli distribution with parameter $\mathbb{E}\left(q_n\hat{K}{}^{\mathbf{X}}_{nij}\right)$ defined above.

Observe that for the totally random model, $\delta(n)$ is a random variable. Thus, we have to to derive a bound on it. In Lemma 4.4.9, we shown that

$$\delta(n)\le t\frac{\log(n)}{n}, \qquad (7.3.11)$$

with probability at least $1-n^{-t}$, where $t\in]0,e[$.

Combining this bound with Theorem 7.3.1 (after conditioning and integrating) applied to the totally random sequence $\{G_{q_n}\}_{n\in\mathbb{N}^*}$, we get the following result.

Theorem 7.3.2. *Suppose that $p\in[1,2[$, $g\in L^2(\Omega)$ and K is a nonnegative measurable, symmetric and bounded mapping. Let $u^\star$ and $u_n^\star$ be the unique minimizers of ($\mathcal{VP}_{\text{nloc}}$) and ($\mathcal{VP}^d_{r,\text{nloc}}$), respectively. Let $p'=\frac{2}{2-p}$.*

(i) There exist positive constants C and C_1 that do not depend on n, such that for any $\beta>0$

$$\left\|I_nu_n^\star-u^\star\right\|^2_{L^2(\Omega)}\le C\left(\left(\beta\frac{\log(n)}{n}+\frac{\max\left(q_n^{-(p'-1)},q_n^{-p'/2}\right)}{n^{p'/2}}\right)^{1/p'}+\left\|g-I_ng_n\right\|^2_{L^2(\Omega)}\right.$$
$$\left.+\left\|g-I_ng_n\right\|_{L^2(\Omega)}+\left\|K-I_n\hat{K}{}^{\mathbf{X}}_n\right\|_{L^{p'}(\Omega^2)}+\left\|u^\star-I_nP_nu^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)}\right), \qquad (7.3.12)$$

with probability at least $1-2n^{-C_1\min\left(q_n^{(2p'-1)},q_n^{p'}\right)\beta}$.

(ii) Assume moreover that $g\in L^\infty(\Omega)\cap\mathrm{Lip}(s,L^q(\Omega))$, with $s\in]0,1]$ and $q\in[2/(3-p),2]$, that $K(x,y)=J(|x-y|)$, $\forall(x,y)\in\Omega^2$, with J a nonnegative bounded measurable mapping on Ω, that $K\in\mathrm{Lip}(s',L^{q'}(\Omega^2))$, $(s',q')\in]0,1]\times[p',+\infty]$ and $q_n\left\|K\right\|_{L^\infty(\Omega^2)}\le 1$. Then there exist positive constants C and C_1 that do not depend on n, such that for any $\beta>0$ and $t\in]0,e[$

$$\left\|I_nu_n^\star-u^\star\right\|^2_{L^2(\Omega)}\le C\left(\left(\beta\frac{\log(n)}{n}+\frac{\max\left(q_n^{-(p'-1)},q_n^{-p'/2}\right)}{n^{p'/2}}\right)^{1/p'}+\left(t\frac{\log(n)}{n}\right)^{\min(sq/2,s')}\right), \qquad (7.3.13)$$

with probability at least $1-\left(2n^{-C_1\min\left(q_n^{(2p'-1)},q_n^{p'}\right)\beta}+n^{-t}\right)$.

PROOF : Again, C will be any positive constant independent of n.

(i) Let

$$\varepsilon'=C\left(\left(\beta\frac{\log(n)}{n}+C\frac{\max\left(q_n^{-(p'-1)},q_n^{-p'/2}\right)}{n^{p'/2}}\right)^{1/p'}+\left\|g-I_ng_n\right\|^2_{L^2(\Omega)}+\left\|g-I_ng_n\right\|_{L^2(\Omega)}\right.$$
$$\left.+\left\|K-I_n\hat{K}{}^{\mathbf{X}}_n\right\|_{L^{p'}(\Omega^2)}+\left\|u^\star-I_nP_nu^\star\right\|_{L^{\frac{2}{3-p}}(\Omega)}\right).$$

Using (7.3.1), and independence of this bound from $\mathbf{x}$, we have

$$\begin{aligned}\mathbb{P}\left(\|I_n u_n^\star - u^\star\|_{L^2(\Omega)}^2 \geq \varepsilon'\right) &= \frac{1}{|\Omega|^n}\int_{\Omega^n} \mathbb{P}\left(\|I_n u_n^\star - u^\star\|_{L^2(\Omega)}^2 \geq \varepsilon' | \mathbf{X} = \mathbf{x}\right) d\mathbf{x} \\ &\leq \frac{1}{|\Omega|^n}\int_{\Omega^n} 2n^{-C_1 \min\left(q_n^{(2p'-1)}, q_n^{p'}\right)\beta} d\mathbf{x} \\ &= 2n^{-C_1 \min\left(q_n^{(2p'-1)}, q_n^{p'}\right)\beta}.\end{aligned}$$

(ii) Recall ε in (7.3.6) and $\kappa = C\left(t\frac{\log(n)}{n}\right)^{\min(sq/2, s')}$. Denote the event

$$A_1 : \left\{\|g - I_n g_n\|_{L^2(\Omega)}^2 + \|g - I_n g_n\|_{L^2(\Omega)} + \|K - I_n \hat{K}_n^{\mathbf{X}}\|_{L^{p'}(\Omega^2)} + \|u^\star - I_n P_n u^\star\|_{L^{\frac{2}{3-p}}(\Omega)} \leq \kappa\right\}.$$

In view of (7.1.4), (7.2.5) and (7.3.11), and that under our assumptions $\hat{K}_n^{\mathbf{X}} = K_n^{\mathbf{X}}$, we have

$$\mathbb{P}(A_1) \geq \mathbb{P}\left(\delta(n) \leq t\frac{\log(n)}{n}\right) \geq 1 - n^{-t}.$$

Let the event

$$A_2 : \left\{\|Z_n\|_{p',n} + \|W_n\|_{p',n} \leq 2\varepsilon\right\},$$

and denote A_i^c the complement of the event A_i. It then follows from (7.3.7) and the union bound that

$$\begin{aligned}\mathbb{P}\left(\|I_n u_n^\star - u^\star\|_{L^2(\Omega)}^2 \leq 2C\varepsilon + \kappa\right) &\geq \mathbb{P}(A_1 \cap A_2) = 1 - \mathbb{P}(A_1^c \cup A_2^c) \\ &\geq 1 - \sum_{i=1}^{2} \mathbb{P}(A_i^c) \geq 1 - \left(2n^{-C_1 \min\left(q_n^{(2p'-1)}, q_n^{p'}\right)\beta} + n^{-t}\right),\end{aligned}$$

which leads to the claimed result.

□

When $p = 1$ (i.e., nonlocal total variation), $g \in L^\infty(\Omega) \cap \mathrm{Lip}(s, L^2(\Omega))$ and K is a sufficiently smooth function, one can deduce from Theorem 7.3.2 that with high probability, the solution to the discrete problem ($\mathcal{VP}_{r,\mathrm{nloc}}^d$) converges to that of the continuous problem ($\mathcal{VP}_{\mathrm{nloc}}$) at the rate $O\left(\left(\frac{\log(n)}{n}\right)^{-\min(1/2, s)}\right)$. Compared to the deterministic graph model, there is overhead due to the randomness of the graph model which is captured in the rate and the extra-logarithmic factor.

7.4 Numerical results

In this section, we will apply the variational regularization problem ($\mathcal{VP}_{\mathrm{nloc}}^d$) to a few applications, and illustrate numerically our bounds.

7.4.1 Minimization algorithm

The algorithm we will describe in this subsection is valid for any $p \in [1, +\infty]$[1]. The minimization problem ($\mathcal{VP}_{\mathrm{nloc}}^d$) can be rewritten in the following form

$$\min_{u_n \in \mathbb{R}^n} \frac{1}{2}\|u_n - g_n\|_2^2 + \frac{\lambda_n}{p}\|\nabla_{K_n} u_n\|_p^p, \tag{7.4.1}$$

where $\lambda_n = \lambda/(2n)$, ∇_{K_n} is the (nonlocal) weighted gradient operator with weights K_{nij}, defined as

$$\nabla_{K_n} : \mathbb{R}^n \to \mathbb{R}^{n\times n}$$

[1] Obviously $\lim_{p\to+\infty} \frac{1}{p}\|\cdot\|_p^p = \iota_{\|u_n\|_\infty \leq 1}$.

$$u_n \mapsto V_n,\ V_{nij} = K_{nij}^{1/p}(u_{nj} - u_{ni}), \forall (i,j) \in [n]^2.$$

This is a linear operator whose adjoint, the (nonlocal) weighted divergence operator denoted div_{K_n}. It is easy to show that

$$\begin{aligned} \mathrm{div}_{K_n} :& \mathbb{R}^{n\times n} \to \mathbb{R}^n \\ & V_n \mapsto u_n,\ u_{ni} = \sum_{m=1}^{n} K_{nmi}^{1/p} V_{nmi} - \sum_{j=1}^{n} K_{nij}^{1/p} V_{nij}, \forall n \in [n]. \end{aligned}$$

Problem (7.4.1) can be easily solved using standard duality-based first-order algorithms. For this we follow [53].

By standard conjugacy calculus, the Fenchel-Rockafellar dual problem of (7.4.1) reads

$$\min_{V_n \in \mathbb{R}^{n\times n}} \frac{1}{2}\left\|g_n - \mathrm{div}_{K_n} V_n\right\|_2^2 + \frac{\lambda_n}{q}\left\|V_n/\lambda_n\right\|_q^q, \tag{7.4.2}$$

where q is the Hölder dual of p, i.e. $1/p + 1/q = 1$. One can show with standard arguments that the dual problem (7.4.2) has a convex compact set of minimizers for any $p \in [1, +\infty[$. Moreover, the unique solution $u_n^\star$ to the primal problem (7.4.1) can be recovered from any dual solution $V_n^\star$ as

$$u_n^\star = g_n - \mathrm{div}_{K_n} V_n^\star.$$

It remains now to solve (7.4.2). The latter can be solved with the (accelerated) FISTA iterative scheme [85, 17, 37] which reads in this case

$$\begin{aligned} W_n^k &= V_n^k + \frac{k-1}{k+b}(V_n^k - V_n^{k-1}) \\ V_n^{k+1} &= \mathrm{prox}_{\gamma\frac{\lambda_n}{q}\left\|\cdot/\lambda_n\right\|_q^q}\left(W_n^k + \gamma\nabla_{K_n}\left(g_n - \mathrm{div}_{K_n}(W_n^k)\right)\right) \\ u_n^{k+1} &= g_n - \mathrm{div}_{K_n} V_n^{k+1}, \end{aligned} \tag{7.4.3}$$

where $\gamma \in \left]0, \left(\sup_{\left\|u_n\right\|_2=1}\left\|\nabla_{K_n} u_n\right\|_2\right)^{-1}\right]$, $b > 2$, and we recall that $\mathrm{prox}_{\tau F}$ is the proximal mapping of the proper lsc convex function F with $\tau > 0$, i.e.,

$$\mathrm{prox}_{\tau F}(W) = \operatorname*{Argmin}_{V\in\mathbb{R}^{n\times n}} \frac{1}{2}\left\|V - W\right\|_2^2 + \tau F(V).$$

The convergence guarantees of scheme (7.4.3) are summarized in the following proposition.

Proposition 7.4.1. *The primal iterates u_n^k converge to $u_n^\star$, the unique minimizer of ($\mathcal{VP}^d_{\mathrm{nloc}}$), at the rate*

$$\left\|u_n^k - u_n^\star\right\|_2 = o(1/k).$$

Proof : Combine [53, Theorem 2] and [9, Theorem 1.1]. □

Let us turn to the computation of the proximal mapping $\mathrm{prox}_{\gamma\frac{\lambda_n}{q}\left\|\cdot/\lambda_n\right\|_q^q}$. Since $\left\|\cdot\right\|_q^q$ is separable, one has that

$$\mathrm{prox}_{\gamma\frac{\lambda_n}{q}\left\|\cdot/\lambda_n\right\|_q^q}(W) = \left(\mathrm{prox}_{\gamma\frac{\lambda_n}{q}|\cdot/\lambda_n|^q}(W_{ij})\right)_{(i,j)\in[n]^2}.$$

Moreover, as $|\cdot|^q$ is an even function on $\mathbb{R}$, $\mathrm{prox}_{\gamma\frac{\lambda_n}{q}|\cdot/\lambda_n|^q}$ is an odd mapping on $\mathbb{R}$, that is,

$$\mathrm{prox}_{\gamma\frac{\lambda_n}{q}|\cdot/\lambda_n|^q}(W_{ij}) = \mathrm{prox}_{\gamma\frac{\lambda_n}{q}|\cdot/\lambda_n|^q}(|W_{ij}|)\,\mathrm{sign}\,(W_{ij})\,.$$

In a nutshell, one has to compute $\mathrm{prox}_{\gamma\frac{\lambda_n}{q}|\cdot/\lambda_n|^q}(t)$ for $t \in \mathbb{R}^+$. We distinguish different situations depending on the value of q:

- $q = +\infty$ (i.e., $p = 1$): this case amounts to computing the orthogonal projector on $[-\lambda_n, \lambda_n]$, which reads

$$t \in \mathbb{R}^+ \mapsto \mathbf{P}_{[-\lambda_n,\lambda_n]}(t) = \min\big(t, \lambda_n\big).$$

- $q = 1$ (i.e., $p = +\infty$): this case corresponds to the well-known soft-thresholding operator, which is given by

$$t \in \mathbb{R}^+ \mapsto \mathrm{prox}_{\gamma|\cdot|}(t) = \max\big(t - \gamma, 0\big).$$

- $q = 2$ (i.e., $p = 2$): it is immediate to see that

$$\mathrm{prox}_{\gamma/(2\lambda_n)|\cdot|^2}(t) = \frac{t}{1 + \gamma/\lambda_n}.$$

- $q \in]1, +\infty[$: in this case, as $|\cdot|^q$ is differentiable, the proximal point $\mathrm{prox}_{\gamma\frac{\lambda_n}{q}|\cdot/\lambda_n|^q}(t)$ is the unique solution $\alpha^\star$ on $\mathbb{R}^+$ of the non-linear equation

$$\alpha - t + \gamma\alpha^{p-1}/\lambda_n = 0.$$

7.4.2 Experimental setup

We apply the scheme (7.4.3) to solve (7.4.1) in two applicative settings with nonlocal regularization on (weighted) graphs. The first one pertains to denoising of a function defined on a 2D point cloud, and the second one to signal denoising. In the first setting, the nodes of the graph are the points in the cloud and u_{ni} is the value of point/vertex index i. For signal denoising, each graph node correspond to a signal sample, and u_{ni} is the signal value at node/sample index i. We chose the nearest neighbour graph with the standard weighting kernel $e^{-|\mathbf{x}-\mathbf{y}|}$ when $|\mathbf{x} - \mathbf{y}| \leq \delta$ and 0 otherwise, where $\mathbf{x}$ and $\mathbf{y}$ are the 2D spatial coordinates of the points for the point cloud[2], and sample index for the signal case.

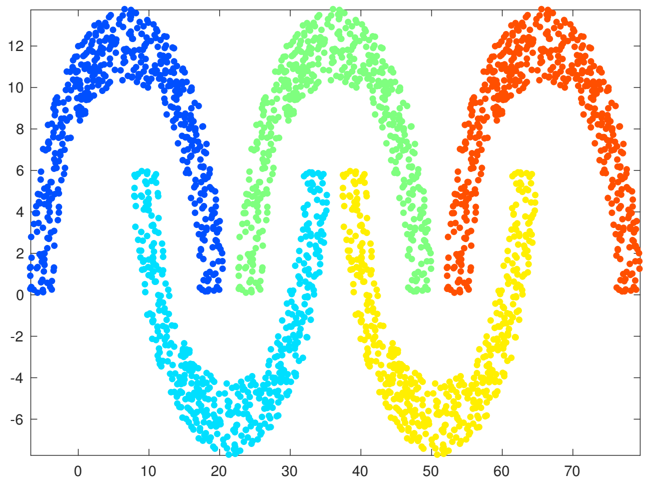


Figure 7.1: Original point cloud with $N = 2500$ points.

Application to point cloud denoising The original point cloud used in our numerical experiments is shown in Figure 7.1. It consists of $N = 2500$ points that do are not on a regular grid. The function on this point cloud, denoted u_N^0, is piecewise-constant taking 5 values (5 clusters) in [5]. A noisy observation g_N (see Figure 7.2(a)) is then generated by adding a white Gaussian noise noise of standard deviation 0.5 to u_N^0. Given the piecewise-constancy of u_N^0, we solved (7.4.1) with the natural choice $p = 1$. The result is shown in Figure 7.2(b). Figure 7.2(c) displays the evolution of $\left\|u_N^k - u_N^\star\right\|_2$ as a function of the iteration counter k, which confirms the theoretical rate $o(1/k)$ predicted above.

[2]For the 2D case, $(\mathbf{x}, \mathbf{y})$ are not to be confused with the "coordinates" (x, y) of the graphon on the continuum, though there is a bijection from one to another.

To illustrate our consistency results, $u^\star$ is needed while it is known in our case. Therefore, we argue as follows. We consider the continuous extension of $I_N u_N^\star$ as a reference and compute $\|u_n^\star - I_N u_N^\star\|_{L^2(\Omega)}$ for varying $n \ll N$, and the corresponding bound is expected to be dominated by that at n. Thus, for each value of $n \in [100, N/8]$, n nodes are drawn uniformly at random in $[N]$ and g_n is generated, which is a sampled version of g_N at those nodes. This is replicated 20 times. For each replication, we solve (7.4.1) with g_n and the same regularization parameter λ, and we compute the mean across the 20 replications of the squared-error $\|I_n u_n^\star - I_N u_N^\star\|_{L^2(\Omega)}^2$. The result is depicted in Figure 7.2(d). The gray-shaded area corresponds to one standard deviation of the error over the 20 replications. One indeed observe that the average error decreases at a rate consistent with the $O(n^{-1/2})$ predicted by our results (see discussion after Theorem 7.2.1 with $s = 1/2$).

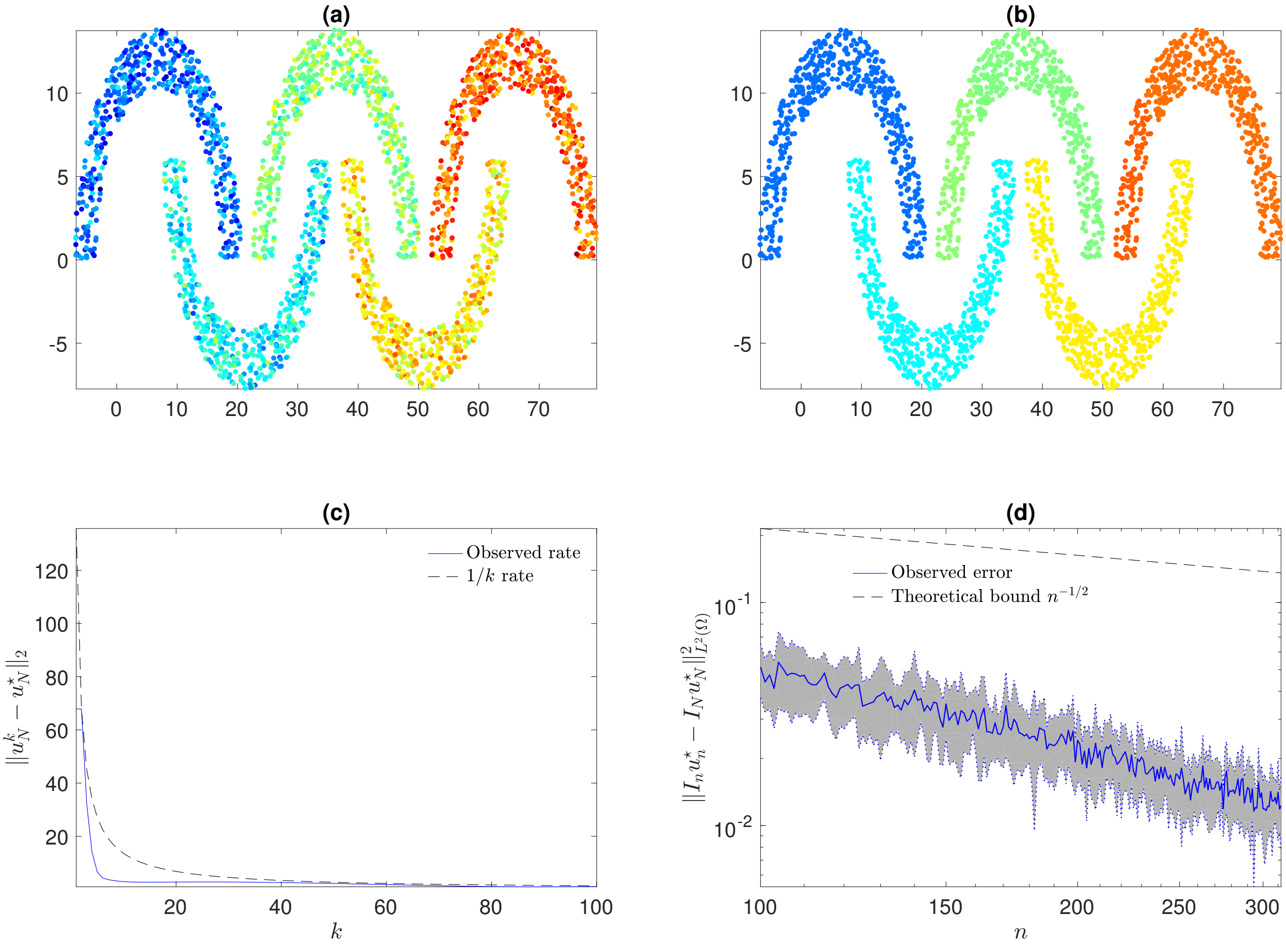


Figure 7.2: Results for point cloud denoising with $p = 1$. (a) Noisy point cloud. (b) Recovered point cloud by solving (7.4.1). (c) Primal convergence criterion $\|u_n^k - u_n^\star\|_2$ as a function of the iteration counter k. (d) Mean error $\|I_n u_n^\star - I_N u_N^\star\|_{L^2(\Omega)}^2$ across replications as a function of n.

Application to signal denoising In this experiment, we choose a piecewise-constant signal shown in Figure 7.3(a) for $N = 1000$ together with its noisy version g_N with additive white Gaussian noise of standard deviation 0.05. Figure 7.3(b) depicts the denoised signal $u_N^\star$ by solving (7.4.1) with $p = 1$ and hand-tuned λ. Figure 7.3(c) also confirms the $o(1/k)$ rate predicted above on $\|u_N^k - u_N^\star\|_2$.

We now illustrate the consistency bound result on a random sequence of graphs $\{G_{q_n}(n, K)\}_{n\in[100,N/4]}$ generated according to Definition 2.1.15 with $q_n = 1$. For each value of $n \in [100, N/4]$, n nodes are drawn uniformly at random in $[N]$, and g_n is generated, which is a sampled version of g_N at those nodes. n^2 independent Bernoulli variables λ_{ij} each with parameter K_{nij} are also generated. This is replicated 20 times. For each replication, we solve (7.4.1) with g_n and the same regularization parameter λ, and we compute the mean across the 20 replications of the squared-error $\left\|I_n u_n^\star - I_N u_N^\star\right\|_{L^2(\Omega)}^2$. The result is reported in Figure 7.3(d). The gray-shaded area indicates one standard deviation of the error over the 20 replications. Again, the average error decreases in agreement with the rate $O\left((\log(n)/n)^{1/2}\right)$ predicted by Theorem 7.3.2.

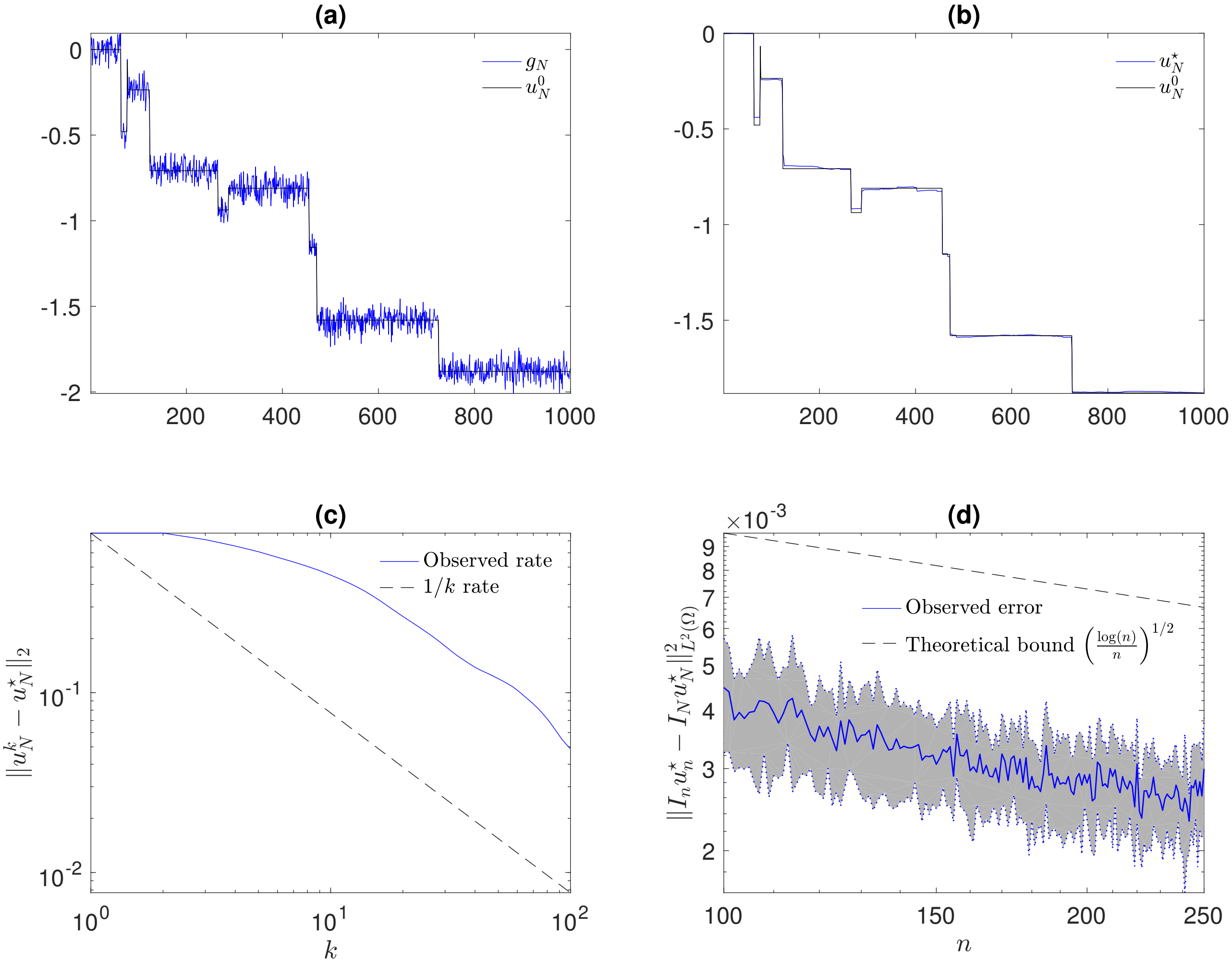


Figure 7.3: Results for signal denoising with $p = 1$. (a) Noisy and original signal. (b) Denoised and original signal foor $N = 1000$. (c) Primal convergence criterion $\left\|u_n^k - u_n^\star\right\|_2$ as a function of the iteration counter k. (d) Mean error $\left\|I_n u_n^\star - I_N u_N^\star\right\|_{L^2(\Omega)}^2$ as a function of n.

Chapter 8

Conclusion and Perspectives

This manuscript provides new results on consistency of evolution and variational nonlocal p-Laplacian problems on graphs along two main standpoints: general error bounds comparing the continuum problems and their discrete approximations on graphs and global convergence rates Our results provide a theoretical and insightful justification to the continuum limit for these nonlocal problems.

Take-away messages: several conclusions and take-away messages can be drawn from this work:

(i) our results reveal that without any extra regularity condition, starting from a bounded initial data, the Neumann nonlocal p-Laplacian evolution problem is consistent.

(ii) Our global nonasymptotic convergence rates for the evolution problem reveal that the approximation error depends on the regularity of the initial data and the *graphon*, and the latter encodes the geometry/structure of the underlying graphs. The more regular the initial data and the *graphon* are, the faster the convergence rate. Especially, for random inhomogeneous graphs, we exhibit different regimes for the convergence rate as a function of the problem parameters. In particular, the convergence rate shows a transition phenomenon at $p = 2$.

(iii) For the variational problem ($\mathcal{VP}_{\text{nloc}}$), we established a global (sharp) error estimate controlling the error between the unique minimizer of the continuum problem and that of the discrete one. The consistency of ($\mathcal{VP}_{\text{nloc}}$) is settled without any regularity assumption, just by supposing that the initial data is in $L^2(\Omega)$ and the kernel K is bounded.

(iv) Under very mild conditions on K and g, typically belonging to Lipschitz functional spaces, precise convergence rates were exhibited. These functional spaces allow to cover a large class of graphs (through K) and initial data g, including functions of bounded variation.

Our research program will not stop here, and many open questions are yet to be answered separately/commonly for both the evolution and variational problems.

8.1 The evolution problem

Other nonlocal operators: beyond the p-Laplacian The analysis developed in this thesis revolves mainly around the p-Laplacian operator. It would be interesting to study other nonlocal operators such as the (nonlocal) fractional Laplacian. i.e;

$$(-\boldsymbol{\Delta})^s u(x) = \frac{C(n,s)}{2} \int_{\mathbb{R}^n} \frac{u(x) - u(x+y) - u(x-y)}{|y|^{n+2s}} dy,$$

$C(n,s)$ is a positive constant, $s \in]0,1[$.

Other nonlocal evolution problems: beyond $(\mathcal{P}_{\text{nloc}})$ It would be also very interesting to extend our results to analyze the consistency of other nonlocal evolution problems such as the nonlocal Hamilton-Jacobi equation; see e.g., [12]. This is the subject of an ongoing work.

One can also think of studying consistency of numerical schemes beyond evolution problems. Typically, we think of the Dirichlet problem.

Other graph sequences Along the entire manuscript and particularly when dealing with networks on convergent graph sequences, we restricted ourselves to bounded *graphons* (we supposed that $K \in L^\infty(\Omega^2)$) and we dealt with a particular graph structure, that is dense graphs (deterministic and inhomogeneous random ones). However, practically many interesting graph models which arise in applications do not have this density property. In fact, our analysis does not accomodate for these graph models. The progress in this direction became possible with the theory of L^p-*graphons* used to define graph limits for sparse graphs of unbounded degree [28]. The goal will be to extend and adapt our arguments and results to this larger class of graphs, which includes directed and undirected, sparse and dense, random and deterministic graphs.

The limiting cases $p = 1$ **and** $p = +\infty$ Starting with the study of the well-posedness and going through the study of the consistency of $(\mathcal{P}_{\text{nloc}})$, excluding the values $p = 1$ and $p = \infty$ was crucial to get our results. Indeed, two main causes stand behind this restriction assumption:

(i) For $p = 1$ and $p = \infty$, the spaces $L^1(\Omega)$ and $L^\infty(\Omega)$ are not reflexive and thus don't have the Radon-Nikodym property. Due to this, one can not get the existence and uniqueness of a *strong* solution to $(\mathcal{P}_{\text{nloc}})$ for these values of p (see [7, Proposition A.35]). However, the authors of [7] have already established the well-posedness (existence and uniqueness of a *strong* solution) of the nonlocal total variation flow.i.e; $(\mathcal{P}_{\text{nloc}})$ with $p = 1$, and the kernel $K(x, y) = J(x - y)$, $x, y \in \Omega$ by taking the limit as $p \searrow 1$ of the solutions of the Neumann Cauchy problem with $p > 1$ that were studied in [7, Chapter 6]. To get the well-posedness of $(\mathcal{P}_{\text{nloc}})$ (for $p = 1$) one has to go a step further by adapting this result for the bivariate kernel K.

(ii) On the other hand, to get our estimate for the problem $(\mathcal{P}_{\text{nloc}})$, Lemma 2.2.16 was fundamental. However, a restrictive assumption was essential to get the desired result, that is to exclude the value $p = 1$. Hence, the error estimates we got are no longer valid for $p = 1$. It would be interesting to find a way to get around this difficulty and establish the consistency of $(\mathcal{P}_{\text{nloc}})$. For $p = \infty$, the definition of the operator $\boldsymbol{\Delta}_p^K$ becomes completely different, many challenges arise in addition to well-posedness.

Consistency of the normalized p**-Laplacian evolution problem** In Chapter 5, we dealt with the discrete in space Neumann evolution problem for the normalized p-Laplacian. We looked only at its well-posedness. It then appears natural to study the continuum counterpart of $(\mathcal{P}_{\text{nloc}}^{\text{Nor},d})$, its well-posedness and consistency of the dicsretization $(\mathcal{P}_{\text{nloc}}^{\text{Nor},d})$. In turn this will allow to study such problems on networks on convergent graph sequences and establish the corresponding convergence rates.

8.2 The variational problem

Inverse problems Beyond $(\mathcal{VP}_{\text{nloc}})$, we can try to extend our results to linear inverse problems where the data fidelity in $E_\lambda(u, g, K)$ is replaced by $\left\|g - \mathrm{A}u\right\|_{L^2(\Sigma)}^2$, and where A is a bounded linear operator from $L^2(\Omega)$ to $L^2(\Sigma)$.

Other nonlocal regularizations It would be interesting to study other nonlocal variational problems beyond the p-Laplacian. More precisely, it would be interesting to get deeper understanding of what are the essential properties of a nonlocal regularizer for our consistency results for instance to hold.

Beyond quadratic fidelity Here we focused on the quadratic data fidelity given its importance in practice for instance in imaging. It would be important to investigate what happens for other data fidelities, including those encountered in machine learning applications.

Convergence rates for $p > 2$ Our consistency results and convergence rates were only established for $p \in [1, 2]$. The extension beyond 2 faces a major obstacle materialized in bounding the term $\|u^\star - I_n P_n u^\star\|_{L^p(\Omega)}$. This is an important challenge.

Other graph sequences In the same vein as for evolution problems (see above), it would be important to extend our consistency results to other graph sequence models.

Solution structure and stability/recovery guarantees Understanding the recovery guarantees (structure of the solution, stability to noise, etc.) of nonlocal regularizers is much less understood than those of local ones (e.g. total variation). This is a whole research program that we believe is important to investigate.

List of Publications

In preparation

Y. Hafiene J.Fadili and A.Elmoataz, *Nonlocal p-Laplacian Variational Problems on Graphs.*

Preprints

Y. Hafiene J.Fadili C.Chesneau and A.Elmoataz, *The Continuum Limit of the Nonlocal p-Laplacian Evolution Problem on Random Inhomogeneous Graphs* submitted to IMA Journal of Numerical Analysis; *arXiv:1805.01754.*

Journal Papers

Y. Hafiene J.Fadili and A.Elmoataz, *Nonlocal p-Laplacian Evolution Problems Graphs*, SIAM Journal on Numerical Analysis, 56(2), 1064–1090, 2018.

Conference Proceedings

(1) Y. Hafiene J. Fadili and A. Elmoataz, *Le p-Laplacien non-local sur graphes: du discret au continu* Colloque sur le Traitement du Signal et des Images (**GRETSI**) ,Juan Les Pins, 2017 (**Oral**).

(2) Y. Hafiene J. Fadili and A. Elmoataz, *Nonlocal p-Laplacian Evolution Problems on Graphs* Colloque ORASIS, Journées Francophones des Jeunes Chercheurs en Vision par Ordinateur,Colleville-sur-Mer, 2017 (**Poster**).

List of Notations

General definitions

$\mathbb{R}$**:** the set of real numbers
$\mathbb{R}_+$**:** positive real numbers
$\bar{\mathbb{R}}$**:** $]-\infty,+\infty[\cup\{+\infty\}$, the extended real value
$\mathbb{N}$**:** set of non-negative integers
$\mathbb{N}^*$**:** set of positive integers
$\mathbb{R}^n, \mathbb{R}^m$**:** finite dimensional real Euclidean spaces

Spaces related

$\mathcal{H}$**:** real Hilbert space
$\mathcal{X}$**:** Banach space
$\Gamma_0(\mathcal{H})$**:** the set of proper convex and lower semicontinuous functions on $\mathcal{H}$
$L^p(\Omega)$**:** the Banach space of p-integrable functions on Ω, $p \in [1,+\infty]$
$C(0,T;\mathcal{X})$**:** the space of functions on $\mathcal{X} \times [0,T]$ which are continuous in the time variable

Sets related

$\iota_{\mathcal{S}}$**:** indicator function of a set $\mathcal{S}$
$\chi_{\mathcal{S}}$**:** charactetistic function of a set $\mathcal{S}$
$N_{\mathcal{S}}$**:** normal cone of a set $\mathcal{S}$
$\mathbf{P}_{\mathcal{S}}$**:** projection operator onto $\mathcal{S}$
$\mathrm{int}(\mathcal{S})$**:** interior of $\mathcal{S}$
$\mathrm{bd}(\mathcal{S})$**:** boundary of $\mathcal{S}$
$\overline{\mathcal{S}}$**:** closure of $\mathcal{S}$
$\mathrm{span}(\mathcal{S})$**:** smallest linear subspace that contains $\mathcal{S}$

Functions related

$\mathrm{dom}(F)$**:** domain of a function F
∇F**:** gradient of F
$\mathrm{prox}_{\gamma F}$**:** proximity operator of F with $\gamma > 0$
∂F**:** subdifferential of function F
$\mathrm{supp}(F)$**:** support of a function F

Operators

$\mathbf{Dom}(A)$**:** domain of the operator A
$\mathbf{R}(A)$**:** range of the operator A
J_A**:** resolvent of the operator A
$\mathbf{I}$**:** identity operator on a space to be understood from the context

Norms

$\|\cdot\|_{L^p(\Omega)}$**:** the norm of functions on $L^p(\Omega)$

$\|\cdot\|_p$**:** the p-norm of a vector in $\mathbb{R}^n$, $p \in [1, +\infty]$

$\|\cdot\|_{p,n}$**:** the normalized p-norm of a vector in $\mathbb{R}^n$, $p \in [1, +\infty]$

List of Figures

Bibliography

[1] Xavier Desquesnes Abderrahim Elmoataz and Mathieu Toutain. On the game p-Laplacian on weighted graphs with applications in image processing and data clustering. *Eurpean Journal of Apllied Mathematics*, 28:922–948, 2017.

[2] A. El Alaoui, X. Cheng, A. Ramdas, M. J. Wainwright, and M. I. Jordan. Asymptotic behavior of ℓ_p-based Laplacian regularization in semi-supervised learning. In *29th Annual Conference on Learning Theory*, pages 879–906, 2016.

[3] Ya.I. Alber, A.N. Iusem, and M.V. Solodov. On the projected subgradient method for nonsmooth convex optimization in a Hilbert space. *Mathematical Programming*, 81:23–35, 1998.

[4] Felipe Alvarez and Juan Peypouquet. Asymptotic equivalence and Kobayashi-type estimates for nonautonomous monotone operators in Banach spaces. *Discrete Contin. Dyn. Syst.*, 25(4):1109–1128, 2009.

[5] F. Andreu, J.M. Mazón, J.D. 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.

[6] Fuensanta Andreu, José Mazón, Julio Rossi, and Julián Toledo. The Neumann problem for nonlocal nonlinear diffusion equations. *Journal of Evolution Equations*, 8(1):189–215, March 2008.

[7] Fuensanta Andreu-Vaillo, José M. Mazón, Julio D. Rossi, and J. Julián Toledo-Melero. *Nonlocal diffusion problems*, volume 165 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2010.

[8] Buades Antoni, Coll Bartomeu, and Morel Jean Michel. On image denoising methods. *SIAM Multiscale Modeling and Simulation*, 4(2):490–530, 2005.

[9] H. Attouch and J. Peypouquet. The rate of convergence of Nesterov's accelerated Forward–Backward method is actually $o(k^{-2})$. *SIAM Journal on Optimization*, 26(3):1824–1834, 2016.

[10] Gilles Aubert and Pierre Kornprobst. *Mathematical Problems in Image Processing*, volume 147. Applied Mathematical Sciences, Springer-Verlag, 2002.

[11] Carrillo José Auel, Figalli Alessio, Mingione Giuseppe, Vázquez Juan Luisntonio, and Del Pino Man. *Nonlocal and Nonlinear Diffusions and Interactions: New Methods and Directions: Cetraro, Italy 2016*, volume 2186. Springer, 2017.

[12] Guy Barles, Pierre Cardaliaguet, Olivier Ley, and Aurélien Monteillet. Uniqueness results for nonlocal hamilton-jacobi equations. *Journal of Functional Analysis*, 257(5):1261–1287, 2008.

[13] John W. Barrett and W. B. Liu. Finite element approximation of the p-laplacian. *Mathematics of compuation*, 61(204):523–537, 1993.

[14] Peter W. Bates and Adam Chmaj. An integrodifferential model for phase transitions: Stationary solutions in higher space dimensions. *Journal of Statistical Physics*, 95(5):1119–1139, 1999.

[15] Peter W. Bates, Paul C. Fife, Xiaofeng Ren, and Xuefeng Wang. Traveling waves in a convolution model for phase transitions. *Archive for Rational Mechanics and Analysis*, 138(2):105–136, 1997.

[16] H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. Springer, 2011.

[17] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.

[18] Ph. Bénilan and M. G. Crandall. Completely accretive operators. *In Semigroup Theory and Evolution Equations (Delft, 1989), Lecture Notes in Applied and Pure Mathematics*, 135:41–75, New York, 1991.

[19] Gabriella Bognar. Numerical and Numerical and Analytic Investigation of Some Nonlinear Problems in Fluid Mechanics. *Computer and Simulation in Moderm Science, WSEAS Press*, 2:172–179, 2008.

[20] Béla Bollobás, Svante Janson, and Oliver Riordan. The phase transition in inhomogeneous random graphs. *Random Struct. Algorithms*, 31(1):3–122, August 2007.

[21] Béla Bollobás and Oliver Riordan. Metrics for sparse graphs. In Sophie Huczynska, James D. Mitchell, and Colva M. Editors Roney-Dougal, editors, *Surveys in Combinatorics 2009*, London Mathematical Society Lecture Note Series, pages 211–288. Cambridge University Press, Cambridge, 2009.

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

[23] 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(6):1801 – 1851, 2008.

[24] Christian Borgs, Jennifer Chayes, L. Lovasz, V. T. Sos, and K. Vesztergombi. Convergent sequences of dense graphs ii: Multiway cuts and statistical physics. *Annals of Mathematics*, 176(2):151–219, December 2012.

[25] Christian Borgs, Jennifer Chayes, László Lovász, Vera Sós, and Katalin Vesztergombi. Limits of randomly grown graph sequences. *European Journal of Combinatorics*, 32(7):985 – 999, 2011.

[26] Christian Borgs, Jennifer Chayes, László Lovász, Vera T. Sós, Balázs Szegedy, and Katalin Vesztergombi. Graph Limits and Parameter Testing. In *Proceedings of the Thirty-eighth Annual ACM Symposium on Theory of Computing*, pages 261–270. ACM, 2006.

[27] Christian Borgs, Jennifer Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Counting graph homomorphisms. In Martin Klazar, Jan Kratochvíl, Martin Loebl, Jiří Matoušek, Pavel Valtr, and Robin Thomas, editors, *Topics in Discrete Mathematics*, pages 315–371, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[28] Christian Borgs, Jennifer T Chayes, Henry Cohn, and Yufei Zhao. An l^p theory of sparse graph convergence i: limits, sparse random graph models, and power law distributions. *arXiv preprint arXiv:1401.2906*, 2014.

[29] Andrea Braides. *Γ-convergence for biginners.* Oxford Lecture Series in Mathematics and Its Applications Series, Oxford University Press, 2002.

[30] H. Brézis and A. Pazy. Convergence and approximation of semigroups of nonlinear operators in Banach spaces. *Journal of Functional Analysis*, 9(1):63 – 74, 1972.

[31] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, 4(2):490–530, 2005.

[32] Antoni Buades, Bartomeu Coll, Jean michel Morel, and Catalina Sbert. Non local demosaicing, 2007.

[33] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. Neighborhood filters and pde's. *Numerische Mathematik*, 105(1):1–34, 2006.

[34] Thomas Bühler and Matthias Hein. Spectral clustering based on the graph p-Laplacian. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pages 81–88, New York, NY, USA, 2009. ACM.

[35] C. Carrillo and P. Fife. Spatial effects in discrete generation population models. *Journal of Mathematical Biology*, 50(2):161–188, 2005.

[36] Francine Catté, Pierre-Louis Lions, Jean-Michel Morel, and Tomeu Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical analysis*, 29(1):182–193, 1992.

[37] A. Chambolle and C. Dossal. On the convergence of the iterates of the "fast iterative shrinkage/thresholding algorithm". *Journal of Optimization Theory and Applications*, 166(3):968–982, 2015.

[38] Tony F. Chan and Jianhong Jackie Shen. *Image Processing and Analysis.* SIAM, 2005.

[39] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the National Academy of Sciences*, 102(21):7426–7431, 2005.

[40] Michael G Crandall and Thomas M Liggett. Generation of semi-groups of nonlinear transformations on general banach spaces. *American Journal of Mathematics*, 93(2):265–298, 1971.

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

[42] J. B. Diaz and R. Výborný. On mean value theorems for strongly continuous vector valued functions. *Contrib. Differential Equations*, 3:107–118, 1964.

[43] Reinhard Diestel. *Graph Theory*, volume 173. Springer-Verlag, New York, 2000.

[44] Pavel Drábek. The p-Laplacian–mascot of nonlinear analysis. *Acta Math. Univ. Comenianae*, 76(1):85–98, 2007.

[45] Abderrahim Elmoataz, Xavier Desquesnes, Zakaria Lakhdari, and Olivier Lézoray. Nonlocal infinity Laplacian equation on graphs with applications in image processing and machine learning. *Mathematics and Computers in Simulation*, 102:153–163, 2014.

[46] Abderrahim Elmoataz, Xavier Desquesnes, and Olivier 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.

[47] Abderrahim Elmoataz, Olivier Lezoray, and Sebastien Bougleux. Nonlocal Discrete Regularization on Weighted Graphs: A Framework for Image and Manifold Processing. *IEEE Transactions on Image Processing*, 17(7), 2008.

[48] Abderrahim Elmoataz, Olivier Lezoray, Sébastien Bougleux, and Vinh Thong Ta. Unifying local and nonlocal processing with partial difference operators on weighted graphs. In *International Workshop on Local and Non-Local Approximation in Image Processing*, pages 11–26, Switzerland, 2008.

[49] Abderrahim Elmoataz, Matthieu Toutain, and Daniel 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.

[50] 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.

[51] Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, 2010.

[52] Gabriele Facciolo, Pablo Arias, Vicent Caselles, and Guillermo Sapiro. Exemplar-based interpolation of sparsely sampled images. In Daniel Cremers, Yuri Boykov, Andrew Blake, and Frank R. Schmidt, editors, *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 331–344, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[53] M. J. Fadili and G. Peyré. Total variation projection with first order schemes. *IEEE Transactions on Image Processing*, 20(3):657–669, 2010.

[54] Driss Meskine Fahd Karami and Khadija Sadik. A new nonlocal model for restoration of textured images. *(submitted)*, 2018.

[55] Kenneth J. Falconer. *Fractal geometry : mathematical foundations and applications*. J. Wiley & sons, Chichester, New York, Weinheim, 1990. Réimpr. en 1993, 1995, 1997, 1999, 2000.

[56] Paul C. Fife and Xuefeng Wang. A convolution model for interfacial motion: the generation and propagation of internal layers in higher space dimensions. *Adv. Differential Equations*, 3(1):85–110, 1998.

[57] P.C. Fife. Some nonclassical trends in parabolic and parabolic-like evolutions. In B. Fiedler, editor, *Trends in Nonlinear Analysis*. Springer-Verlag, 2002.

[58] Alessio Figalli, Ireneo Peral, and Enrico Valdinoci. *Partial Differential Equations and Geometric Measure Theory*. Springer International Publishing, 2018.

[59] Nicolás García Trillos and Dejan Slepčev. Continuum Limit of Total Variation on Point Clouds. *Archive for Rational Mechanics and Analysis*, 220(1):193–241, April 2016.

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

[61] Guy Gilboa, Jerome Darbon, Stanley Osher, and Tony Chan. Nonlocal convex functionals for image regularization. *UCLA CAM-report*, pages 06–57, 2006.

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

[63] Guy Gilboa and Stanley Osher. Nonlocal operators with applications to image processing. *SIAM J. Multiscale Modeling and Simulation*, 7(3):1005–1028, 2008.

[64] M. Gunzburger and R.B. Lehoucq. A nonlocal vector calculus with applications to nonlocal boundary value problems. *SIAM J. Multiscale Modeling and Simulation*, 8:1581–1620, 2010.

[65] Yosra Hafiene, Jalal Fadili, Christophe Chesneau, and Abderrahim Elmoataz. Continuum limit of the nonlocal p-Laplacian evolution problem on random inhomogeneous graphs. *arXiv:1805.01754*, 2018.

[66] Yosra Hafiene, Jalal Fadili, and Abderrahim Elmoataz. Nonlocal p-Laplacian evolution problems on graphs. *SIAM Journal on Numerical Analysis*, 56(2):1064–1090, 2018.

[67] Yosra Hafiene, Jalal Fadili, and Abderrahim Elmoataz. Nonlocal p-Laplacian variational problems on graphs. in preparation, 2018.

[68] B. Hinds and P. Radu. Dirichlet's principle and well-posedness of solutions for a nonlocal Laplacian system. *Applied Mathematics and Computation*, 219:1411–1419, 11 2012.

[69] Rustam Ibragimov and Shaturgun 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.

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

[71] T. Kato. Nonlinear semigroups and evolution equations. *Journal of the Mathematical Society of Japan*, 19:508–520, 1967.

[72] Bernd Kawohl. Variations on the p-Laplacian. *Nonlinear Elliptic Partial Differential Equations. Contemporary Mathmatics*, 540:35–46, 2011.

[73] Yoshikazu Koabayashi. Difference approximation of Cauchy problems for quasi-dissipative operators and generation of nonlinear semigroups. *J. Math. Soc. Japan*, 27(4):640–665, 10 1975.

[74] Norbert Kusolitsch. Why the theorem of Scheffé should be rather called a theorem of Riesz. *Periodica Mathematica Hungarica*, 61(1):225–229, 2010.

[75] Carolina A.Mosquera Leandro M.Del Pezzo and Julio D. Rossi. Existence, uniqueness and decay rates for evolution equations on trees. *Arxiv:1309.7634v2*, 2014.

[76] Jong-Sen Lee. Digital image smoothing and the sigma filter. *Computer Vision, Graphics, and Image Processing*, 24(2):255 – 269, 1983.

[77] Peter Lindqvist. Notes on the p-Laplace equation. University of Jyväskylä, 2006.

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

[79] László Lovász and Vera T. Sós. Generalized quasirandom graphs. *Journal of Combinatorial Theory, Series B*, 98(1):146 – 163, 2008.

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

[81] François Lozes, Abderrahim Elmoataz, and Olivier Lézoray. Partial difference operators on weighted graphs for image processing on surfaces and point clouds. *IEEE Transactions on Image Processing*, 23(9):3896–3909, 2014.

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

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

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

[85] Y. Nesterov. A method for solving the convex programming problem with convergence rate $O(1/k^2)$. *Dokl. Akad. Nauk SSSR*, 269(3):543–547, 1983.

[86] Y. Nesterov. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer, 2004.

[87] Adam M Oberman. Finite difference methods for the infinity laplace and p-laplace equations. *Journal of Computational and Applied Mathematics*, 254:65–80, 2013.

[88] Stanley Osher and Nikos Paragios. *Geometric Level Set Methods in Imaging, Vision, and Graphics*. Springer Verlag, 2003.

[89] Etienne Pardoux. Cours intégration et probabilité. Lecture notes (Aix-Marseille Universtité), November 2009.

[90] Yuval Peres and Scott Sheffield. Tug-of-War with noise: A game-theoretic view of the p-Laplacian. *Duke Mathematical Journal*, 145(1):91–120, 2008.

[91] Mayte Pérez-LLanos and Julio D. Rossi. Numerical approximations for a nonlocal evolution equation. *SIAM J. on Numerical Analysis*, 49(5):2103–2123, 2011.

[92] G. Peyré. Image processing with nonlocal spectral bases. *SIAM J. Multiscale Modeling and Simulation*, 7(2):703–730, 2008.

[93] Gabriel Peyré. Résolution numérique d'équations intégrales exemple de la radiosité. Lecture notes, October 2001.

[94] Gabriel Peyré, Sébastien Bougleux, and Laurent Cohen. Non-local regularization of inverse problems. *Inverse Problems and Imaging*, 5(1930):511, 2011.

[95] Simeon Reich. Weak convergence theorems for nonexpansive mappings in Banach spaces. *Journal of Mathematical Analysis and Applications*, 67(2):274 – 276, 1979.

[96] Simeon Reich and David Shoiykhet. *Nonlinear semigroups, fixed points, and geometry of domains in Banach spaces*. Imperial College Press, 2005.

[97] Rolf-Dieter Reiss. *Approximate Distributions of Order Statistics with Applications to Nonparametric Statistics.* Springer-Verlag, New York, 1989.

[98] Matthew Rudd. Statistical exponential formulas for homogeneous diffusion. *arXiv preprint arXiv :1403.1853*, 2014.

[99] Dejan Slepĉev and Matthew Thorpe. Analysis of p-Laplacian Regularization in Semi-Supervised Learning. *arXiv:1707.06213v2*, October 2017.

[100] Stephen M. Smith and J. Michael Brady. Susan—a new approach to low level image processing. *International Journal of Computer Vision*, 23(1):45–78, May 1997.

[101] A. Spira, R. Kimmel, and N. Sochen. A short- time beltrami kernel for smoothing images and manifolds. *IEEE Transactions on Image Processing*, 16(6):1628–1636, June 2007.

[102] Karthik Sridharan. A Gentle Introduction to Concentration Inequalities.

[103] Stanley Osher Stefan Kindermann and Peter W. Jones. Deblurring and Denoising of Images by Nonlocal Functionals. *SIAM Multiscale Modeling and Simulations*, 4:25, 2006.

[104] Arthur Szlam, Mauro Maggioni, and Ronald Coifman. A general framework for adaptive regularization based on diffusion processes on graphs. *Journal of Machine Learning Research*, 19:1711–1739, 2008.

[105] Vinh Thong Ta, Abderrahim Elmoataz, and Olivier Lezoray. Nonlocal PDEs-Based Morphology on Weighted Graphs for Image and Data Processing. *IEEE Transactions on Image Processing*, 20(6), 2011.

[106] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pages 839–846, Jan 1998.

[107] Mathieu Toutain. *EdP géométriques pour le traitement et la classification de données sur graphes.* PhD thesis, Université Caen Normandie, 2015.

[108] Luminita A Vese and Stanley J Osher. Modeling textures with total variation minimization and oscillating patterns in image processing. *Journal of scientific computing*, 19(1-3):553–572, 2003.

[109] Jingyue Wang and Bradley J. Lucier. Error bounds for finite-difference methods for rudin–osher–fatemi image smoothing error Bounds for Finite-Difference Methods for Rudin-Osher-Fatemi Image Smoothing. *SIAM Journal on Numerical Analysis*, 49, 2011.

[110] Xuefeng Wang. Metastability and stability of patterns in a convolution model for phase transitions. *Journal of Differential Equations*, 183(2):434 – 461, 2002.

[111] Hong-Kun Xu. Inequalities in Banach spaces with applications. *Nonlinear Analysis: Theory, Methods & Applications*, 16(12):1127–1138, 1991.

[112] Z. Yang and M. Jacob. Nonlocal regularization of inverse problems: A unified variational framework. *IEEE Transactions on Image Processing*, 22(8):3192–3203, Aug 2013.

[113] L. P. Yaroslavsky. *Digital Picture Processing — an Introduction.* Springer, Berlin, 1985.

[114] W. H. Young and G. C. Young. On derivatives and the theorem of the mean. *Quart. J. Pure Appl. Math.*, 40:1–26, 1909.

[115] Scott Sheffield Yuval Peres, Oded Schramm and David B. Wilson. Tug-of-War and the infinity Laplacian. *Journal of the American Mathematical Society*, 22:167–210, 2009.