Discrete-to-Continuum Rates of Convergence for Nonlocal p-Laplacian Evolution Problems

Adrien Weihs[1], Jalal Fadili[2], and Matthew Thorpe[3]

[1]Department of Mathematics,
University of California Los Angeles,
Los Angeles, CA 90095, USA.

[2]Normandie Université, ENSICAEN, UNICAEN, CNRS, GREYC,
Caen, France

[3]Department of Statistics,
University of Warwick,
Coventry, CV4 7AL, UK.

Abstract

Higher-order regularization problem formulations are popular frameworks used in machine learning, inverse problems and image/signal processing. In this paper, we consider the computational problem of finding the minimizer of the Sobolev $\mathrm{W}^{1,p}$ semi-norm with a data-fidelity term. We propose a discretization procedure and prove convergence rates between our numerical solution and the target function. Our approach consists of discretizing an appropriate gradient flow problem in space and time. The space discretization is a nonlocal approximation of the p-Laplacian operator and our rates directly depend on the localization parameter ε_n and the time mesh-size τ_n. We precisely characterize the asymptotic behaviour of ε_n and τ_n in order to ensure convergence to the considered minimizer. Finally, we apply our results to the setting of random graph models.

Keywords and phrases. convergence rates, evolution problems, nonlocal variational problems, p-Laplacian, random graph models, asymptotic consistency, regularization
Mathematics Subject Classification. 65N12, 34G20, 65M06, 35R02

1 Introduction

In machine learning, inverse problems and image/signal processing, one is often confronted with finding smooth solutions to regression problems. This leads to the introduction of regularization terms in the problem formulation (see for example [42], [9], [35]). In this paper, we investigate the computation of the solution to the following regularization problem:

$$u_\infty \in \operatorname*{argmin}_{v\in \mathrm{W}^{1,p}(\Omega)} \mathcal{F}(v) := \frac{\mu}{p}\|\nabla v\|^p_{\mathrm{L}^p(\Omega)} + \frac{1}{2}\|\mathcal{A}v-\ell\|^2_{\mathrm{L}^2(\Omega)} \tag{1}$$

for some $\mu > 0$, linear operator $\mathcal{A}$, data function $\ell \in \mathrm{L}^2(\Omega)$, $\Omega \subseteq \mathbb{R}^d$ and $p \geq 1$. The first term on the right-hand side of the latter enforces some regularization upon the functions v while the second term is a data-fidelity term. For $\mathcal{A} = \mathrm{Id}$, this is also called a fully supervised machine learning problem or a denoising problem in image processing. More precisely, in this paper, we prove convergence rates between our numerical discrete solution and u_∞. This type of regularization has been considered for example in [24], [51], [34] and [75]; for examples involving nonlocal versions of the $\mathrm{W}^{1,p}$ semi-norm we refer to [3], [43] and references therein; for other higher-order regularization problems we refer for example to [67] and references therein.

The aim of this paper can be seen as the converse to a large part of the discrete-to-continuum work in recent years: indeed, in [61], [31], [39], [72] and [73] for example, one starts with a discrete problem and analyses the large data limit of the latter – a point of interest in machine learning settings. Similarly to [33], [44] and [43], we start from the continuum and explain how the appropriate discretization should be designed as is usually the case in numerical analysis. In particular, this means that in the former case, emphasis is partly placed on having as few constraints as possible on different parameters as the latter are usually inferred from the data at hand: in our case, as we are choosing all the parameters, this concern is less relevant.

The approach we choose in this paper is to discretize the gradient flow (7) associated to (1) which contains the p-Laplacian operator $\Delta_p u = \mathrm{div}(|\nabla u|^{p-2}\nabla u)$ (see [49] and references therein). We note two important consequences from this reformulation of the task at hand: adding a time dependence to our problem will allow us to leverage the theory of semigroups in Banach spaces (see Section 2.2) in order to derive rates; the main problem is to obtain convergence rates between the p-Laplacian operator and its discrete counterpart.

The plan for the discretization follows loosely the strategy in [39] and [61]: after choosing an appropriate kernel K, we start by introducing a nonlocal version of the p-Laplacian operator in the continuum and then discretize the latter. The former is inspired by [13], [56], [48] where finite-difference approximations of Sobolev norms are discussed in the continuum, while the latter step allows us to use existing results on convergence results in the nonlocal setting as in [33], [44], [43].

Passing from the nonlocal setting to the local one requires a kernel that is appropriately scaled. For some length-scale $\varepsilon_n \to 0$, it is shown for example in [61], [18] and [2] that, ignoring regularity assumptions,

$$\frac{1}{\varepsilon_n^{d+p}} \int K\left(\frac{|y-x|}{\varepsilon_n}\right) |u(y)-u(x)|^{p-2}(u(y)-u(x))\,\mathrm{d}y \to \Delta_p u(x).$$

The finite-difference structure of the above nonlocal approximation is essential for obtaining rates in the continuum: indeed, as in [18], for smooth enough functions, we pass from finite-differences to derivatives and the rates follow from a conceptually basic Taylor expansion.

The discrete-to-continuum step requires more subtle techniques. In fact, one needs a way to compare functions $\bar{u}_n \in \mathbb{R}^{D_n}$ and $u : \Omega \mapsto \mathbb{R}$. We discuss various alternatives in Section 1.2. In this paper, we choose to partition our space Ω into D_n cells and elements of $\mathbb{R}^{D_n}$ are injected through the operator $\mathcal{I} : \mathbb{R}^{D_n} \to \mathrm{L}^1(\Omega)$ in the continuum through piecewise constant approximations while continuum functions are projected through the operator $\mathcal{P} : \mathrm{L}^1(\Omega) \to \mathbb{R}^{D_n}$ onto our cells by averaging on each cell. Using this method, establishing rates between a continuum function and its injected discrete approximation relies on tools from approximation theory, depending on the partition chosen and the regularity assumption of the continuum function. This topic is discussed in greater detail in Section 2.4. The other central tool for the discrete-to-continuum rates, where the discretization is both in space and time, is the semigroup structure of our solutions to the nonlocal gradient flows. Indeed, relying on some favourable properties of the nonlocal p-Laplacian, we are able to deduce strong contraction properties as discussed in Section 2.2.

Combining the discrete-to-continuum rates in the nonlocal setting to the nonlocal-to-local rates in the continuum, we obtain in Corollary 3.10 that for $p \geq 3$ and $\mathcal{A} = \mathrm{Id}$,

$$\begin{aligned}\|\mathcal{I}_n \bar{u}_n^N - u_\infty\|_{\mathrm{L}^2} \leq C\Bigg(&\varepsilon_n \log(\varepsilon_n^{-\kappa}) + \varepsilon_n^{\kappa/4}(\mathcal{F}(u_0) - \mathcal{F}(u_\infty))^{1/2} \\ &+ \varepsilon_n^{-\kappa}\left[n^{-\alpha_1} + n^{-\alpha_2} + \frac{\log(\varepsilon_n^{-\kappa})^{(p-1)}}{\varepsilon_n^{d+p+\alpha_3} n^{\alpha_3}} + \tau_n \frac{\log(\varepsilon_n^{-\kappa})^{2p-3}}{\varepsilon_n^{2(d+p)}}\right]\Bigg)\end{aligned} \tag{2}$$

where u_∞ is the solution to (1), $\mathcal{I}_n \bar{u}_n^N$ is the injected discretized solution on a partition indexed by n in space and N in time, τ_n is the maximum step-size of the forward Euler time-discretization (we pick $0 = t^0 < t^1 < \cdots < t^N \approx \log(\varepsilon_n^{-\kappa})$ so that $N \geq \log(\varepsilon_n^{-\kappa})/\tau_n$), $C > 0$ is a constant that is independent of n, $\kappa > 0$ and $\alpha_i > 0$ are chosen numerical constants depending on the regularity of the initial condition of the gradient flow problem (7) u_0, the kernel K and the data ℓ.

First, we note that each term in the right-hand side of (2) corresponds to a specific error source, namely (from left to right) the continuum nonlocal-to-local approximation, the gradient flow convergence, the discrete-to-continuum approximation of u_0, $\mathcal{A}^*\ell$ and K as well as a general discretization error. Furthermore, as the choices of α_i and κ are left to the practitioner, the rates can be enhanced upon implementation.

Second, while we give the precise statement of this result in Section 3.3, we note that the right-hand side of (2) tending to 0 involves finding the right interplay between $\tau_n \to 0$ and $\varepsilon_n \to 0$: this is similar to Courant-Friedrichs-Lewy (CFL) conditions [25] and [29]. In particular, while the classical CFL conditions correspond to $\tau_n \ll \varepsilon_n^2$ for the heat equation with the forward Euler time-discretization, we will show in Theorem 3.7 that our requirement is roughly $\tau_n \ll \varepsilon_n^{2(d+p)}$. We also find that ε_n admits a lower bound and this is analogous to results in semi-supervised learning discrete-to-continuum phenomena in [61], [31] and [73].

Third, we note that (2) does not cover the linear case of $p = 2$. This is due to a technicality and indeed, our well-posedness results both in the nonlocal case (Theorem 3.6) and the local case (Theorem 4.9) only require $p \geq 2$. The $p \geq 2$ assumption is particularly helpful since it allows one to have $\mathrm{L}^{p/(p-1)}(\Omega) \subseteq \mathrm{L}^p(\Omega)$ and $\mathrm{W}^{1,p}(\Omega) \subseteq \mathrm{L}^2(\Omega)$. When establishing continuum rates in Theorem 4.10 however, similarly to what is presented in [18], we will have to consider a third-order Taylor expansion of the function $x \mapsto |x|^{p-2}x$, hence requiring $p \geq 3$. We nevertheless remark that the choice of p is left to the practitioner and, as explained in Remark 3.2, should be made in accordance with the dimension of Ω in order to have a small ratio d/p.

Fourth, we point out that the condition $\mathcal{A} = \mathrm{Id}$ is not constraining. Indeed, our results in Theorem 3.7 allow us to obtain similar rates for more general $\mathcal{A}$. We refer to the proof of the above-mentioned theorem and Remark 4.16 for more details.

Lastly, coming back to a more data-centric approach, we apply the above-mentioned results to a random graph model inspired by graphons. The latter are defined for points in $\Omega = [0, 1]$, appear in several applications and we refer to Section 2.5 for further details. We obtain results equivalent to the ones displayed above in Theorem 3.13: the main difference is an additional term accounting for the discrete random-to-deterministic approximation error.

1.1 Contributions

Our main contributions in the paper which we discuss in greater detail in Section 3 are:

1. A rigourous proof of the well-posedness of the nonlocal continuum gradient flow defined in (5);
2. The establishment of rates of convergence between the discrete gradient flow and u_∞ through a precise characterization of the discretization parameters in Corollary 3.10;
3. An application to random graph models in Section 4.3.

1.2 Related works

p-Laplacian operator approximations. Approximating the p-Laplacian operator has been explored in [6], [41], [71], [5] but always in the context of finite elements. This simplifies some part of the analysis as described in Remark 3.1 but has the disadvantage of being difficult to apply in higher dimensions. In contrast to these works references, our paper deals with nonlocal discrete approximations. While this will require some increased regularity considerations (see Remark 3.1), establishing the rate of convergence is conceptually simple as explained previously.

In the continuum setting, nonlocal-to-local convergence of gradient flows involving the p-Laplacian operator is shown in [2]. This relies heavily on [13] and is a consistency result, without rates. Some rates are established between the nonlocal and local operators in [18].

Much of the discrete-to-continuum work in recent years has dealt with similar problems in many ways. The closest results are to be found in [33], [44] and [43] where rates are established for some discrete-to-continuum problems involving the p-Laplacian in the nonlocal case using the same discretization procedures.

Discrete-to-continuum work. Other energies and operators have also been studied under the discrete-to-continuum lens, for example the eikonal equation [36], total variation [39], the Ginzburg-Landau functional [10], [68], [27], [62], the Mumford-Shah functional [22], an application in empirical risk minimization [65], various Sobolev semi-norms [61], [66], [31], [73] or variants of the Laplacian operator [19], [16], [58], [21].

Discrete-to-continuum consistency and convergence have been studied under various topologies. The simplest one is pointwise convergence. In particular, the idea is to consider the limit of the discrete energies applied to a (sufficiently regular) continuum function restricted to the discrete domain. This has been explored for example in [7], [23], [40], [46], [45], [60], [63]. A more elaborate approach is to consider the convergence of the operators directly, namely convergence of their spectral properties (eigenvalues and eigenfunctions); we refer to [38], [20], [66], [69], [55], [70], [28] for illustrations of this type of approach. As eigenfunctions are obviously themselves functions, these papers rely on pointwise convergence or TL^p-convergence which was introduced in [39]: in the latter, one starts with discrete functions and extends them to the continuum through the use of optimal transport theory.

The rest of the paper is organized as follows: in Section 2 we introduce the main theoretical tools required for the proofs; in Section 3 we present our main results; in Section 4 we prove our results. For some technical results, we will also refer to the supplementary material file whose sections are indexed by roman numbers.

2 Background

2.1 General notation

For $\Omega \subseteq \mathbb{R}^d$, we denote by $\overrightarrow{n}$ the outward normal vector to the boundary $\partial\Omega$ of Ω. We denote the identity operator by Id and the indicator function of a set A by χ_A. We will write $\mathrm{cl}(\Omega)$ to denote the closure of Ω. For $T > 0$, let λ_t and λ_x respectively be the Lebesgue measure on $[0, T]$ and Ω. Elements of a discrete space will be over-lined, for example $\bar{u} \in \mathbb{R}^d$. The i-th component of $\bar{u}$ is denoted by $(\bar{u})_i$. We will denote Lebesgue spaces by L^p, the space of functions with k-th continuous derivatives by C^k, Hölder spaces by $\mathrm{C}^{k,\alpha}$ and Sobolev spaces by $\mathrm{W}^{k,p}$. We write $\|\cdot\|_{\mathrm{L}^p(A)}$ for the L^p-norm over a space A and $\|\cdot\|_{\mathrm{L}^p}$ when we take the norm over the whole space Ω or $\Omega \times \Omega$ depending on the domain of the function. We also write that $u_n \rightharpoonup u$ if u_n converges weakly to u.

Functions We define Lambert's $W : [0, \infty) \mapsto [0, \infty)$ function [53] as the inverse of the function xe^x: for every $y > 0$, we have $W(y)e^{W(y)} = y$. It is clear that W is increasing on $[0, \infty)$. From this, it follows that the function $\mathrm{expW} : [0, \infty) \mapsto [1, \infty)$ defined as $\mathrm{expW}(y) = e^{W(y)}$ is also increasing and solves $\mathrm{expW}(y) \log(\mathrm{expW}(y)) = y$ for every $y > 0$.

Asymptotics For two functions $f : \mathbb{N} \mapsto [0, \infty)$ and $g : \mathbb{N} \mapsto [0, \infty)$, we will write $f(n) \gg g(n)$ if $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$. Therefore, $1 \gg f$ means that $\lim_{n\to\infty} f(n) = 0$.

2.2 Nonlinear semigroup theory

We now introduce a few elements of nonlinear semigroup theory that will be useful in the rest of the paper. However, we stress that a thorough and proper treatment of the subject can, for example, be found in [1], [4], [26], [15], [8], [54] and references therein.

In the following, V is a Banach space with its topological dual V^* and norm $\|\cdot\|$. The notation $A : V \to 2^V$ means that A is a set-valued operator on V. The domain and range of A are defined respectively as $\mathrm{dom}(A) = \{v \in V \,|\, Av \neq \emptyset\}$ and $\mathrm{ran}(A) = \{y \,|\, y \in Av \text{ for some } v \in \mathrm{dom}(A)\}$. The graph of A is $\mathrm{gph}(A) = \{(v, w) \in V \times V \,|\, w \in Av\}$.

Let us consider the abstract Cauchy problem for an operator A on V:

$$(\mathrm{CP}_{f,u_0}) \qquad \begin{cases} u'(t) + Au(t) \ni f(t) & \text{on } t \in (0, T), \\ u(0) = u_0 \end{cases}$$

for some $f : (0, T) \mapsto V$ and $u_0 \in V$. Various concepts of solution have been designed for the problem (5) and we refer to [1] for a brief review of abstract Cauchy problems. We will rely on the following notion of solution as in [1, Definition A.3]. We note that the following type of solutions are called strong solutions in the nonlinear semigroup literature.

Definition 2.1 (Solution to (CP_{f,u_0})). *A function u is called a strong solution to (CP_{f,u_0}) if $u(t, x) \in C([0, T]; V) \cap \mathrm{W}^{1,1}_{\mathrm{loc}}((0, T); V)$, $u(0, \cdot) = u_0$ and $u'(t) + Au(t) \ni f(t)$ λ_t-a.e..*

Our aim is to introduce a class of operators that will be of particular interest when solving the abstract Cauchy problem (CP_{f,u_0}).

Definition 2.2 (Accretive operator)**.** *We say that an operator A on V is accretive if*

$$\|v - w\| \leq \|v - w + \lambda(\hat{v} - \hat{w})\|$$

for all $\lambda > 0$ and $(v, \hat{v}), (w, \hat{w}) \in \mathrm{gph}(A)$.

One can see that A is accretive if and only if $(\mathrm{Id} + \lambda A)^{-1}$ is a single-valued nonexpansive (i.e., 1-Lipschitz) map for $\lambda > 0$. $(\mathrm{Id} + \lambda A)^{-1}$ is known as the the resolvent of λA. Observe that $\mathrm{dom}((\mathrm{Id} + \lambda A)^{-1}) = \mathrm{ran}(\mathrm{Id} + \lambda A)$.

We recall that an operator A on a Hilbert space H endowed with the inner product $\langle \cdot, \cdot \rangle_H$, is monotone if $\langle v - w, \hat{v} - \hat{w} \rangle_H \geq 0$ for all $(v, \hat{v}), (w, \hat{w}) \in \mathrm{gph}(A)$. Accretivity is a generalization of monotonicity in Hilbert spaces.

Proposition 2.3 (Equivalence of accretivity and monotonicity)**.** *An operator A on a Hilbert space H is accretive if and only if it is monotone.*

Definition 2.4 (m-accretive operators)**.** *We say that an operator A on V is m-accretive if A is accretive and $\mathrm{ran}(\mathrm{Id} + \lambda A) = V$ for all $\lambda > 0$.*

A m-accretive operator A in V is maximal accretive in the sense that there exists no other accretive operator whose graph properly contains $\mathrm{gph}(A)$. In general, the converse is not true, but it is true in Hilbert spaces due to the celebrated Minty theorem [15].

The m-accretivity property of operators is a sufficient condition for the existence of a solution to (CP_{f,u_0}) (see [54, Chapter 2, Theorem 10.2]).

Theorem 2.5 (Existence of a solution to (CP_{f,u_0}))**.** *Let V be a reflexive Banach space, A a m-accretive operator, $u_0 \in \mathrm{dom}(A)$ and $f \in \mathrm{W}^{1,1}(0, T; V)$, then there exists a unique solution as in Definition 2.1 to* (CP_{f,u_0}).

In the remainder of this section, we specialize our discussion by considering the Banach spaces $\mathrm{L}^p(\Omega)$, where $\Omega \subseteq \mathbb{R}^d$. In order to present these results, we first introduce the following elements. Let $\mathrm{L}^0(\Omega)$ be the set of measurable functions on Ω that map to $\mathbb{R}$ and define the two sets of functions

$$\mathcal{H} = \{h \in \mathrm{C}^\infty(\mathbb{R}) \,|\, 0 \leq h' \leq 1,\ \mathrm{supp}(h') \text{ is compact, } 0 \notin \mathrm{supp}(h)\}$$

and

$$\mathcal{J} = \{j : \mathbb{R} \mapsto [0, +\infty] \,|\, j \text{ is convex, lower semi-continuous and satisfies } j(0) = 0\}.$$

Then, we can introduce the following relation for two functions $v, w \in \mathrm{L}^0(\Omega)$:

$$v \overset{\mathrm{sgt}}{\ll} w \quad \text{if and only if } \int_\Omega j(v(x))\,\mathrm{d}x \leq \int_\Omega j(w(x))\,\mathrm{d}x \text{ for all } j \in \mathcal{J}.$$

Definition 2.6 (Completely accretive operators)**.** *We say that an operator A on $\mathrm{L}^0(\Omega)$ is completely accretive if $v - w \overset{\mathrm{sgt}}{\ll} v - w + \lambda(\hat{v} - \hat{w})$ for all $\lambda > 0$ and $(v, \hat{v}), (w, \hat{w}) \in \mathrm{gph}(A)$.*

We note that completely accretive operators need not be defined on Banach spaces in contrast to accretive operators. However, if the graph of a completely accretive operator A is contained in $\mathrm{L}^p(\Omega) \times \mathrm{L}^p(\Omega)$, and since $\mathrm{L}^p(\Omega) \subseteq \mathrm{L}^0(\Omega)$, then A is accretive in $\mathrm{L}^p(\Omega)$. We refer to Lemma II.2 in Section II for additional details.

In Section 4.1, we will consider operators A such that $\mathrm{gph}(A) \subseteq \mathrm{L}^1(\Omega) \times \mathrm{L}^1(\Omega)$ with $\lambda_x(\Omega) < \infty$. In this case, another useful characterization of completely accretive operators on $\mathrm{L}^1(\Omega)$ uses $\mathcal{H}$ (see [1, Corollary A.38]).

Proposition 2.7 (Characterization of completely accretive operators)**.** *Let A be an operator with $\operatorname{gph}(A) \subseteq \mathrm{L}^1(\Omega) \times \mathrm{L}^1(\Omega)$ and $\lambda_x(\Omega) < \infty$. Then, A is completely accretive if and only if*

$$\int_\Omega h(v(x) - w(x))(\hat{v}(x) - \hat{w}(x) \, \mathrm{d}x \geq 0$$

for any $h \in \mathcal{H}$ and $(v, \hat{v}), (w, \hat{w}) \in \operatorname{gph}(A)$.

One can combine Definition 2.4 and Definition 2.6 to define m-completely accretive operators. We now present a variant of [1, Theorem A.20] whose proof can be found in Section II.

Lemma 2.8 (Contraction property for completely accretive operators)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be bounded, $p \geq 1$, A be a completely accretive operator with $\operatorname{gph}(A) \subseteq \mathrm{L}^p(\Omega) \times \mathrm{L}^p(\Omega)$ and u and v be solutions as in Definition 2.1 that respectively solve (CP_{f,u_0}) and (CP_{g,v_0}). Then, for any $1 \leq r \leq \infty$ and $0 \leq t \leq T$, we have:*

$$\|u(t) - v(t)\|_{\mathrm{L}^r} \leq \|u_0 - v_0\|_{\mathrm{L}^r} + \int_0^t \|f(s) - g(s)\|_{\mathrm{L}^r} \, \mathrm{d}s. \tag{3}$$

In particular, (CP_{f,u_0}) has a unique strong solution.

Our motivation for the use of nonlinear semigroup theory really stems from Lemma 2.8. Indeed, we will consider Cauchy problems of the form (CP_{f,u_0}) for our gradient flows. One could be tempted to use direct results for the existence of gradient flows such as described in [59, Section 8] for example. However, we would not obtain the very strong contraction property (3) which will be essential to establish our rates in Section 4.

2.3 Nonlinear problems

In this subsection, we will introduce several results related to nonlinear problems which will be used in Section 4.1.

Definition 2.9 (Hemicontinuity)**.** *Let $A : V \to V^*$ be an operator. We say that A is hemicontinous if for all $v, w, z \in V$, the function $\lambda \mapsto \langle A(v + \lambda w), z \rangle_{V^*}$ is continuous from $\mathbb{R}$ to $\mathbb{R}$.*

A weak version of [37, Lemma 6.2] is as follows and will be used for the well-posedness result of the solution to the problem (7) in the proof of Theorem 4.9.

Lemma 2.10 (Continuity implies hemicontinuity)**.** *Let $A : V \to V^*$ be an operator. If A is continuous, then A is hemicontinuous.*

The next result deals with the existence of a solution to a nonlinear stationary problem [50, Chapter 2, Theorem 2.1]. This will be an essential step in the proof of the range condition of our operators in Proposition 3.5.

Theorem 2.11 (Solutions to nonlinear stationary problems)**.** *Let V be a separable reflexive Banach space and $A : V \to V^*$ an operator which is bounded, hemicontinuous and satisfies*

$$\langle A(v) - A(w), v - w \rangle_{V^*} \geq 0 \qquad \text{for all } v, w \in V \text{ as well as}$$

$$\frac{\langle A(v), v \rangle_{V^*}}{\|v\|_V} \to +\infty \qquad \text{when } \|v\|_V \to +\infty.$$

Then, for every $f \in V^*$, there exists $u \in V$ such that

$$A(u) = f.$$

For the existence of the local problem, we will use the following result on nonlinear evolution problems which can be found in [50, Chapter 2, Theorem 1.4 and Remark 1.13].

Theorem 2.12 (Solutions to nonlinear evolution problems)**.** *Let H be an Hilbert space and V_i reflexive Banach spaces with $V_i \subseteq H$ and V_i dense in H for $1 \leq i \leq q$. Assume that $\cap_{i=1}^q V_i$ is separable and dense in H. For $1 \leq i \leq q$, let $A_i : V_i \to V_i^*$ be an operator which is bounded, hemicontinuous and, for some $1 < p_i < \infty$ and $c_i > 0$, satisfies:*

1. $\|A_i(v)\|_{V_i^*} \leq c_i \|v\|_{V_i}^{p_i - 1}$*;*

2. $\langle A_i(v) - A_i(w), v - w\rangle_{V_i^*} \geq 0$ *for all* $v, w \in V_i$*;*

3. there exists a seminorm $[\cdot]_i$ *on* V_i *with constants* $\alpha_i, \lambda_i, \beta_i > 0$ *such that* $[\cdot]_i + \lambda_i \|\cdot\|_H \geq \beta_i \|\cdot\|_{V_i}$ *on* V_i *and* $\langle A_i(v), v\rangle_{V_i^*} \geq \alpha_i [v]_i^{p_i}$ *for all* $v \in V_i$*.*

Then, for $u_0 \in H$ *and* $f \in \sum_{i=1}^q \mathrm{L}^{p_i^*}(0, T; V_i^*)$ *where* $1/p_i + 1/p_i^* = 1$*, there exists a unique function* u *such that*

$$u \in \bigcap_{i=1}^{q} \mathrm{L}^{p_i}(0, T; V_i), \quad u \in \mathrm{L}^\infty(0, T; H)$$

that satisfies

$$\frac{\partial u}{\partial t} + \sum_{i=1}^q A_i(u) = f \quad \textit{and} \quad u(0) = u_0.$$

2.4 Piecewise constant approximations

We start by introducing a space of functions which will be relevant for our discrete-to-continuum approximations. We refer to [64] for a detailed discussion.

Definition 2.13 (Lipschitz spaces)**.** *Let* Ω *be an open bounded subset of* $\mathbb{R}^d$*. For* $g \in \mathrm{L}^q(\Omega)$ *with* $q \in [1, +\infty)$*, we define the (first-order)* $\mathrm{L}^q(\Omega)$ *modulus of smoothness by*

$$\omega(g, h)_q = \sup_{z \in \mathbb{R}^d, |z| < h} \left(\int_{x, x+z \in \Omega} |g(x+z) - g(x)|^q \, \mathrm{d}x \right)^{1/q}.$$

For $0 < s \leq 1$*, the Lipschitz spaces* $\mathrm{Lip}(s, \mathrm{L}^q(\Omega))$ *consist of all functions* $g \in \mathrm{L}^q(\Omega)$ *for which*

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

Lipschitz spaces contain functions with, roughly speaking, s "derivatives" in $\mathrm{L}^q(\Omega)$. We also note that we restrict ourselves to $0 < s \leq 1$ since for $s \geq 1$, the only functions in $\mathrm{Lip}(s, \mathrm{L}^q(\Omega))$ are constants by [30, Chapter 2, Proposition 7.1]. Lipschitz spaces allow for a broad range of functions and namely, $\mathrm{Lip}(1, \mathrm{L}^1(\Omega))$ contains functions of bounded variation, for example [30, Chapter 2, Lemma 9.2].

We will be considering the error rate between a function and its piecewise constant approximation. The results presented below are part of the broader literature on approximation theory and in particular, spline approximations (we refer to [30] for a review of such topics). We begin by defining operators that will allow us to connect discrete and continuum spaces.

Let $\Omega \subseteq \mathbb{R}^d$ be some bounded set and let $\Pi = \{\pi_i\}_{i=1}^{|\Pi|}$ be a disjoint partition of Ω with cardinality $|\Pi|$. We define the projection operator $\mathcal{P}_\Pi : \mathrm{L}^1(\Omega) \mapsto \mathbb{R}^{|\Pi|}$ and the injection operator $\mathcal{I}_\Pi : \mathbb{R}^{|\Pi|} \mapsto \mathrm{L}^1(\Omega)$ as

$$(\mathcal{P}_\Pi u)_i = \frac{1}{\lambda_x(\pi_i)} \int_{\pi_i} u(x) \, \mathrm{d}x \quad \text{and} \quad (\mathcal{I}_\Pi \bar{u})(x) = \sum_{i=1}^{|\Pi|} \bar{u}_i \chi_{\pi_i}(x)$$

respectively for $u \in \mathrm{L}^1(\Omega)$ and $i = 1, \ldots, |\Pi|$, $\bar{u} \in \mathbb{R}^{|\Pi|}$ and $x \in \Omega$. When using the partition $\{\pi_i \times \pi_j\}_{i,j=1}^{|\Pi|}$ of $\Omega \times \Omega$, by an abuse of notation, we have

$$(\mathcal{P}_\Pi v)_{i,j} = \frac{1}{\lambda_x(\pi_i)\lambda_x(\pi_j)} \int_{\pi_i} \int_{\pi_j} v(x, y) \, \mathrm{d}y \mathrm{d}x \quad \text{and} \quad (\mathcal{I}_\Pi \bar{v})(x, y) = \sum_{i,j=1}^{|\Pi|} \bar{v}_{i,j} \chi_{\pi_i}(x) \chi_{\pi_j}(y)$$

for $v \in \mathrm{L}^1(\Omega \times \Omega)$, $\bar{v} \in \mathbb{R}^{|\Pi| \times |\Pi|}$ and $(x, y) \in \Omega \times \Omega$.

For a function u, in order to obtain quantitative rates for $\|u - \mathcal{I}_\Pi \mathcal{P}_\Pi u\|_{\mathrm{L}^p}$ the choice of Π and the regularity of u are essential. We now describe one construction of a partition Π but refer to [28] and references therein for more examples. In order to simplify the discussion, we consider $\Omega = (0,1)^d$ although some of the results described below hold in more general cases.

For $n \in \mathbb{N}$, define $[n] = 1, \ldots, n$ and let $\boldsymbol{i} = (i_1, i_2, \ldots, i_d) \in [n]^d$ be a multi-index. We partition Ω into n^d hypercubes with sides of length n^{-1} and denote this partition by $\Pi_{\mathrm{uni},n} = \{\Omega_{n,\boldsymbol{i}}\}_{\boldsymbol{i}\in[n]^d}$. We have the following approximation lemma whose proof can be found in Section II.

Lemma 2.14 (Approximation on the uniform partition). *Let $\Omega = (0,1)^d$, $g \in \mathrm{Lip}(s, \mathrm{L}^q(\Omega))$ and assume the partition $\Pi_{\mathrm{uni},n}$ on Ω. For $0 < s \leq 1$ and $1 \leq q \leq \infty$, we have*

$$\|g - \mathcal{I}_{\Pi_{\mathrm{uni},n}} \mathcal{P}_{\Pi_{\mathrm{uni},n}} g\|_{\mathrm{L}^q} \leq C |g|_{\mathrm{Lip}(s,\mathrm{L}^q(\Omega))} n^{-s},$$

for some constant $C > 0$ depending only on d.

Furthermore, if g is in $\mathrm{C}^{0,\alpha}$ for $0 < \alpha \leq 1$ and $\varepsilon > 0$, then:

$$\|g(\cdot/\varepsilon) - \mathcal{I}_{\Pi_{\mathrm{uni},n}} \mathcal{P}_{\Pi_{\mathrm{uni},n}} g(\cdot/\varepsilon)\|_{\mathrm{L}^q} \leq C \varepsilon^{-\alpha} n^{-\alpha}$$

for some $C > 0$ dependent on d.

Remark 2.15 (Rates for functions on the product space). We note that the conclusions of Lemma 2.14 also apply to functions g defined on $\Omega \times \Omega$.

2.5 Random graph models

Whenever we deal with random graph models, we will assume that we have the uniform partition $\Pi_{\mathrm{uni},n} = \{\Omega_{n,i}\}_{i=1}^n$ of $(0,1)$ and shall therefore write $\mathcal{P}_n = \mathcal{P}_{\Pi_{\mathrm{uni},n}}$ and $\mathcal{I}_n = \mathcal{I}_{\Pi_{\mathrm{uni},n}}$.

In some applications, data is represented in the form of an undirected graph. One approach to understanding the underlying structure is to analyse the convergence properties of the graph as the number of vertices goes to infinity. We therefore consider graphs through their weight function defined on $[0,1]^2$: for a graph G with vertices labelled by $[n]$ and weight matrix $\{\bar{K}_{n,ij}\}_{i,j=1}^n$ with $\bar{K}_{n,ii} = 0$ and $\bar{K}_{n,ji} = \bar{K}_{n,ij} \geq 0$ for $i \neq j$, for $(x,y) \in \Omega_{n,i} \times \Omega_{n,j}$, we define $\tilde{K}_n(x,y) = \bar{K}_{n,ij}$. The objective is now to analyze the limit of the step-functions $\tilde{K}_n$ as $n \to \infty$: this is well-known to be a graphon, i.e. a symmetric kernel function in $\mathrm{L}^1([0,1]^2)$ (see [12] and references therein). Given a sequence of graphs from a certain graph model, i.e. a sequence of graphs for which there is a systematic way to determine the graph weights $\{\bar{K}_{n,ij}\}_{i,j=1}^n$, we can establish the convergence of the sequence to the corresponding graphon in an appropriate metric. In the data-centric approach, one tries to fit graph models to data therewith estimating the underlying graphon (see for example [11], [74]): the intuition here is that graph sequences that have related graphons as limiting points should share similarities.

In this paper, we consider a general sparse random graph model originally introduced and studied by [12].

Definition 2.16 (Random graph models). *For $n \in \mathbb{N}$, let $\rho_n > 0$ and $\bar{K} \in \mathbb{R}^{n\times n}$ be a symmetric matrix with non-negative entries. Assume $\rho_n \bar{K}_{ij} \leq 1$ for all $i,j \in [n]$ and $\bar{K}_{ii} = 0$ for all $i \in [n]$. Let*

$$\left(\bar{\Lambda}_n\right)_{ij} = \begin{cases} \frac{1}{\rho_n} & \textit{with probability } \rho_n \bar{K}_{ij} \\ 0 & \textit{else.} \end{cases}$$

We define the random graph $G(n, \bar{\Lambda}_n)$ to be the random graph with vertices $[n]$ and weight matrix $\bar{\Lambda}_n$.

As a simple example, consider the case where $\rho_n = 1$ and the entries of $\bar{K}$ are $\bar{K}_{ij} = p \in (0,1)$ for $1 \leq i,j \leq n$. Then, vertices i and j are connected with probability p and $G(n, \bar{\Lambda}_n)$ is the Erdős-Rényi graph model $G(n,p)$.

Note that in previous papers (for example [33, Section 3.2]), the weight matrix only has 0-1 weights but, when the graphs are used in evolution problems, the quantities are normalized by ρ_n^{-1} which is related to the average degree of the graph. Altering the definition as we did in Definition 2.16 leads to more a straight-forward and clearer problem setting.

In our paper, we will be considering the limit to a local problem and are therefore less concerned with statements related to the convergence of our random graph models. We refer to [12] and [33, Proposition 3.1] for more details on the topic however.

3 Main results

3.1 Assumptions

In this section, we enumerate all the assumptions on the space, operators, kernels and length-scale that we will be using throughout the paper.

Assumptions 1 (Assumptions on the space).

S.1 The space Ω is a bounded open subset of $\mathbb{R}^d$.
S.2 The space Ω has Lipschitz boundary.

We recall that for an operator $\mathcal{A}$, we denote its adjoint by $\mathcal{A}^*$.

Assumptions 2 (Assumptions on the operator $\mathcal{A}$).

O.1 The operator $\mathcal{A} : \mathrm{L}^2(\Omega) \mapsto \mathrm{L}^2(\Omega)$ is bounded and linear. We write $C_{\mathrm{op}} = \|\mathcal{A}\|_{\mathrm{op}}$, the operator norm of $\mathcal{A}$.
O.2 The operator $\mathcal{A}^*\mathcal{A}$ is order-preserving: for all $u \in \mathrm{L}^2(\Omega)$ and almost every $x \in \Omega$, $\mathrm{sign}(\mathcal{A}^*\mathcal{A}u(x)) = \mathrm{sign}(u(x))$ where sign is the sign function.
O.3 For $n \in \mathbb{N}$ and a partition Π_n, there exists a positive semi-definite linear operator $\bar{G}_n : \mathbb{R}^{|\Pi_n|} \mapsto \mathbb{R}^{|\Pi_n|}$ such that $\mathcal{I}_n(\bar{G}_n(\bar{u})) = \mathcal{A}^*\mathcal{A}(\mathcal{I}_n\bar{u})$.

From [57, Theorem VI.2 and Theorem VI.3]), if Assumption **O.1** holds, it is easy to see that $\|\mathcal{A}^*\mathcal{A}v\|_{\mathrm{L}^2} \leq C_{\mathrm{op}}^2\|v\|_{\mathrm{L}^2}$ for all $v \in \mathrm{L}^2(\Omega)$. Assumption **O.3** reflects the fact that we require the discretization $\bar{G}_n$ of $\mathcal{A}^*\mathcal{A}$ to commute with the injection operator $\mathcal{I}_n$. This property is for example satisfied in the following case: if $\mathcal{A}$ is unitary, then $\mathcal{A}^*\mathcal{A} = \mathrm{Id}$ so that $\bar{G}_n$ can be chosen to be the identity as well. Furthermore, in Corollary 4.8, we will use similar properties of the differential operators introduced in Section 3.2.

Assumptions 3 (Assumptions on the kernel K).

K.1 The kernel $K : [0,\infty) \mapsto [0,\infty)$ is bounded in L^∞.
K.2 The kernel $K : [0,\infty) \mapsto [0,\infty)$ has $\mathrm{supp}(K) = [0,1]$.

The next set of assumptions concerns the regularity of the solutions to our various gradient flows. While we rigorously introduce the latter in Section 3.2, for the reader's convenience, we repeat the continuum nonlocal and local variants, respectively (5) and (7), below:

$$\begin{cases} \frac{\partial}{\partial t}u(t,x) + \mathcal{E}^K_{\mathcal{A},f}(u(t,x)) = 0 & \text{on} \quad (0,T)\times\Omega, \\ u(0,x) = u_0(x) \end{cases}$$

and

$$\begin{cases} \frac{\partial}{\partial t}u(t,x) + \mu\Delta_p u(t,x) + \mathcal{A}^*\mathcal{A}u(t,x) = \mathcal{A}^*\ell(x), & \text{on} \quad (0,T)\times\Omega \\ |\nabla u(t,x)|^{p-2}\nabla u(t,x)\cdot\overrightarrow{n} = 0, & \text{on} \quad (0,T)\times\partial\Omega \\ u(0,x) = u_0(x). \end{cases}$$

Assumptions 4 (Assumptions on regularity).

R.1 The function $u \in \mathrm{C}^0([0,T];\mathrm{L}^p(\Omega))$ solving (5) satisfies $u(t) \in \mathrm{W}^{1,p}(\Omega)$ for $0 < t < T$.
R.2 For $h \geq 1$, $s > 3 + d/p$ and $r > 2 + d/p$, the function $u \in \mathrm{L}^p(0,T;\mathrm{W}^{1,p}(\Omega)) \cap \mathrm{L}^2(0,T;\mathrm{L}^2(\Omega))$ solving (7) satisfies $u \in \mathrm{L}^h(0,T;\mathrm{W}^{s,p}(\Omega)) \cap \mathrm{L}^\infty(0,T;\mathrm{W}^{r,p}(\Omega))$.
R.3 For $h \geq 1$ and $s > 3 + d/p$, there exists $\tilde{C}$ independent of T such that the function

$$u \in \mathrm{L}^p(0,T;\mathrm{W}^{1,p}(\Omega)) \cap \mathrm{L}^2(0,T;\mathrm{L}^2(\Omega))$$

solving (7) satisfies

$$u \in \mathrm{L}^h(0,T;\mathrm{W}^{s,p}(\Omega)) \quad \text{and} \quad \max\left\{\sup_{t\in(0,T)}\|\nabla u\|_{\mathrm{L}^\infty}, \sup_{t\in(0,T)}\|\nabla^2 u\|_{\mathrm{L}^\infty}\right\} \leq \tilde{C}$$

for all $T > 0$.

Remark 3.1 (Higher regularity). In Theorem 4.10, we will assume higher regularity on the solutions of (5) and (7), namely Assumptions **R.1**, **R.2** and **R.3**. Previous attempts at deriving rates in related problems [6], [41], [71], [5] have led to similar assumptions.

In the nonlocal case, we require the solutions to be in the same space as the local solution, namely $\mathrm{W}^{1,p}(\Omega)$. This is a natural requirement for establishing rates which is fulfilled when approximating local problems by finite elements as in [41]. This is in contrast with our finite-differences approach for which conservation of regularity properties of the local problem is not inherent. For a similar problem, regularity of a nonlocal solution is studied in [3].

In the local case, we extend the regularity from $\mathrm{W}^{1,p}(\Omega)$ to some fractional Sobolev space $\mathrm{W}^{s,p}(\Omega)$. This will allow us to consider a function u with continuous first, second and third derivatives. In order to align for example with [41, Section 6] where functions are taken in $\mathrm{W}^{2,p}(\Omega)$, and due to the regularity of Sobolev spaces we would like to have $s = k + \delta$ for $k = 2, 3$. As it will turn out, this uncovers a subtle interplay between regularity and dimension of the underlying space.

Remark 3.2 (Embeddings of $\mathrm{W}^{k+\delta,p}(\Omega)$ into $\mathrm{C}^k(\Omega)$). In view of Remark 3.1, we are interested in the embedding of $\mathrm{W}^{k+\delta,p}(\Omega)$ into $\mathrm{C}^k(\Omega)$. In fact, $\mathrm{W}^{k+\delta,p}(\Omega)$ embeds into $\mathrm{C}^k(\Omega)$ only when $\delta > d/p$: the weaker we want our regularity assumption on the local solution to be, the higher the regularization parameter p needs to be.

Assumptions 5 (Assumptions on the length-scale).

L.1 The length scale $\varepsilon = \varepsilon_n$ is positive and converges to 0, i.e. $0 < \varepsilon_n \to 0$ as $n \to \infty$.

3.2 Setting

3.2.1 Discrete problem

Given a partition Π of Ω and a symmetric $\bar{K} \in \mathbb{R}^{|\Pi|\times|\Pi|}$, we define the discrete nonlocal p-Laplacian operator for $\bar{u} \in \mathbb{R}^{|\Pi|}$ and $1 \le i \le |\Pi|$ as follows:

$$(\Delta_{p,\Pi}^{\bar{K}} \bar{u})_i = -\sum_{j=1}^{|\Pi|} \lambda_x(\pi_j) \bar{K}_{i,j} |(\bar{u})_j - (\bar{u})_i|^{p-2}((\bar{u})_j - (\bar{u})_i).$$

Let $\bar{f} \in \mathbb{R}^{|\Pi|}$ and $\bar{u}_0 \in \mathbb{R}^{|\Pi|}$. We will consider the fully discrete nonlocal evolution problem

$$\begin{cases} \frac{\bar{u}^k - \bar{u}^{k-1}}{\tau^{k-1}} + \mu \Delta_{p,\Pi}^{\bar{K}} \bar{u}^k + \bar{G}(\bar{u}^k) = \bar{f}, \quad \text{for} \quad 1 \le k \le N \\ \bar{u}(0) = \bar{u}_0 \end{cases} \tag{4}$$

for some $\mu > 0$, a positive sequence $\{\tau^k\}_{k=1}^{N-1}$ (with $\sum_{k=0}^{N-1} \tau^k = T$) and linear operator $\bar{G} : \mathbb{R}^{|\Pi|} \to \mathbb{R}^{|\Pi|}$. We also define $\tau = \max_{k=1,\dots,N} \tau^k$ and will write τ_n for τ when our sequence $\{\tau^k\}_{k=1}^{N-1}$ will be indexed by n. We say that $\bar{u}$ solves (4) with parameters $\bar{K}$, $\bar{f}$ and $\bar{u}_0$.

Later we choose $\tau^k = t^{k+1} - t^k$ for a discretisation $0 = t^0 < t^1 < \dots < t^N = T$ of $[0, T]$ and we will need the following two quantities. First, we define the time interpolated version of the injected vectors $\{\bar{u}^k\}_{k=1}^N$:

$$\mathrm{u}_{\mathrm{TimeInt}}(t, x) = \frac{t^k - t}{\tau^{k-1}} \left(\mathcal{I}_{\Pi_n} \bar{u}^{k-1}\right)(x) + \frac{t - t^{k-1}}{\tau^{k-1}} \left(\mathcal{I}_{\Pi_n} \bar{u}^k\right)(x) \text{ for } (t, x) \in (t^{k-1}, t^k] \times \Omega$$

and $\mathrm{u}_{\mathrm{TimeInt}}(0, x) = \mathcal{I}_{\Pi_n} \bar{u}_0$. Second, we define the time injected version of the injected vectors $\{\bar{u}^k\}_{k=1}^N$:

$$\mathrm{u}_{\mathrm{TimeInj}}(t, x) = \sum_{i=1}^{N} \left(\mathcal{I}_{\Pi_n} \bar{u}^k\right)(x) \chi_{(t^{k-1}, t^k]}(t).$$

Most of our results will holds for any partition Π but, keeping in mind that we are ultimately interested in convergence results, meaning the case where $\max_{j \le |\Pi|} \lambda_x(\pi_j) \to 0$, we will index our partition with a parameter n. We denote the latter by $\Pi_n = \{\pi_j^n\}_{j=1}^{|\Pi_n|}$. As an example, one can consider $\Pi_{\mathrm{uni},n}$. For ease of notation we will write $\Delta_{p,n}^{\bar{K}} = \Delta_{p,\Pi_n}^{\bar{K}}$, $\mathcal{I}_n = \mathcal{I}_{\Pi_n}$, $\mathcal{P}_n = \mathcal{P}_{\Pi_n}$ and similarly for other quantities.

3.2.2 Nonlocal problem

Given a function a kernel function $K : [0, \infty) \to [0, \infty)$, we define the nonlocal p-Laplacian operator Δ_p^K for a function $u \in \mathrm{L}^1(\Omega)$ and $x \in \Omega$ as follows:

$$\Delta_p^K u(x) = -\int_\Omega K(|x-y|)|u(y)-u(x)|^{p-2}(u(y)-u(x))\,\mathrm{d}y.$$

For a kernel K, we define the function $\tilde{K} : \Omega \times \Omega \mapsto [0, \infty)$ by $\tilde{K}(x, y) = K(|x-y|)$. For a general function $v : \Omega \times \Omega \mapsto [0, \infty)$, we naturally have:

$$\Delta_p^v u(x) = -\int_\Omega v(x,y)|u(y)-u(x)|^{p-2}(u(y)-u(x))\,\mathrm{d}y.$$

Furthermore, for $f \in \mathrm{L}^2(\Omega)$, $p \geq 2$ and q such that $p^{-1} + q^{-1} = 1$, we define the following evolution operator $\mathcal{E}_{\mathcal{A},f}^K : \mathrm{L}^p(\Omega) \mapsto \mathrm{L}^q(\Omega)$ for $\mu > 0$, a function $u \in \mathrm{L}^p(\Omega)$ and some linear operator $\mathcal{A} : \mathrm{L}^2(\Omega) \mapsto \mathrm{L}^2(\Omega)$:

$$\mathcal{E}_{\mathcal{A},f}^K(u) = \mu\Delta_p^K u + \mathcal{A}^*\mathcal{A}u - f.$$

The well-posedness of this evolution operator will be discussed in Theorem 3.6.

We will consider the following nonlocal evolution problem:

$$\begin{cases} \frac{\partial}{\partial t}u(t,x) + \mathcal{E}_{\mathcal{A},f}^K(u(t,x)) = 0 & \text{on} \quad (0,T)\times\Omega, \\ u(0,x) = u_0(x). \end{cases} \tag{5}$$

We say that u solves (5) with parameters K, f and u_0. In particular, in order to link the above problem with (1), we will be interested in (5) when $\mathcal{E}_{\mathcal{A},f}^K(u) = \mathcal{E}_{\mathcal{A},\mathcal{A}^*\ell}^K(u)$ for some $\ell \in \mathrm{L}^2(\Omega)$.

We are interested in the following solution which is just of Definition 2.1 using the evolution operator appearing in (5).

Definition 3.3 (Nonlocal problem solution)**.** *Assume that Assumptions **S.1**, **O.1**, **O.2** and **K.1** hold. For $p \geq 2$, $T > 0$, $\mu > 0$, $f \in \mathrm{L}^p(\Omega)$, $u_0 \in \mathrm{L}^p(\Omega)$, a function $u(t,x)$ is a solution to the nonlocal problem* (5) *if $u(t,x) \in C([0,T]; \mathrm{L}^p(\Omega)) \cap W_{\mathrm{loc}}^{1,1}((0,T); \mathrm{L}^p(\Omega))$, $u(0,x) = u_0(x)$ λ_x-a.e. on Ω and λ_t-a.e.:*

$$\frac{\partial}{\partial t}u(t,x) + \mu\Delta_p^K u(t,x) + \mathcal{A}^*\mathcal{A}u(t,x) = f(x) \quad \lambda_x\text{-a.e.}.$$

If we are further given a positive sequence $\{\varepsilon_n\}_{n=1}^\infty$, we will write $K_{\varepsilon_n} = \frac{2}{c(p,d)\varepsilon_n^{d+p}} K(\cdot/\varepsilon_n)$ where

$$c(p,d) = \int_{\mathbb{R}^d} K(|x|)|x_d|^p\,\mathrm{d}x. \tag{6}$$

3.2.3 Local problem

The local p-Laplacian operator is defined as

$$\Delta_p u = -\operatorname{div}(|\nabla u|^{p-2}\nabla u).$$

For $p \geq 2$, $\mu > 0$, $\ell \in \mathrm{L}^2(\Omega)$ and $u_0 \in \mathrm{L}^p(\Omega)$, we will consider the following local evolution problem:

$$\begin{cases} \frac{\partial}{\partial t}u(t,x) + \mu\Delta_p u(t,x) + \mathcal{A}^*\mathcal{A}u(t,x) = \mathcal{A}^*\ell(x), & \text{on} \quad (0,T)\times\Omega \\ |\nabla u(t,x)|^{p-2}\nabla u(t,x)\cdot\overrightarrow{n} = 0, & \text{on} \quad (0,T)\times\partial\Omega \\ u(0,x) = u_0(x) \end{cases} \tag{7}$$

for some linear operator $\mathcal{A} : \mathrm{L}^2(\Omega) \mapsto \mathrm{L}^2(\Omega)$.

The following definition is inspired by the results of [50, Chapter 2, Theorem 1.4] which we have stated as Theorem 2.12.

Definition 3.4 (Local weak solution)**.** *Assume that Assumptions **S.1** and **O.1** hold. For* $p \geq 2$*,* $T > 0$*,* $\ell \in \mathrm{L}^2(\Omega)$*,* $u_0 \in \mathrm{L}^2(\Omega)$*, a function* $u(t,x)$ *is a weak solution to* (7) *if* $u \in \mathrm{L}^p(0,T;\mathrm{W}^{1,p}(\Omega)) \cap \mathrm{L}^2(0,T;\mathrm{L}^2(\Omega))$*,* $u \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))$*,* $u(0,\cdot) = u_0$ λ_x*-a.e. on* Ω *and if* λ_t*-a.e.:*

$$\int_\Omega \frac{\partial}{\partial t} u(t,x)\zeta(x)\,\mathrm{d}x + \mu \int_\Omega |\nabla u(t,x)|^{p-2} \nabla u(t,x) \cdot \nabla \zeta(x)\,\mathrm{d}x + \int_\Omega \mathcal{A}^*\mathcal{A}u(t,x)\zeta(x)\,\mathrm{d}x = \int_\Omega \mathcal{A}^*\ell(x)\zeta(x)\,\mathrm{d}x$$

for all $\zeta \in \mathrm{W}^{1,p}(\Omega)$*.*

The derivative $\frac{\partial}{\partial t}u$ makes sense thanks to Theorem 2.12. Indeed, applying the latter to (7), and in view of assumption **O.1**, we have that $\frac{\partial}{\partial t}u \in \mathrm{L}^{p^*}(0,T;\mathrm{W}^{-1,p^*}(\Omega)) + \mathrm{L}^2(0,T;\mathrm{L}^2(\Omega))$. It then follows using the same arguments as in [50, Chapter 2, Remark 1.2] that the solution u is (up to a modification on a Lebesgue measure zero set) a continuous function from $[0,T] \to \mathrm{L}^2(\Omega)$ in such a way that the initial condition in (7) makes perfectly sense.

3.2.4 Random graphs problem

Given a graph $G(n, \bar{\Lambda}_n)$, $\bar{u}_0 \in \mathbb{R}^n$ and $\bar{f} \in \mathbb{R}^n$, we can also consider the following evolution problem:

$$\begin{cases} \frac{\bar{u}^k - \bar{u}^{k-1}}{\tau^{k-1}} + \mu \Delta_{p,n}^{\bar{\Lambda}_n} \bar{u}^k + \bar{G}(\bar{u}^k) = \bar{f}, \quad \text{for} \quad 1 \leq k \leq N \\ \bar{u}(0) = \bar{u}_0 \end{cases} \tag{8}$$

for some $\mu > 0$, linear operator $\bar{G} : \mathbb{R}^n \to \mathbb{R}^n$, partition $0 = t^0 < t^1 < \cdots < t^N = T$ and where $\tau^{k-1} = t^k - t^{k-1}$. We say that $\bar{u}$ solves (4) with parameters $\bar{\Lambda}_n$, $\bar{f}$ and $\bar{u}_0$.

3.3 Main results

3.3.1 Well-posedness of the nonlocal continuum gradient flow

In order to prove well-posedness, we loosely follow the strategy in [2]. Our main contribution lies in the verification of the range condition in Proposition 3.5: we propose an alternative proof for a generalization of the commonly used [2, Theorem 2.4] based on principles related to Γ-convergence (see [14] or [52]). In particular, showing the range condition boils down to solving a PDE which is unsolvable by direct methods. We therefore modify the latter by adding a term that will make the operators involved coercive. We then use a compactness argument to show that the solutions to the PDE approximations converge to a limiting function which solves the initial PDE by the $\liminf$-inequality of Γ-convergence.

Proposition 3.5 (Complete accretivity and range condition)**.** *Assume that Assumptions **S.1**, **O.1**, **O.2** and **K.1** hold. Let* $p \geq 2$ *and assume that* $f \in \mathrm{L}^p(\Omega)$*. Then, the evolution operator* $\mathcal{E}^K_{\mathcal{A},f}$ *is completely accretive and satisfies the range condition* $\mathrm{L}^p(\Omega) \subseteq \mathrm{ran}(\mathrm{Id} + \lambda \mathcal{E}^K_{\mathcal{A},f})$ *for* $\lambda > 0$*.*

The proof of the proposition is given in Section 4.1.1. From the proposition one easily deduces the existence of solutions to the nonlocal problem. The proof of the theorem is also given in Section 4.1.1.

Theorem 3.6 (Existence and uniqueness of a solution for the nonlocal problem)**.** *Assume that Assumptions **S.1**, **O.1**, **O.2** and **K.1** hold. Let* $p \geq 2$*,* $T > 0$*,* $u_0 \in \mathrm{L}^p(\Omega)$ *and assume that* $f \in \mathrm{L}^p(\Omega)$*. Then, there exists a unique solution* u *as in Definition 3.3 to the evolution problem* (5) *with the operator* $\mathcal{E}^K_{\mathcal{A},f}$ *and initial value* u_0*.*

Furthermore, if v *is a solution as above solving* (5) *with the operator* $\mathcal{E}^K_{\mathcal{A},g}$ *and initial value* v_0*, we have*

$$\|u(t) - v(t)\|_{\mathrm{L}^r} \leq \|u_0 - v_0\|_{\mathrm{L}^r} + t\|f - g\|_{\mathrm{L}^r} \tag{9}$$

for $1 \leq r \leq \infty$ *and* $0 \leq t \leq T$*.*

In addition, we also have

$$\|u(t)\|_{\mathrm{L}^r} \leq \|u_0\|_{\mathrm{L}^r} + t\|f\|_{\mathrm{L}^r} \tag{10}$$

for $1 \leq r \leq \infty$ *and* $0 \leq t \leq T$*.*

3.3.2 Rates of convergence

The rates are established by combining two intermediate results. On one hand, we obtain nonlocal-to-local continuum rates by switching from finite-differences to derivatives which allows one to prove the convergence of the nonlocal p-Laplacian operator to the local one. On the other hand, the central point of the proofs is to leverage the fact that the injected discrete gradient flow solution solves a nonlocal continuum gradient flow with particular parameters. We then rely on the contraction properties (9) and (10) to express the rates of convergence of discrete-to-continuum solutions in the nonlocal setting in terms of the discretization error of the initial condition, data and kernel functions.

The next result precisely describes the interplay between our space-localization parameter ε_n and our time-discretization parameter τ_n in order to ensure convergence of the discrete solution to the solution of (7).

Theorem 3.7 (Discrete-to-continuum local rates)**.** *Assume that assumptions **S.1**, **O.1**, **O.2**, **O.3**, **K.1** and **L.1** hold. Let $p \geq 2$, $\mu > 0$, $T > 0$, $u_0 \in \mathrm{L}^p(\Omega)$, $\ell \in \mathrm{L}^2(\Omega)$, Ω' be compactly contained in Ω and assume that $\mathcal{A}^*\ell \in \mathrm{L}^p(\Omega)$. Furthermore, let $n \in \mathbb{N}$ and define $\bar{K}_{\varepsilon_n} = \mathcal{P}_n \tilde{K}_{\varepsilon_n}$, $\bar{f} = \mathcal{P}_n \mathcal{A}^*\ell$, $\bar{u}_0 = \mathcal{P}_n u_0$.*

Then, for any partition $0 = t^0 < t^1 < \cdots < t^N = T$, there exists a sequence $\{\bar{u}_n^k\}_{k=0}^N$ satisfying (4) *with $\bar{K}_{\varepsilon_n}$, $\bar{f}$, $\bar{u}_0$ and $\bar{G}_n$ chosen as above, a solution u_{ε_n} to* (5) *with kernel K_{ε_n} and a solution u to* (7)*.*

*In addition, assume that Assumptions **S.2** and **K.2** hold, $p \geq 3$, that we are using the partition $\Pi_{\mathrm{uni},n}$, that u_{ε_n} satisfies Assumption **R.1** for all $T > 0$ and that u satisfies Assumption **R.3**. For some $1 \leq q_1 < \infty$ and $1 \leq q_2 < \infty$ and $0 < \alpha_i \leq 1$ for $1 \leq i \leq 3$, assume furthermore that $u_0 \in \mathrm{Lip}(\alpha_1, \mathrm{L}^{q_1}(\Omega)) \cap \mathrm{L}^\infty(\Omega)$, $\mathcal{A}^*\ell \in \mathrm{Lip}(\alpha_2, \mathrm{L}^{q_2}(\Omega)) \cap \mathrm{L}^\infty(\Omega)$ and $K \in \boldsymbol{C}^{0,\alpha_3}(\Omega)$. Then, if for some $\kappa > 0$ we set $T(n) = \left(\frac{1}{1+C_{\mathrm{op}}^4}\right)\log(\varepsilon_n^{-\kappa})$ and assume that*

$$\tau_n \ll \frac{\varepsilon_n^{2(d+p)+\kappa}}{\log(\varepsilon_n^{-\kappa})^{(2p-3)}} \tag{11}$$

as well as

$$\varepsilon_n \gg \max\left\{ n^{-\alpha_1/\kappa}, n^{-\alpha_2/\kappa}, \left[\mathrm{expW}\left(n^{\frac{\alpha_3}{\max(1+(d+p+\alpha_3)/\kappa, p-1)}}\right)\right]^{-1/\kappa} \right\},$$

for n large enough, we have:

$$\begin{aligned}\sup_{1\leq k\leq N}\ \sup_{t\in(t^{k-1},t^k]} \|\mathcal{I}_n \bar{u}_n^k - u(t,\cdot)\|_{\mathrm{L}^2(\Omega')} \leq C\Bigg(& \varepsilon_n \log(\varepsilon_n^{-\kappa}) \\ & + \varepsilon_n^{-\kappa}\left[n^{-\alpha_1} + n^{-\alpha_2} + \frac{\log(\varepsilon_n^{-\kappa})^{(p-1)}}{\varepsilon_n^{d+p+\alpha_3} n^{\alpha_3}} + \tau_n \frac{\log(\varepsilon_n^{-\kappa})^{2p-3}}{\varepsilon_n^{2(d+p)}}\right]\Bigg)\end{aligned} \tag{12}$$

for some $C > 0$ that might be dependent on Ω (and d), u_0 and $\mathcal{A}^*\ell$ and the latter right-hand side tends to 0 as $n \to \infty$.

The proof of the theorem can be found in Section 4.2.3.

Remark 3.8 (Asymptotics of N). It is a natural requirement that $N \to \infty$ as $T(n) \to \infty$ as in Theorem 3.7. This is indeed the case (and also will be for Corollary 3.12 by the same argument) since by definition $N \geq T(n)/\tau_n$ and (11) ensures that $\tau_n \to 0$ with $\varepsilon_n \to 0$.

Remark 3.9 (Parameters in Theorem 3.7). Despite their appearances, the conditions on ε_n and τ_n in Theorem 3.7 are not constraining. Indeed, all parameters involved are chosen by the practitioner prior to the implementation of the numerical procedure.

The regularity requirements of both the nonlocal solution and the local solution are discussed in greater detail in Remarks 3.1 and 3.2. The nonlocal regularity assumption is linked to our approximation through finite differences. Indeed, the expression (5) does not contain any differential term in the space component leading our solution to be in $\mathrm{L}^p(\Omega)$ as opposed to $\mathrm{W}^{1,p}(\Omega)$: when using finite elements, the approximations are in the same space as the solutions to original problem, which in our case is $\mathrm{W}^{1,p}$ by Theorem 4.9. The local regularity assumption is a typical one and actually induces an interesting relationship between the regularity requirement and the regularization parameter p.

The rates are formulated for n large enough. This is discussed in Remark 4.11 and is essentially dependent on the choice of our compactly contained set $\Omega' \subseteq \Omega$.

A simple additional step allows one to deduce the convergence of the discrete gradient flow to u_∞. The proof is given in Section 4.2.3.

Theorem 3.10 (Rates for u_∞)**.** *Assume the same setting as in Theorem 3.7 and furthermore that $\mathcal{A} = \mathrm{Id}$. Then, for n large enough, we have*

$$\begin{aligned} \|\mathcal{I}_n \bar{u}_n^N - u_\infty\|_{\mathrm{L}^2(\Omega')} \leq C \Bigg(& \varepsilon_n \log(\varepsilon_n^{-\kappa}) + \varepsilon_n^{\kappa/4} (\mathcal{F}(u_0) - \mathcal{F}(u_\infty))^{1/2} \\ & + \varepsilon_n^{-\kappa} \left[n^{-\alpha_1} + n^{-\alpha_2} + \frac{\log(\varepsilon_n^{-\kappa})^{(p-1)}}{\varepsilon_n^{d+p+\alpha_3} n^{\alpha_3}} + \tau_n \frac{\log(\varepsilon_n^{-\kappa})^{2p-3}}{\varepsilon_n^{2(d+p)}} \right] \Bigg) \end{aligned} \tag{13}$$

for some $C > 0$ that might be dependent on Ω (and d), u_0 and ℓ and the latter right-hand side tends to 0 as $n \to \infty$.

Remark 3.11 (Curse of dimensionality). We note the presence of terms of the form $\varepsilon_n^{-\gamma-\delta d}$ with γ, $\delta > 0$ in the right-hand sides of both (12) and (13). For fixed n and $\varepsilon_n < 1$, as $d \to \infty$, $\varepsilon_n^{-\gamma-\delta d}$ will tend to infinity. While we can control this behaviours when d is fixed, for high dimensions the rates deteriorate.

3.3.3 Application to random graphs

Everything discussed until now was determinisitic. In particular, the discretization procedure was based on partitioning the space into cells in a pre-defined way that would allow us to control the discretization error as described in Section 2.4. An alternative setting is the one of random graph models present in various applications.

Obtaining results in the random graph setting is split in two steps: (1) prove rates of convergence between the discrete random and deterministic gradient flows and then (2) use the deterministic rates of Proposition 4.13 and Theorem 4.10. The first part is conceptually similar to how the results in the nonlocal case are derived while adding the necessary probabilistic estimates. The main results are to be found in Theorems 3.12 and 3.13. The proofs of the latter are given in Section 4.3.

Theorem 3.12 (Discrete random-to-continuum local rates)**.** *Assume that Assumptions **O.1**, **O.2**, **O.3** and **L.1** hold. Let $p \geq 2$, $\mu > 0$, $T > 0$, $u_0 \in \mathrm{L}^p(\Omega)$, $\ell \in \mathrm{L}^2(\Omega)$, Ω' be compactly contained in Ω and assume that $\mathcal{A}^*\ell \in \mathrm{L}^p(\Omega)$. Furthermore, let $n \in \mathbb{N}$ and define $\bar{K}_{\varepsilon_n} = \mathcal{P}_n \tilde{K}_{\varepsilon_n}$, $\bar{f} = \mathcal{P}_n \mathcal{A}^*\ell$, $\bar{u}_0 = \mathcal{P}_n u_0$. We also suppose that ρ_n is a positive sequence with $\rho_n \to 0$ and $\rho_n \ll \varepsilon_n^{1+p}$. Let $\Lambda_n \in \mathbb{R}^{n\times n}$ be the weight matrix defined as in Definition 2.16 with $\bar{K} = \bar{K}_{\varepsilon_n}$.*

Then, for any partition $0 = t^0 < t^1 < \cdots < t^N = T$, there exists a sequence $\{\bar{u}_n^k\}_{k=0}^N$ solving (8) *with parameters $\bar{\Lambda}_n$, $\bar{f}$ and $\bar{u}_0$, a solution u_{ε_n} to* (5) *with kernel K_{ε_n} and a solution u to* (7)*.*

*In addition, assume that Assumptions **S.2** and **K.2** hold, $p \geq 3$, that we are using the partition $\Pi_{\mathrm{uni},n}$, that u_{ε_n} satisfies Assumption **R.1** for all $T > 0$ and that u satisfies Assumption **R.3**. For some $1 \leq q_1 < \infty$ and $1 \leq q_2 < \infty$ and $0 < \alpha_i \leq 1$ for $1 \leq i \leq 3$, assume furthermore that $u_0 \in \mathrm{Lip}(\alpha_1, \mathrm{L}^{q_1}(\Omega)) \cap \mathrm{L}^\infty(\Omega)$, $\mathcal{A}^*\ell \in \mathrm{Lip}(\alpha_2, \mathrm{L}^{q_2}(\Omega)) \cap \mathrm{L}^\infty(\Omega)$ and $K \in C^{0,\alpha_3}(\Omega)$. For some $\kappa > 0$, let $T(n) = \left(\frac{2}{2+3C_{\mathrm{op}}^4}\right) \log(\varepsilon_n^{-\kappa})$ and assume that*

$$\tau_n \ll \frac{1}{\log(\varepsilon_n^{-\kappa})^{(2p-3)}} \varepsilon_n^{2+2p+\kappa},$$

$$\begin{aligned} \varepsilon_n \gg \max\Bigg\{ & n^{-\alpha_1/\kappa}, n^{-\alpha_2/\kappa}, \left[\mathrm{expW}\left(n^{\frac{\alpha_3}{\max(1+(1+p+\alpha_3)/\kappa, p-1)}}\right)\right]^{-1/\kappa}, \\ & \left[\mathrm{expW}\left((n\log(n))^{1/\max(4(p-1), 4+(2+4p)/\kappa)}\right)\right]^{-1/\kappa} \Bigg\} \end{aligned}$$

and

$$\frac{\log(n)\varepsilon_n^{2p}}{n} \ll \rho_n^2 \ll \varepsilon_n^{2(p+1)} \quad \text{as well as} \quad \frac{\log(\varepsilon_n^{-\kappa})^{2(p-1)}}{\varepsilon_n^{1+2p}\log(n)^{1/2}n^{1/2}} \ll \theta_n^2 \ll \varepsilon_n^{2\kappa}.$$

Then, for n large enough, we have:

$$\begin{aligned}\sup_{1\leq k\leq N}\sup_{t\in(t^{k-1},t^k]} \|\mathcal{I}_n\bar{u}_n^k - u(t,\cdot)\|_{\mathrm{L}^2(\Omega')} &\leq C\Bigg(\varepsilon_n\log(\varepsilon_n^{-\kappa}) \\ &+ \varepsilon_n^{-\kappa}\left[n^{-\alpha_1} + n^{-\alpha_2} + \frac{\log(\varepsilon_n^{-\kappa})^{(p-1)}}{\varepsilon_n^{1+p+\alpha_3}n^{\alpha_3}} + \tau_n\frac{\log(\varepsilon_n^{-\kappa})^{2p-3}}{\varepsilon_n^{2(1+p)}} + \theta_n\right]\Bigg)\end{aligned} \tag{14}$$

for some $C > 0$ that might be dependent on Ω, u_0 and $\mathcal{A}^*\ell$ and with probability larger than

$$1 - \frac{C\log(\varepsilon_n^{-\kappa})^{2(p-1)}\left(1+\frac{3C_{\mathrm{op}}^4}{2}\right)^{2(1-p)}}{\theta_n^2\varepsilon_n^{1+p}n\rho_n}.$$

Furthermore, the right-hand side of (14) *tends to 0 as $n\to\infty$ and the probability tends to* 1.

Theorem 3.13 (Rates for u_∞ using the random graph model)**.** *Assume the same setting as in Theorem 3.12 and furthermore that $\mathcal{A} = \mathrm{Id}$. Then, for n large enough, we have:*

$$\begin{aligned}\|\mathcal{I}_n\bar{u}_n^N - u_\infty\|_{\mathrm{L}^2(\Omega')} \leq C\Bigg(&\varepsilon_n\log(\varepsilon_n^{-\kappa}) + \varepsilon_n^{\kappa/5}(\mathcal{F}(u_0)-\mathcal{F}(u_\infty))^{1/2} \\ &+ \varepsilon_n^{-\kappa}\left[n^{-\alpha_1} + n^{-\alpha_2} + \frac{\log(\varepsilon_n^{-\kappa})^{(p-1)}}{\varepsilon_n^{1+p+\alpha_3}n^{\alpha_3}} + \tau_n\frac{\log(\varepsilon_n^{-\kappa})^{2p-3}}{\varepsilon_n^{2(1+p)}} + \theta_n\right]\Bigg)\end{aligned}$$

for some $C > 0$ that might be dependent on Ω (and d), u_0 and ℓ and with probability larger than

$$1 - \frac{C\log(\varepsilon_n^{-\kappa})^{2(p-1)}}{\theta_n^2\varepsilon_n^{1+p}n\rho_n}.$$

Furthermore, the latter right-hand side tends to 0 as $n\to\infty$ and the probability tends to 1.

4 Proofs

4.1 Well-posedness

Well-posedness of all our gradient flows is a central question. For the continuum nonlocal case, we will make use of nonlinear semigroup theory. The discrete case will follow from the continuum nonlocal case by using the interplay between the p-Laplacian and injection operators. Lastly, the continuum local case will follow from classical results.

4.1.1 Nonlocal problem

The following proposition, an extension of [47, Theorem 3.9], will allow us to characterize certain functions both by an equation they satisfy as well as a variational problem they minimize. The proof follows the above-mentioned reference but has been included in Section III.1 for completeness.

Proposition 4.1 (Dirichlet principles)**.** *Assume that Assumptions **S.1**, **K.1** and **O.1** hold. Let $p \geq 2$, $\mu > 0$ and $f \in \mathrm{L}^2(\Omega)$. Given $n \in \mathbb{N}$, $\lambda > 0$ and functions $u, \phi \in \mathrm{L}^p(\Omega)$, consider the equations*

$$\frac{|u|^{p-2}u}{n} + \lambda\left(\mu\Delta_p^K u + \mathcal{A}^*\mathcal{A}u - f\right) + u - \phi = 0 \tag{15}$$

and

$$\lambda(\mu\Delta_p^K u + \mathcal{A}^*\mathcal{A}u - f) + u - \phi = 0 \tag{16}$$

as well as their variational counterparts

$$E_{n,\lambda,\mathcal{A},f}(u) = \frac{1}{np}\int_\Omega |u|^p \,\mathrm{d}x + \frac{\lambda\mu}{2p}\int_{\Omega\times\Omega} K(|x-y|)|u(x)-u(y)|^p \,\mathrm{d}x\,\mathrm{d}y + \frac{\lambda}{2}\int_\Omega (\mathcal{A}u)^2 \,\mathrm{d}x - \lambda\int_\Omega fu\,\mathrm{d}x + \frac{1}{2}\int_\Omega (u-\phi)^2\,\mathrm{d}x \tag{17}$$

and

$$E_{\lambda,\mathcal{A},f}(u) = \frac{\lambda\mu}{2p}\int_\Omega\int_\Omega K(|x-y|)|u(x)-u(y)|^p \,\mathrm{d}y\mathrm{d}x + \frac{\lambda}{2}\int_\Omega (\mathcal{A}u)^2 \,\mathrm{d}x - \lambda\int_\Omega fu\,\mathrm{d}x + \frac{1}{2}\int_\Omega (u-\phi)^2\,\mathrm{d}x. \tag{18}$$

1. *If $u \in \mathrm{L}^p(\Omega)$ satisfies* (15) *λ_x-a.e., then we have $E_{n,\lambda,\mathcal{A},f}(u) \leq E_{n,\lambda,\mathcal{A},f}(v)$ for all $v \in \mathrm{L}^p(\Omega)$.*
2. *If for $u \in \mathrm{L}^p(\Omega)$ we have $E_{\lambda,\mathcal{A},f}(u) \leq E_{\lambda,\mathcal{A},f}(v)$ for all $v \in \mathrm{L}^p(\Omega)$, then u satisfies* (16) *λ_x-a.e..*

Remark 4.2 (Dirichlet principles). It is clear from the proof of Proposition 4.1 that the converse of the two statements of Proposition 4.1 can be shown analogously.

In the next step, our aim will be to show that if $u_n \rightharpoonup u$, we have $\liminf_{n\to\infty} E_{n,\lambda,\mathcal{A},f}(u_n) \geq E_{\lambda,\mathcal{A},f}(u)$, where $E_{n,\lambda,\mathcal{A},f}$ and $E_{\lambda,\mathcal{A},f}$ are respectively defined in (17) and (18).

Proposition 4.3 ($\liminf$-inequality for $E_{n,\lambda,\mathcal{A},f}$)**.** *Assume that Assumptions **S.1**, **K.1** and **O.1** hold. Let $p \geq 2$, $\mu > 0$, $f \in \mathrm{L}^2(\Omega)$, $\lambda > 0$ and $\phi \in \mathrm{L}^p(\Omega)$. Let $\{u_n\}_{n=1}^\infty \subset \mathrm{L}^p(\Omega)$ and $u \in \mathrm{L}^p(\Omega)$ be functions so that $u_n \rightharpoonup u$. Then,*

$$E_{\lambda,\mathcal{A},f}(u) \leq \liminf_{n\to\infty} E_{n,\lambda,\mathcal{A},f}(u_n)$$

where $E_{n,\lambda,\mathcal{A},f}$ and $E_{\lambda,\mathcal{A},f}$ are respectively defined in (17) *and* (18)*.*

Proof. In the proof $C > 0$ will denote a constant that can be arbitrarily large, (which might be) dependent on the kernel K, p, λ, μ, $\mathcal{A}$ and/or Ω, that may change from line to line.

We start by recalling (17) for $v \in \mathrm{L}^p(\Omega)$:

$$\begin{aligned} E_{n,\lambda,\mathcal{A},f}(v) &= \frac{1}{np}\int_\Omega |v|^p \,\mathrm{d}x + \frac{\lambda\mu}{2p}\int_{\Omega\times\Omega} K(|x-y|)|v(x)-v(y)|^p \,\mathrm{d}x\,\mathrm{d}y + \frac{\lambda}{2}\int_\Omega (\mathcal{A}v)^2 - 2fv\,\mathrm{d}x \\ &\quad + \frac{1}{2}\int_\Omega (v-\phi)^2\,\mathrm{d}x \\ &=: T_1(v) + T_2(v) + T_3(v) + T_4(v). \end{aligned}$$

For the T_1 term, by weak lower semi-continuity of norms, we have

$$\liminf_{n\to\infty} T_1(u_n) = \liminf_{n\to\infty} \frac{1}{np}\|u_n\|_{\mathrm{L}^p}^p = 0.$$

For the T_2 term, we first note that T_2 is proper and convex. Next, let $0 < r < R$. We claim that T_2 is bounded on the ball $B_{\mathrm{L}^p}(u,R)$. Indeed, for $v \in B_{\mathrm{L}^p}(u,R)$:

$$|T_2(v)| \leq C\int_\Omega\int_\Omega 2^{p-1}(|v(y)|^p + |v(x)|^p)\,\mathrm{d}y\,\mathrm{d}x \leq C(R + \|u\|_{\mathrm{L}^p}) \leq C.$$

Combining the above, by [52, Proposition 5.11], we have that T_2 is Lipschitz continuous on $B_{\mathrm{L}^p}(u,r)$ which in turn implies that T_2 is continuous and convex in $\mathrm{L}^p(\Omega)$. By [32, Corollary 2.2], we deduce that T_2 is weakly lower-semicontinuous. Analogously, T_3 and T_4 are convex and continuous and therefore weakly lower semi-continuous by [32, Corollary 2.2].

Collecting all the latter results, we obtain:

$$\liminf_{n\to\infty} E_{n,\lambda,\mathcal{A},f}(u_n) \geq T_2(u) + T_3(u) + T_4(u) = E_{\lambda,\mathcal{A},f}(u). \qquad \square$$

The next properties are easily checked: in particular, monotony and coercivity follow from [2, Lemma 2.3] and [17, Lemma 3.4 and Lemma 3.6]. For completeness, the proof may be found in Section III.1.

Lemma 4.4 (Properties of $\mathcal{E}_{n,\lambda,\mathcal{A},f}$)**.** *Assume Assumptions* **S.1**, **K.1** *and* **O.1** *hold. Let* $p \geq 2$, $\mu > 0$ *and* $f \in \mathrm{L}^2(\Omega)$. *For* $n \in \mathbb{N}$ *and* $\lambda > 0$, *we define the operator:*

$$\mathcal{E}_{n,\lambda,\mathcal{A},f}(u) = \frac{|u|^{p-2}u}{n} + u + \lambda(\mu\Delta_p^K u + \mathcal{A}^*\mathcal{A}u - f). \tag{19}$$

The following properties are satisfied:

1. *For* q *such that* $p^{-1} + q^{-1} = 1$, $\mathcal{E}_{n,\lambda,\mathcal{A},f} : \mathrm{L}^p(\Omega) \mapsto \mathrm{L}^q(\Omega)$ *and*

$$\|\mathcal{E}_{n,\lambda,\mathcal{A},f}(u)\|_{\mathrm{L}^q} \leq C\left(\left(1 + \frac{1}{n}\right)\|u\|_{\mathrm{L}^p}^{\frac{p}{q}} + \|u\|_{\mathrm{L}^p} + \|f\|_{\mathrm{L}^2}\right);$$

2. $\mathcal{E}_{n,\lambda,\mathcal{A},f}$ *is hemicontinuous, monotone and coercive.*

We now proceed to show a range condition on our evolution operator which will allow us to deduce the existence of a solution to the nonlocal problem.

Proof of Proposition 3.5. In the proof $C > 0$ will denote a constant that can be arbitrarily large, (which might be) dependent on the kernel K and/or Ω, that may change from line to line.

We begin by showing complete accretivity of $\mathcal{E}_{\mathcal{A},f}^K$. Let $u, v \in \mathrm{L}^p(\Omega)$, $h \in \mathcal{H}$ and consider:

$$\begin{aligned}\int_\Omega \left(\mathcal{E}_{\mathcal{A},f}^K(u) - \mathcal{E}_{\mathcal{A},f}^K(v)\right) h(u-v)\,\mathrm{d}x &= \mu \int_\Omega \left(\Delta_p^K u - \Delta_p^K v\right) h(u-v)\,\mathrm{d}x \\ &\quad + \int_\Omega (\mathcal{A}^*\mathcal{A}u - \mathcal{A}^*\mathcal{A}v)\, h(u-v)\,\mathrm{d}x \\ &=: T_1 + T_2.\end{aligned}$$

For the T_1 term, we obtain:

$$\begin{aligned}T_1 &= \mu \int_{\Omega\times\Omega} K(|x-y|)\Big(|u(x)-u(y)|^{p-2}(u(x)-u(y)) - |v(x)-v(y)|^{p-2}(v(x)-v(y))\Big) \\ &\qquad h(u(x)-v(x))\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{\mu}{2}\int_{\Omega\times\Omega} K(|x-y|)\Big(|u(x)-u(y)|^{p-2}(u(x)-u(y)) - |v(x)-v(y)|^{p-2}(v(x)-v(y))\Big) \\ &\qquad [h(u(x)-v(x)) - h(u(y)-v(y))]\,\mathrm{d}x\,\mathrm{d}y.\end{aligned}$$

Since h and $t \mapsto |t|^{p-2}t$ are both increasing then by splitting the latter equation in cases where $u(y) - v(y) \geq u(x) - v(x)$ (and conversely) we see that $T_1 \geq 0$. By Assumptions **O.1** and **O.2**, we know that $T_2 \geq 0$ and therefore, by Proposition 2.7 and Assumption **S.1**, $\mathcal{E}_{\mathcal{A},f}^K$ is completely accretive.

For the range condition, let $\phi \in \mathrm{L}^p(\Omega) \subseteq \mathrm{L}^q(\Omega)$ where $q = p/(p-1)$ since $p \geq 2$. We first show that there exists a solution to the equation

$$\frac{|u|^{p-2}u}{n} + u + \lambda\left(\mu\Delta_p^K u + \mathcal{A}^*\mathcal{A}u - f\right) = \mathcal{E}_{n,\lambda,\mathcal{A},f}(u) = \phi.$$

where $\mathcal{E}_{n,\lambda,\mathcal{A},f}$ is defined in (19). By Lemma 4.4, the operator $\mathcal{E}_{n,\lambda,\mathcal{A},f}$ satisfies all the conditions required to apply Theorem 2.11 and we deduce that for all $n \in \mathbb{N}$, there exists a function $u_n \in \mathrm{L}^p(\Omega)$ such that $\mathcal{E}_{n,\lambda,\mathcal{A},f}(u_n) = \phi$, or formulated otherwise: u_n satisfies (15).

Next, we claim that $u_n \overset{\mathrm{sgt}}{\ll} \phi + \lambda f$. Indeed, let $h \in \mathcal{H}$ and, using (15), we compute as follows:

$$\int_\Omega (\phi + \lambda f)\, h(u_n)\,\mathrm{d}x = \int_\Omega u_n h(u_n)\,\mathrm{d}x + \lambda\mu\int_\Omega \Delta_p^K u_n h(u_n)\,\mathrm{d}x$$

$$
\begin{aligned}
&+\frac{1}{n}\int_\Omega |u_n|^{p-2}u_n h(u_n)\,\mathrm{d}x + \lambda\int_\Omega \mathcal{A}^*\mathcal{A}u_n h(u_n)\,\mathrm{d}x \\
&\geq \int_\Omega u_n h(u_n)\,\mathrm{d}x
\end{aligned}
\tag{20}
$$

where we used the same argument as to show $T_1 \geq 0$ (using $v = 0$) to infer $\int_\Omega \Delta_p^K u_n h(u_n)\,\mathrm{d}x \geq 0$, the fact that $\mathrm{sign}(u_n) = \mathrm{sign}(h(u_n))$ and the Assumption **O.2** for (20). By Lemma II.2 and since $f \in \mathrm{L}^p(\Omega)$, we therefore obtain $u_n \overset{\mathrm{sgt}}{\ll} \phi + \lambda f \in \mathrm{L}^p(\Omega)$ and consequently $\|u_n\|_{\mathrm{L}^p} \leq \|\phi + \lambda f\|_{\mathrm{L}^p} \leq C$. From this, we deduce that the set $\{u_n\}_{n=1}^\infty$ is uniformly bounded in $\mathrm{L}^p(\Omega)$ and hence, there exists $u^* \in \mathrm{L}^p(\Omega)$ and a subsequence $\{u_{n_m}\}_{m=1}^\infty$ such that $u_{n_m} \rightharpoonup u^*$ in $\mathrm{L}^p(\Omega)$.

By Proposition 4.1, for each n, u_n minimizes $E_{n,\lambda,\mathcal{A},f}$ in $\mathrm{L}^p(\Omega)$. We will now show that u^* minimizes $E_{\lambda,\mathcal{A},f}$ in $\mathrm{L}^p(\Omega)$. Let $v \in \mathrm{L}^p(\Omega)$. In fact, by Proposition 4.3 we have

$$
E_{\lambda,\mathcal{A},f}(u^*) \leq \liminf_{m\to\infty} E_{n_m,\lambda,\mathcal{A},f}(u_{n_m}) \leq \liminf_{m\to\infty} E_{n_m,\lambda,\mathcal{A},f}(v) = E_{\lambda,\mathcal{A},f}(v).
$$

Therefore, by Proposition 4.1, u^* satisfies (16) which concludes the proof of the range condition. □

The straight-forward proof of the next result may be found in Section III.1.

Corollary 4.5 (Special cases of $\mathcal{E}_{\mathcal{A},f}^K$)**.** *Assume that Assumptions **S.1**, **O.1**, **O.2** and **K.1** hold. Let $p \geq 2$, $\mu > 0$ and assume that $f \in \mathrm{L}^p(\Omega)$. Then, the operators $\mathcal{E}_{\mathcal{A},f}^K$, $\mathcal{E}_{\mathcal{A},0}^K$ and $\mathcal{E}_{0,0}^K$ are m-completely accretive.*

Finally, we can deduce the following existence and uniqueness result.

Proof of Theorem 3.6. First, we note that if we find a solution to

$$
\begin{cases}
\frac{\partial}{\partial t}u + \mathcal{E}_{\mathcal{A},0}^K(u) = f \\
u(0,\cdot) = u_0,
\end{cases}
\tag{21}
$$

then, u solves (5) with the operator $\mathcal{E}_{\mathcal{A},f}^K$. From Corollary 4.5, we know that $\mathcal{E}_{\mathcal{A},0}^K$ is m-completely accretive. We can therefore apply Theorem 2.5 to deduce the existence of a unique solution u as in Definition 2.1. We obtain (9) by applying Lemma 2.8.

For the last part, we note that $v = 0$ solves (21) with $f = 0$ and $u_0 = 0$, so that by inserting $v = 0$ in (9), we obtain (10). □

The next result is a stability result for solutions to (5) which can be considered an extension of [33, Theorem 5.1] whose proof can be found in Section III.1.

Proposition 4.6 (Stability of solutions to (5))**.** *Assume that Assumptions **S.1**, **O.1** and **O.2** hold. Let $p \geq 2$, $\mu > 0$ and $T \geq 1$. Furthermore, for $i = 1, 2$, let K_i satisfy Assumption **K.1**, $u_{0,i} \in \mathrm{L}^p(\Omega)$ and $f_i \in \mathrm{L}^p(\Omega)$. Then, for $i = 1, 2$, there exists a unique solution u_i to the nonlocal problem* (5) *with evolution operator $\mathcal{E}_{\mathcal{A},f_i}^{K_i}$ and initial condition $u_{0,i}$ and we have the following stability estimates for some $C > 0$ dependent on Ω, u_0 and f_i:*

1. *if, for either $i = 1$ or $i = 2$, we have $u_{0,i} \in \mathrm{L}^{2(p-1)}(\Omega)$ and $f_i \in \mathrm{L}^{2(p-1)}(\Omega)$ then*

$$
\|u_2(t,\cdot) - u_1(t,\cdot)\|_{\mathrm{L}^2} \leq Ct^p \left(\sup_{x\in\Omega} \|K_2(x-\cdot) - K_1(x-\cdot)\|_{\mathrm{L}^2} + \|f_1 - f_2\|_{\mathrm{L}^2}\right) + \|u_{0,2} - u_{0,1}\|_{\mathrm{L}^2};
$$

2. *if, for either $i - 1$ or $i = 2$, we have $u_{0,i} \in \mathrm{L}^\infty(\Omega)$ and $f_i \in \mathrm{L}^\infty(\Omega)$ then*

$$
\|u_2(t,\cdot) - u_1(t,\cdot)\|_{\mathrm{L}^2} \leq Ct^p \left(\|K_2 - K_1\|_{\mathrm{L}^2} + \|f_1 - f_2\|_{\mathrm{L}^2}\right) + \|u_{0,2} - u_{0,1}\|_{\mathrm{L}^2}.
$$

Remark 4.7 (Generality of f in (5)). In Section 4.1, we have considered a general function f in (5). We could continue to do so below, but for ease of exposition, from now on, we will only consider $f = \mathcal{A}^*\ell$ or $f = \mathcal{P}_n\mathcal{A}^*\ell$.

4.1.2 Discrete nonlocal problem

As a corollary of Proposition 3.5, we can prove the well-posedness of the discrete problem (4) in the case that is of interest to us. The proof of the next corollary also displays the importance of accretivity of our operator.

Corollary 4.8 (Well-posedness of (4)). *Assume that Assumptions **S.1**, **O.1**, **O.2**, **O.3** and **K.1** hold. Let* $p \geq 2$, $\mu > 0$, $T > 0$, $u_0 \in \mathrm{L}^p(\Omega)$, $\ell \in \mathrm{L}^2(\Omega)$ *and assume that* $\mathcal{A}^*\ell \in \mathrm{L}^p(\Omega)$. *Furthermore, let* $n \in \mathbb{N}$ *and define* $\bar{K} = \mathcal{P}_n\tilde{K}$, $\bar{f} = \mathcal{P}_n\mathcal{A}^*\ell$, $\bar{u}_0 = \mathcal{P}_n u_0$. *Then, for any partition* $0 = t^0 < t^1 < \cdots < t^N = T$, *there exists a sequence* $\{\bar{u}_n^k\}_{k=0}^N$ *satisfying* (4) *with the above parameters that is well-defined and unique. We also have*

$$\|\mathcal{I}_n\bar{u}_n^k\|_{\mathrm{L}^r} \leq \|u_0\|_{\mathrm{L}^r} + T\|\mathcal{A}^*\ell\|_{\mathrm{L}^r} \tag{22}$$

for $1 \leq r \leq \infty$.

Proof. We first start by considering the well-posedness of the sequence $\{u_n^k\}_{k=0}^N$ defined iteratively by $u_n^0 = \mathcal{I}_n\bar{u}_0 = \mathcal{I}_n\mathcal{P}_n u_0$ and

$$(\mathrm{Id} + \tau^{k-1}\mathcal{E}_{\mathcal{A},\mathcal{I}_n\bar{f}}^{\mathcal{I}_n\bar{K}})(u_n^k) = u_n^k + \tau^{k-1}(\mu\Delta_p^{\mathcal{I}_n\bar{K}}u_n^k + \mathcal{A}^*\mathcal{A}u_n^k - \mathcal{I}_n\bar{f}) = u_n^{k-1}. \tag{23}$$

This can be reformulated as $u_n^k = (\mathrm{Id} + \tau^{k-1}\mathcal{E}_{\mathcal{A},0}^{\mathcal{I}_n\bar{K}})^{-1}(u_n^{k-1} + \tau^{k-1}\mathcal{I}_n\bar{f})$

By [33, Lemma 2.1], we have that $\|\mathcal{I}_n\bar{K}\|_{\mathrm{L}^\infty} \leq \|K\|_{\mathrm{L}^\infty} < \infty$, so we can apply Corollary 4.5 to deduce that $\mathcal{E}_{\mathcal{A},0}^{\mathcal{I}_n\bar{K}}$ is m-completely accretive. In particular, by Proposition 3.5 and [8, Section 2], for $\lambda > 0$, we have that $(\mathrm{Id} + \lambda\mathcal{E}_{\mathcal{A},0}^{\mathcal{I}_n\bar{K}})^{-1}$ is single-valued on $\mathrm{dom}((\mathrm{Id} + \lambda\mathcal{E}_{\mathcal{A},0}^{\mathcal{I}_n\bar{K}})^{-1}) = \mathrm{ran}(\mathrm{Id} + \lambda\mathcal{E}_{\mathcal{A},0}^{\mathcal{I}_n\bar{K}}) = \mathrm{L}^p(\Omega)$ and, for $1 \leq r \leq \infty$:

$$\|(\mathrm{Id} + \lambda\mathcal{E}_{\mathcal{A},0}^{\mathcal{I}_n\bar{K}})^{-1}(g_1) - (\mathrm{Id} + \lambda\mathcal{E}_{\mathcal{A},0})^{-1}(g_2)\|_{\mathrm{L}^r} \leq \|g_1 - g_2\|_{\mathrm{L}^r}$$

for any $g_i \in \mathrm{L}^p(\Omega)$. We have that $(\mathrm{Id} + \lambda\mathcal{E}_{\mathcal{A},0}^{\mathcal{I}_n\bar{K}})^{-1}(0) = 0$ so that, combining with the above,

$$\|(\mathrm{Id} + \lambda\mathcal{E}_{\mathcal{A},0}^{\mathcal{I}_n\bar{K}})^{-1}(g)\|_{\mathrm{L}^r} \leq \|g\|_{\mathrm{L}^r}$$

for any $g \in \mathrm{L}^p(\Omega)$.

Since by [33, Lemma 2.1], we have that $\|\mathcal{I}_n\bar{f}\|_{\mathrm{L}^r} \leq \|\mathcal{A}^*\ell\|_{\mathrm{L}^r}$ and $\|\mathcal{I}_n\bar{u}_0\|_{\mathrm{L}^p} \leq \|u_0\|_{\mathrm{L}^p}$, we can now proceed to show that u_n^k is well-posed by induction. For $k = 0$, we have that $u_n^k = \mathcal{I}_n\bar{u}_0 \in \mathrm{L}^p(\Omega)$. Now, assume that for $1 \leq m \leq k-1$, $u_n^m \in \mathrm{L}^p(\Omega)$ is well-defined. Then, $u_n^{k-1} + \tau^{k-1}\mathcal{I}_n\bar{f} \in \mathrm{L}^p(\Omega)$ and, since $(\mathrm{Id} + \tau^{k-1}\mathcal{E}_{\mathcal{A},0}^{\mathcal{I}_n\bar{K}})^{-1}$ is single-valued on $\mathrm{L}^p(\Omega)$, u_n^k is well-defined with

$$\begin{aligned}
\|u_n^k\|_{\mathrm{L}^r} &= \left\|\left(\mathrm{Id} + \tau^{k-1}\mathcal{E}_{\mathcal{A},0}^{\mathcal{I}_n\bar{K}}\right)^{-1}(u_n^{k-1} + \tau^{k-1}\mathcal{I}_n\bar{f})\right\|_{\mathrm{L}^r} \\
&\leq \left\|u_n^{k-1} + \tau^{k-1}\mathcal{I}_n\bar{f}\right\|_{\mathrm{L}^r} \\
&\leq \left\|u_n^{k-1}\right\|_{\mathrm{L}^r} + \tau^{k-1}\left\|\mathcal{I}_n\bar{f}\right\|_{\mathrm{L}^r}.
\end{aligned}$$

By induction

$$\|u_n^k\|_{\mathrm{L}^r} \leq \|u_n^0\|_{\mathrm{L}^r} + \sum_{m=0}^{k-1}\tau^m\|\mathcal{I}_n\bar{f}\|_{\mathrm{L}^r} = \|u_n^0\|_{\mathrm{L}^r} + t^k\|\mathcal{I}_n\bar{f}\|_{\mathrm{L}^r} \leq \|u_n^0\|_{\mathrm{L}^r} + T\|\mathcal{I}_n\bar{f}\|_{\mathrm{L}^r}.$$

The well-posedness implies that $\{u_n^k\}_{k=0}^N$ is the unique sequence that is defined iteratively by (23) and such that $u_n^0 = \mathcal{I}_n\bar{u}_0$.

Now, assume that there exists $\{\bar{u}_n^k\}_{k=0}^N$ that solves (4) with $\bar{f}$, $\bar{u}_0$, $\bar{G}_n$ and $\bar{K}$ as defined above. Then, we have $\mathcal{I}_n\bar{u}_0 = \mathcal{I}_n\mathcal{P}_n u_0 = u_n^0$ and

$$(\mathrm{Id} + \tau^{k-1}\mathcal{E}_{\mathcal{A},0}^{\mathcal{I}_n\bar{K}})\mathcal{I}_n\bar{u}_n^k = \mathcal{I}_n\bar{u}_n^k + \tau^{k-1}\left(\mathcal{I}_n(\mu\Delta_{p,n}^{\bar{K}}\bar{u}_n^k) + \mathcal{I}_n(\bar{G}_n(\bar{u}_n^k))\right) \tag{24}$$

$$= \sum_{i=1}^{|\Pi_n|}\chi_{\pi_i^n}\left[(\bar{u}_n^{k-1})_i + \tau^{k-1}(\bar{f})_i\right] \tag{25}$$

$$= \mathcal{I}_n \bar{u}_n^{k-1} + \tau^{k-1}\mathcal{I}_n \bar{f}$$

where we used Assumption **O.3** and [33, Lemma 6.1] for (24) and (4) for (25). By the uniqueness of the sequence $\{u_n^k\}_{k=0}^N$, we have that $u_n^k = \mathcal{I}_n \bar{u}_n^k$.

To conclude the proof, we show the existence of $\{\bar{u}_n^k\}_{k=0}^N$. First, recall from (4) that $\bar{u}_n^0 = \bar{u}_0$ and

$$\bar{u}_n^k = (\mathrm{Id} + \tau^{k-1}(\mu\Delta_{p,n}^{\bar{K}} + \bar{G}_n))^{-1}(\bar{u}_n^{k-1} + \tau^{k-1}\bar{f}).$$

By the same argument as above, we need to show that the operator $\mu\Delta_{p,n}^{\bar{K}} + \bar{G}_n$ is accretive on $\mathbb{R}^{|\Pi_n|}$ for $(\mathrm{Id} + \tau^{k-1}(\mu\Delta_{p,n}^{\bar{K}} + \bar{G}_n))^{-1}$ to be well-defined and unique. It is clear that $\mathrm{ran}(\Delta_{p,n}^{\bar{K}} + \bar{G}_n) \subseteq \mathbb{R}^{|\Pi_n|}$. By Proposition 2.3, we know that accretivity is equivalent to monotony in $\mathbb{R}^{|\Pi_n|}$ and it therefore only remains to verify that for $\bar{v}, \bar{w} \in \mathbb{R}^{|\Pi_n|}$:

$$\begin{aligned}\langle(\mu\Delta_{p,n}^{\bar{K}} + \bar{G}_n)(\bar{v}) - (\mu\Delta_{p,n}^{\bar{K}} + \bar{G}_n)(\bar{w}), \bar{v} - \bar{w}\rangle &= \langle\mu\Delta_{p,n}^{\bar{K}}(\bar{v}) - \mu\Delta_{p,n}^{\bar{K}}(\bar{w}), \bar{v} - \bar{w}\rangle \\ &\quad + \langle\bar{G}_n(\bar{v}) - \bar{G}_n(\bar{w}), \bar{v} - \bar{w}\rangle \\ &=: T_1 + T_2 \geq 0.\end{aligned}$$

The proof of $T_1 \geq 0$ is analogous to what we have shown for the continuum nonlocal Laplacian in the proof of Proposition 3.5; $T_2 \geq 0$ is due to Assumption **O.3**. □

4.1.3 Local problem

Theorem 4.9 (Well-posededness of (7)). *Assume that Assumptions **S.1** and **O.1** hold. Let* $p \geq 2$, $\mu > 0$, $T > 0$, $\ell \in \mathrm{L}^2(\Omega)$ *and* $u_0 \in \mathrm{L}^p(\Omega)$. *Then, there exists a unique weak solution* $u(t, x)$ *to the evolution problem* (7).

Proof. We begin the proof by verifying some properties of the operator $\mathcal{A}^*\mathcal{A}$. By Assumption **O.1**, $\mathcal{A}^*\mathcal{A}$ is linear and we know that $\mathcal{A}^*\mathcal{A}$ is bounded, hence continuous and therefore hemicontinuous by Lemma 2.10. Since $\mathcal{A}$ is linear, we can define the seminorm $\mathcal{S}(v) := \|\mathcal{A}v\|_{\mathrm{L}^2}$ on $\mathrm{L}^2(\Omega)$. By the boundedness of $\mathcal{A}$, we then note that $\mathcal{S}(\cdot) + \|\cdot\|_{\mathrm{L}^2}$ and $\|\cdot\|_{\mathrm{L}^2}$ are equivalent. Finally, we have that $\mathcal{A}^*\mathcal{A}$ is monotone since for $u, v \in \mathrm{L}^2(\Omega)$:

$$\langle\mathcal{A}^*\mathcal{A}u - \mathcal{A}^*\mathcal{A}v, u - v\rangle_{\mathrm{L}^2} = \langle\mathcal{A}^*\mathcal{A}(u - v), u - v\rangle_{\mathrm{L}^2} = \|\mathcal{A}(u - v)\|_{\mathrm{L}^2}^2 \geq 0.$$

We deduce that $\mathcal{A}^*\mathcal{A}$ satisfies the assumptions of Theorem 2.12 with $V_1 = H = V_1^* = \mathrm{L}^2(\Omega)$ and $p_1 = 2$.

We now define the operator $\mathcal{D}$ by

$$\mathcal{D}(u)(v) = \int_\Omega |\nabla u|^{p-2}\nabla u \cdot \nabla v \, \mathrm{d}x$$

for $u, v \in \mathrm{W}^{1,p}$. It is straight-forward (similarly to Lemma 4.4) to check that $\mathcal{D} : \mathrm{W}^{1,p} \mapsto (\mathrm{W}^{1,p})^*$ is bounded, hemicontinuous and satisfies $\|\mathcal{D}(u)\|_{(\mathrm{W}^{1,p})^*} \leq C\|u\|_{\mathrm{W}^{1,p}}^{p-1}$ as well as $(\mathcal{D}(u) - \mathcal{D}(v))(u - v) \geq 0$ for all $u, v \in \mathrm{W}^{1,p}$ (the last claim follows from [17, Lemma 3.6]). Furthermore, we define the Sobolev seminorm $[v] = \|\nabla v\|_{\mathrm{L}^p}$ and one can show that there exists $\lambda > 0$ such that $[v] + \lambda\|v\|_{\mathrm{L}^2} \geq \|v\|_{\mathrm{W}^{1,p}}$ (see for example [50, Chapter 2, Section 1.5.1]) for all $v \in \mathrm{W}^{1,p}$. It is also clear that $\mathcal{D}(v)(v) \geq [v]^p$. Hence $\mathcal{D}$ satisfies the assumptions of Theorem 2.12 with $V_2 = \mathrm{W}^{1,p}(\Omega) \subseteq \mathrm{L}^2(\Omega) \subseteq \mathrm{W}^{1,p}(\Omega)^*$ (where the inclusion is made possible by the fact that $p \geq 2$) and $p_2 = p$.

After an application of Theorem 2.12, we therefore obtain the existence of a unique function

$$u \in \mathrm{L}^p(0, T; \mathrm{W}^{1,p}(\Omega)) \cap \mathrm{L}^2(0, T; \mathrm{L}^2(\Omega)) \text{ and } u \in \mathrm{L}^\infty(0, T; \mathrm{L}^2(\Omega))$$

satisfying Definition 3.4 (see [50, Chapter 2, Section 1.5.1 or Example 1.7.2] for the treatment of the boundary term). □

4.2 Rates

We now turn our attention to establishing rates between our gradient flows defined on $[0, T]$. However, all of our rates will be expressed explicitly as a function of T as our aim will be to take $T \to \infty$ as to approximate the solution of (1). We begin with the continuum-to-continuum case.

4.2.1 Continuum nonlocal to local rates

The proof of the next result follows the same structure as the proof of Proposition 4.13 in order to apply Gronwall's lemma. The terms involving the application of $\Delta_p^{K^{\varepsilon_n}} - \Delta_p$ on a regular function give rise to rates through Taylor expansions as in [18, Theorem A.1]. For completeness, the full proof is included in Section III.2.

Theorem 4.10 (Continuum nonlocal-to-local rates)**.** *Assume that Assumptions **S.1**, **O.1**, **O.2**, **K.1** and **L.1** hold. Let $p \geq 2$, $\mu > 0$, $T > 0$, $\ell \in \mathrm{L}^2(\Omega)$, $u_0 \in \mathrm{L}^p(\Omega)$, Ω' be compactly contained in Ω and assume that $\mathcal{A}^*\ell \in \mathrm{L}^p(\Omega)$. Then, for all n, there exists a solution u_{ε_n} to* (5) *with kernel K_{ε_n} and $f = \mathcal{A}^*\ell$ and a solution u to* (7)*.*

*In addition, assume that Assumptions **S.2** and **K.2** hold, $p \geq 3$, that u_{ε_n} satisfies Assumption **R.1** and u satisfies Assumption **R.2**. Then, for n large enough, we have:*

$$\|u_{\varepsilon_n}(t,\cdot) - u(t,\cdot)\|_{\mathrm{L}^2(\Omega')} \leq \mathcal{O}\left(\varepsilon_n t C_1^{p-3}\left[C_1 + C_2^2\right]\right) \tag{26}$$

where $C_1 = \sup_{t\in(0,T)} \|\nabla u(t,\cdot)\|_{\mathrm{L}^\infty}$ and $C_2 = \sup_{t\in(0,T)} \|\nabla^2 u(t,\cdot)\|_{\mathrm{L}^\infty}$.

Remark 4.11 (Asymptotic rates in Theorem 4.10). In Theorem 4.10, the rates hold for n large enough, but this is not constraining in practice. Indeed, the latter condition is derived from Lemma I.1. This implies that the rates are relevant as soon as $\varepsilon_n < m/c$ where m is the minimum distance between the closure of the compactly contained set and the boundary of Ω and c is chosen so that $\mathrm{cl}(\Omega') \subset B(0,c)$.

4.2.2 Discrete-to-continuum nonlocal rates

Here we loosely follow [33, Section 6.2.2]. By using [33, Lemma 6.1], we easily check that the time interpolated version of the injected discrete problem satisfies an evolution problem. The proof can be found in Section III.2.

Lemma 4.12 (Evolution problem for $\mathrm{u}_{\mathrm{TimeInt}}$)**.** *Assume that Assumptions **S.1**, **O.1**, **O.2**, **O.3** and **K.1** hold. Let $p \geq 2$, $\mu > 0$, $T > 0$, $u_0 \in \mathrm{L}^p(\Omega)$, $\ell \in \mathrm{L}^2(\Omega)$ and assume that $\mathcal{A}^*\ell \in \mathrm{L}^p(\Omega)$. Furthermore, let $n \in \mathbb{N}$ and define $\bar{K} = \mathcal{P}_n\tilde{K}$, $\bar{f} = \mathcal{P}_n\mathcal{A}^*\ell$, $\bar{u}_0 = \mathcal{P}_n u_0$. Then, for any partition $0 = t^0 < t^1 < \cdots < t^N = T$, the sequence $\{\bar{u}_n^k\}_{k=0}^N$ is unique and well-defined by* (4) *with the above parameters. Furthermore, $\mathrm{u}_{\mathrm{TimeInt}}$ solves the following evolution problem:*

$$\begin{cases} \frac{\partial}{\partial t}\mathrm{u}_{\mathrm{TimeInt}} + \mu\Delta_p^{\mathcal{I}_n\bar{K}}(\mathrm{u}_{\mathrm{TimeInj}}) + \mathcal{A}^*\mathcal{A}(\mathrm{u}_{\mathrm{TimeInj}}) = \mathcal{I}_n\bar{f} & \text{in } (0,T)\times\Omega \\ \mathrm{u}_{\mathrm{TimeInt}}(0,\cdot) = \mathcal{I}_n\bar{u}_0. \end{cases}$$

While our final results in Theorems 3.7 and 3.10 will be concerned with letting $T \to \infty$, $\tau_n \to 0$ and $\varepsilon_n \to 0$, the next proposition presents more general non-asymptotic results that are valid for any kernel K, time-discretization τ and time T.

Proposition 4.13 (Discrete-to-continuum nonlocal rates)**.** *Assume that Assumptions **S.1**, **O.1**, **O.2**, **O.3** and **K.1** hold. Let $p \geq 2$, $\mu > 0$, $T > 0$, $u_0 \in \mathrm{L}^p(\Omega)$, $\ell \in \mathrm{L}^2(\Omega)$ and assume that $\mathcal{A}^*\ell \in \mathrm{L}^p(\Omega)$. Furthermore, let $n \in \mathbb{N}$ and $\bar{K} = \mathcal{P}_n\tilde{K}$, $\bar{f} = \mathcal{P}_n\mathcal{A}^*\ell$, $\bar{u}_0 = \mathcal{P}_n u_0$. Then, for any partition $0 = t^0 < t^1 < \cdots < t^N = T$, there exists a sequence $\{\bar{u}_n^k\}_{k=0}^N$ satisfying* (4) *with the above parameters. Furthermore, there exists a solution u_K to* (5)*. We also have the following rates for some $C > 0$ dependent on Ω, u_0 and $\mathcal{A}^*\ell$:*

1. *if $u_0 \in \mathrm{L}^{2p-2/(p-1)}(\Omega)$ and $\mathcal{A}^*\ell \in \mathrm{L}^{2p-2/(p-1)}(\Omega)$:*

$$\begin{aligned} \sup_{1\leq k\leq N}\sup_{t\in(t^{k-1},t^k]} \|\mathcal{I}_n\bar{u}_n^k - u_K(t,\cdot)\|_{\mathrm{L}^2} \leq Ce^{(1+C_{\mathrm{op}}^4)T}\Bigg(& \|\mathcal{I}_n\mathcal{P}_n u_0 - u_0\|_{\mathrm{L}^2} + \|\mathcal{I}_n\mathcal{P}_n\mathcal{A}^*\ell - \mathcal{A}^*\ell\|_{\mathrm{L}^2} \\ & + (1+T^{p-1})\sup_{x\in\Omega}\|\mathcal{I}_n\mathcal{P}_n K(|x-\cdot|) - K(|x-\cdot|)\|_{\mathrm{L}^2} + \tau(1+\|K\|_{\mathrm{L}^\infty})(1+T^{p-1}) \\ & + \tau^{p/(2p-1)}\|K\|_{\mathrm{L}^\infty}^{p/(2p-1)}(1+\|K\|_{\mathrm{L}^\infty})^{p/(2p-1)}(1+T^{p-1})^{p/(2p-1)}(1+T^{p-1-1/p})^{p/(2p-1)} \\ & + \tau^{(p+1)/(2p)}\|K\|_{\mathrm{L}^\infty}^{1/2}\left(1+T^{p-1-1/p}\right)^{1/2}(1+\|K\|_{\mathrm{L}^\infty})^{(p+1)/(2p)}(1+T^{p-1})^{(p+1)/(2p)}\Bigg); \end{aligned} \tag{27}$$

2. *if* $u_0 \in \mathrm{L}^\infty(\Omega)$ *and* $\mathcal{A}^*\ell \in \mathrm{L}^\infty(\Omega)$*:*

$$\sup_{1\leq k\leq N} \sup_{t\in(t^{k-1},t^k]} \|\mathcal{I}_n \bar{u}_n^k - u_K(t,\cdot)\|_{\mathrm{L}^2} \leq Ce^{(1+C_{\mathrm{op}}^4)T}\Bigg(\|\mathcal{I}_n\mathcal{P}_n u_0 - u_0\|_{\mathrm{L}^2} + \|\mathcal{I}_n\mathcal{P}_n\mathcal{A}^*\ell - \mathcal{A}^*\ell\|_{\mathrm{L}^2}$$

$$+ (1+T^{p-1})\|\mathcal{I}_n\mathcal{P}_n\tilde{K} - \tilde{K}\|_{\mathrm{L}^2(\Omega\times\Omega)} \tag{28}$$

$$+ \tau(1+\|K\|_{\mathrm{L}^\infty})(1+T^{p-1})\left[1+\|K\|_\infty(1+T^{p-2}) + \left(\|K\|_{\mathrm{L}^\infty}(1+T^{p-2})\right)^{1/2}\right]\Bigg).$$

Proof. In the proof $C>0$ will denote a constant that can be arbitrarily large, (which might be) dependent on Ω, u_0 or/and $\mathcal{A}^*\ell$, that may change from line to line.

The existence and well-posedness of $\{\bar{u}_n^k\}_{k=0}^N$ and u_K follow from Corollary 4.8 and Theorem 3.6 respectively. Also, note that for $p\geq 2$, we have $2p-2/(p-1)\geq 2(p-1)\geq p$. Let $\alpha_i>0$ for $1\leq i\leq 3$ be such that $\alpha_1+\mu(\alpha_2+\alpha_3)=\frac{1}{2}$.

We start by noticing that for $t\in(t^{k-1},t^k]$, we have

$$\|\mathrm{u}_{\mathrm{TimeInt}}(t,\cdot)\|_{\mathrm{L}^p} \leq \frac{t^k-t}{\tau^{k-1}}\|\mathcal{I}_n\bar{u}_n^{k-1}\|_{\mathrm{L}^p} + \frac{t-t^{k-1}}{\tau^{k-1}}\|\mathcal{I}_n\bar{u}_n^k\|_{\mathrm{L}^p} \leq C\left(\|u_0\|_{\mathrm{L}^p}+T\|\mathcal{A}^*\ell\|_{\mathrm{L}^p}\right) < C(C+TC)$$

where we used (22) for the second inequality and also:

$$\|\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot)\|_{\mathrm{L}^p} = \|\mathcal{I}_n\bar{u}_n^{k-1}\|_{\mathrm{L}^p} \leq \|u_0\|_{\mathrm{L}^p}+T\|\mathcal{A}^*\ell\|_{\mathrm{L}^p} < C+TC \tag{29}$$

again by (22). This, together with $\|\mathcal{I}_n\bar{u}_0\|_{\mathrm{L}^2}\leq\|u_0\|_{\mathrm{L}^p}$ by [33, Lemma 2.1], implies that for any $0\leq t<T$, we have that $\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot), \mathrm{u}_{\mathrm{TimeInt}}(t,\cdot)\in\mathrm{L}^p(\Omega)$ uniformly in t.

Define $\zeta_{\mathrm{TimeInt}}(t,x)=\mathrm{u}_{\mathrm{TimeInt}}(t,x)-u_K(t,x)$ and compute as follows:

$$\begin{aligned}
\frac{1}{2}\frac{\partial}{\partial t}\|\zeta_{\mathrm{TimeInt}}(t,\cdot)\|_{\mathrm{L}^2}^2 &= -\mu\int_\Omega\left(\Delta_p^{\mathcal{I}_n\bar{K}}\mathrm{u}_{\mathrm{TimeInj}}(t,x)-\Delta_p^{\mathcal{I}_n\bar{K}}u_K(t,x)\right)(\mathrm{u}_{\mathrm{TimeInj}}(t,x)-u_K(t,x))\,\mathrm{d}x\\
&\quad-\mu\int_\Omega\left(\Delta_p^{\mathcal{I}_n\bar{K}}\mathrm{u}_{\mathrm{TimeInj}}(t,x)-\Delta_p^{\mathcal{I}_n\bar{K}}u_K(t,x)\right)(\mathrm{u}_{\mathrm{TimeInt}}(t,x)-\mathrm{u}_{\mathrm{TimeInj}}(t,x))\,\mathrm{d}x\\
&\quad-\mu\int_\Omega\left(\Delta_p^{\mathcal{I}_n\bar{K}}u_K(t,x)-\Delta_p^{K}u_K(t,x)\right)\zeta_{\mathrm{TimeInt}}(t,x)\,\mathrm{d}x\\
&\quad-\int_\Omega\mathcal{A}^*\mathcal{A}(\mathrm{u}_{\mathrm{TimeInj}}(t,x)-u_K(t,x))\zeta_{\mathrm{TimeInt}}(t,x)\,\mathrm{d}x\\
&\quad+\int_\Omega(\mathcal{I}_n\bar{f}(x)-\mathcal{A}^*\ell(x))\zeta_{\mathrm{TimeInt}}(t,x)\,\mathrm{d}x\\
&=: T_1+\mu T_2+\mu T_3+T_4+T_5
\end{aligned}\tag{30}$$

where we used Lemma 4.12, (5) and the fact that $\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot),\mathrm{u}_{\mathrm{TimeInt}}(t,\cdot)\in\mathrm{L}^p(\Omega)$ for (30). Arguing as in Proposition 3.5 or Proposition 4.6 (which relies on [2, Lemma 2.3]), we obtain that $T_1\leq 0$. Furthermore, by Young's inequality for products,

$$T_5 \leq C\|\mathcal{I}_n\mathcal{P}_n\mathcal{A}^*\ell-\mathcal{A}^*\ell\|_2^2+\alpha_1\|\zeta_{\mathrm{TimeInt}}(t)\|_{\mathrm{L}^2}^2. \tag{31}$$

We continue our estimates by showing some auxiliary results first. For $t\in(t^{k-1},t^k]$:

$$\begin{aligned}
\|\mathrm{u}_{\mathrm{TimeInt}}(t,\cdot)-\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot)\|_{\mathrm{L}^2} &= |t^k-t|\left\|\frac{\mathcal{I}_n\bar{u}_n^k-\mathcal{I}_n\bar{u}_n^{k-1}}{\tau^{k-1}}\right\|_{\mathrm{L}^2}\\
&= |t^k-t|\left\|\mu\Delta_p^{\mathcal{I}_n\bar{K}}\mathcal{I}_n\bar{u}_n^k+\mathcal{A}^*\mathcal{A}\mathcal{I}_n\bar{u}_n^k-\mathcal{I}_n\bar{f}\right\|_{\mathrm{L}^2} &(32)\\
&\leq \tau\left[\mu\|\Delta_p^{\mathcal{I}_n\bar{K}}\mathcal{I}_n\bar{u}_n^k\|_{\mathrm{L}^2}+\|\mathcal{A}^*\mathcal{A}\mathcal{I}_n\bar{u}_n^k\|_{\mathrm{L}^2}+\|\mathcal{I}_n\bar{f}\|_{\mathrm{L}^2}\right] &(33)
\end{aligned}$$

where we used the proof of Corollary 4.8 for (32). Now, by Assumption **O.1** and (29) we have that

$$\|\mathcal{A}^*\mathcal{A}\mathcal{I}_n\bar{u}_n^k\|_{\mathrm{L}^2} = \|\mathcal{A}^*\mathcal{A}\,\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot)\|_{\mathrm{L}^2} \leq C\|\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot)\|_{\mathrm{L}^2} \leq C\|\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot)\|_{\mathrm{L}^p} \leq C(C+TC)$$

as well as $\|\mathcal{I}_n\bar{f}\|_{\mathrm{L}^2} \leq \|\mathcal{A}^*\ell\|_{\mathrm{L}^2} \leq C\|\mathcal{A}^*\ell\|_{\mathrm{L}^p} \leq C$ by [33, Lemma 2.1]. Furthermore, re-using [33, Lemma 2.1] and Assumption **S.1** we have

$$\|\Delta_p^{\mathcal{I}_n\bar{K}}\mathcal{I}_n\bar{u}_n^k\|_{\mathrm{L}^2} = \|\Delta_p^{\mathcal{I}_n\bar{K}}\,\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot)\|_{\mathrm{L}^2} \leq C\|K\|_{\mathrm{L}^\infty}\|\,\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot)\|_{\mathrm{L}^{2(p-1)}}^{p-1}. \tag{34}$$

Now, in either case, $u_0 \in \mathrm{L}^{2p-2/(p-1)}(\Omega)$ and $\mathcal{A}^*\ell \in \mathrm{L}^{2p-2/(p-1)}(\Omega)$ or $u_0 \in \mathrm{L}^\infty(\Omega)$ and $\mathcal{A}^*\ell \in \mathrm{L}^\infty(\Omega)$, we get from (22) that $\|\,\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot)\|_{\mathrm{L}^{2(p-1)}}^{p-1} \leq C + T^{p-1}C$ as was shown in (29). Injecting the above observation in (34) and then starting from (33), we obtain that:

$$\|\,\mathrm{u}_{\mathrm{TimeInt}}(t,\cdot) - \mathrm{u}_{\mathrm{TimeInj}}(t,\cdot)\|_{\mathrm{L}^2} \leq \tau(C + TC + \|K\|_{\mathrm{L}^\infty}(C + T^{p-1}C)). \tag{35}$$

For the T_4 term, by Young's inequality of products:

$$\begin{aligned}
|T_4| &\leq \frac{1}{2}\|\zeta_{\mathrm{TimeInt}}(t)\|_{\mathrm{L}^2}^2 + \frac{1}{2}\|\mathcal{A}^*\mathcal{A}(\mathrm{u}_{\mathrm{TimeInj}}(t) - u_K(t))\|_{\mathrm{L}^2}^2 \\
&\leq \frac{1}{2}\|\zeta_{\mathrm{TimeInt}}(t)\|_{\mathrm{L}^2}^2 + \frac{C_{\mathrm{op}}^4}{2}\|\,\mathrm{u}_{\mathrm{TimeInj}}(t) - \mathrm{u}_{\mathrm{TimeInt}} + \mathrm{u}_{\mathrm{TimeInt}} - u_K(t)\|_{\mathrm{L}^2}^2 \qquad (36)\\
&\leq \left(\frac{1}{2} + C_{\mathrm{op}}^4\right)\|\zeta_{\mathrm{TimeInt}}(t)\|_{\mathrm{L}^2}^2 + C\|\,\mathrm{u}_{\mathrm{TimeInj}}(t) - \mathrm{u}_{\mathrm{TimeInt}}\,\|_{\mathrm{L}^2}^2 \\
&\leq \left(\frac{1}{2} + C_{\mathrm{op}}^4\right)\|\zeta_{\mathrm{TimeInt}}(t)\|_{\mathrm{L}^2}^2 + \left[\tau(C + TC + \|K\|_{\mathrm{L}^\infty}(C + T^{p-1}C))\right]^2 \qquad (37)
\end{aligned}$$

where we used Assumption **O.1** for (36) and (35) for (37).

We will now tackle the T_2 and T_3 terms. First, assume that $u_0,\ \mathcal{A}^*\ell \in \mathrm{L}^{2p-2/(p-1)}(\Omega)$.We want to estimate $T_6 := \|\Delta_p^{\mathcal{I}_n\bar{K}}\,\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot) - \Delta_p^{\mathcal{I}_n\bar{K}}u_K(t,\cdot)\|_{\mathrm{L}^2}$:

$$\begin{aligned}
T_6^2 &\leq C\|\mathcal{I}_n\bar{K}\|_{\mathrm{L}^\infty}^2 \int_\Omega\int_\Omega \big|\,|\,\mathrm{u}_{\mathrm{TimeInj}}(y) - \mathrm{u}_{\mathrm{TimeInj}}(x)|^{p-2}(\mathrm{u}_{\mathrm{TimeInj}}(y) - \mathrm{u}_{\mathrm{TimeInj}}(x)) \\
&\qquad - |u_K(y) - u_K(x)|^{p-2}(u_K(y) - u_K(x))\big|^2\,\mathrm{d}y\mathrm{d}x \\
&\leq C\|K\|_{\mathrm{L}^\infty}^2 \int_\Omega\int_\Omega |\,\mathrm{u}_{\mathrm{TimeInj}}(y) - \mathrm{u}_{\mathrm{TimeInj}}(x) - u_K(y) + u_K(x)|^{2/p} \\
&\qquad \times (|\,\mathrm{u}_{\mathrm{TimeInj}}(y) - \mathrm{u}_{\mathrm{TimeInj}}(x)| + |u_K(y) - u_K(x)|)^{2(p-1)-2/p}\,\mathrm{d}y\mathrm{d}x \qquad (38)\\
&\leq C\|K\|_{\mathrm{L}^\infty}^2 \left[\int_\Omega\int_\Omega |\,\mathrm{u}_{\mathrm{TimeInj}}(y) - \mathrm{u}_{\mathrm{TimeInj}}(x) - u_K(y) + u_K(x)|^2\,\mathrm{d}y\mathrm{d}x\right]^{1/p} \\
&\qquad \times \left[\int_\Omega\int_\Omega (|\,\mathrm{u}_{\mathrm{TimeInj}}(y) - \mathrm{u}_{\mathrm{TimeInj}}(x)| + |u_K(y) - u_K(x)|)^{2p-2/(p-1)}\,\mathrm{d}y\mathrm{d}x\right]^{1/q} \qquad (39)\\
&\leq C\|K\|_{\mathrm{L}^\infty}^2 \left[\int_\Omega |\,\mathrm{u}_{\mathrm{TimeInj}}(x) - u_K(x)|^2\,\mathrm{d}x\right]^{1/p} \\
&\qquad \times \left[\int_\Omega\int_\Omega (|\,\mathrm{u}_{\mathrm{TimeInj}}(y)| + |\,\mathrm{u}_{\mathrm{TimeInj}}(x)| + |u_K(y)| + |u_K(x)|)^{2p-2/(p-1)}\,\mathrm{d}y\mathrm{d}x\right]^{1/q} \qquad (40)\\
&\leq C\|K\|_{\mathrm{L}^\infty}^2\|\,\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot) - u_K(t,\cdot)\|_{\mathrm{L}^2}^{2/p}\|\,|\,\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot)| + |u_K(t,\cdot)|\,\|_{\mathrm{L}^{2p-2/(p-1)}}^{2(p-1)-2/p} \qquad (41)\\
&\leq C\|K\|_{\mathrm{L}^\infty}^2\|\,\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot) - u_K(t,\cdot)\|_{\mathrm{L}^2}^{2/p}\left(C + T^{2(p-1)-2/p}C\right) \qquad (42)
\end{aligned}$$

where we used [33, Lemma 4.1, (ii)] with $\alpha = 1/p$ for (38), Hölder's inequality for (39), Assumption **S.1** for (40) and (41), and (22) as well as (10) with $r = 2p - 2/(p-1)$ for (42).

We now return to T_2 and estimate as follows:

$$\begin{aligned}
|T_2| &\leq \|\Delta_p^{\mathcal{I}_n\bar{K}}\,\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot) - \Delta_p^{\mathcal{I}_n\bar{K}}u_K(t,\cdot)\|_{\mathrm{L}^2}\|\,\mathrm{u}_{\mathrm{TimeInt}}(t,\cdot) - \mathrm{u}_{\mathrm{TimeInj}}(t,\cdot)\|_{\mathrm{L}^2} \\
&\leq \tau\|K\|_{\mathrm{L}^\infty}\left(C + TC + \|K\|_{\mathrm{L}^\infty}(C + T^{p-1}C)\right)\left(C + T^{(p-1)-1/p}C\right)\|\,\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot) - u_K(t,\cdot)\|_{\mathrm{L}^2}^{1/p} \qquad (43)\\
&\leq \tau\|K\|_{\mathrm{L}^\infty}\left(C + TC + \|K\|_{\mathrm{L}^\infty}(C + T^{p-1}C)\right)\left(C + T^{(p-1)-1/p}C\right)\Big[\|\,\mathrm{u}_{\mathrm{TimeInt}}(t,\cdot) - u_K(t,\cdot)\|_{\mathrm{L}^2}^{1/p}
\end{aligned}$$

$$+\left(\tau(C+TC+\|K\|_{\mathrm{L}^\infty}(C+T^{p-1}C))\right)^{1/p}\Big] \tag{44}$$

$$\leq \alpha_2\|\zeta_{\text{TimeInt}}(t)\|_{\mathrm{L}^2}^2 + \left[\tau\|K\|_{\mathrm{L}^\infty}\left(C+TC+\|K\|_{\mathrm{L}^\infty}(C+T^{p-1}C)\right)\left(C+T^{(p-1)-1/p}C\right)\right]^{2p/(2p-1)}$$

$$+\left[\tau\left(\|K\|_{\mathrm{L}^\infty}\left(C+T^{(p-1)-1/p}C\right)\right)^{p/(p+1)}\left(C+TC+\|K\|_{\mathrm{L}^\infty}(C+T^{p-1}C)\right)\right]^{(p+1)/p} \tag{45}$$

where we used (35) and (42) for (43), (35) again for (44) and Young's inequality for products for (45).

For T_3, relying on the fact that for $p \geq 2$ we have $2p-2/(p-1) \geq 2(p-1)$, we proceed as in Proposition 4.6 to obtain (9):

$$|T_3| \leq C(C+T^{p-1}C)\left[\sup_{x\in\Omega}\|\mathcal{I}_n\bar{K}(|x-\cdot|)-K(|x-\cdot|)\|_{\mathrm{L}^2}\right]\|\zeta_{\text{TimeInt}}(t)\|_{\mathrm{L}^2}$$

$$\leq (C+T^{2(p-1)}C)\sup_{x\in\Omega}\|\mathcal{I}_n\bar{K}(|x-\cdot|)-K(|x-\cdot|)\|_{\mathrm{L}^2}^2+\alpha_3\|\zeta_{\text{TimeInt}}(t)\|_{\mathrm{L}^2}^2 \tag{46}$$

using Young's inequality for products for (46).

Combining (45), (46), (37) and (31), we obtain:

$$\begin{aligned}\frac{\partial}{\partial t}\|\zeta_{\text{TimeInt}}(t)\|_{\mathrm{L}^2}^2 &\leq 2\left(\alpha_1+\mu(\alpha_2+\alpha_3)+\frac{1}{2}+C_{\mathrm{op}}^4\right)\|\zeta_{\text{TimeInt}}(t)\|_{\mathrm{L}^2}^2\\ &+\left[\tau(C+TC+\|K\|_{\mathrm{L}^\infty}(C+T^{p-1}C))\right]^2\\ &+\left[\tau\|K\|_{\mathrm{L}^\infty}\left(C+TC+\|K\|_{\mathrm{L}^\infty}(C+T^{p-1}C)\right)\left(C+T^{(p-1)-1/p}C\right)\right]^{2p/(2p-1)}\\ &+\left[\tau\left(\|K\|_{\mathrm{L}^\infty}\left(C+T^{(p-1)-1/p}C\right)\right)^{p/(p+1)}\left(C+TC+\|K\|_{\mathrm{L}^\infty}(C+T^{p-1}C)\right)\right]^{(p+1)/p}\\ &+(C+T^{2(p-1)}C)\sup_{x\in\Omega}\|\mathcal{I}_n\bar{K}(|x-\cdot|)-K(|x-\cdot|)\|_{\mathrm{L}^2}^2+C\|\mathcal{I}_n\mathcal{P}_n\mathcal{A}^*\ell-\mathcal{A}^*\ell\|_{\mathrm{L}^2}^2\\ &=:\left(2+2C_{\mathrm{op}}^4\right)\|\zeta_{\text{TimeInt}}(t)\|_{\mathrm{L}^2}^2+T_7.\end{aligned}$$

We continue by applying Gronwall's lemma on the latter to deduce:

$$\|\zeta_{\text{TimeInt}}(t)\|_{\mathrm{L}^2} \leq e^{\left(1+C_{\mathrm{op}}^4\right)T}\left(\|\mathcal{I}_n\mathcal{P}_nu_0-u_0\|_{\mathrm{L}^2}+C\cdot T_7^{1/2}\right). \tag{47}$$

We conclude that

$$\begin{aligned}\sup_{1\leq k\leq N}\sup_{t\in(t^{k-1},t^k]}\|\mathcal{I}_n\bar{u}_n^k-u_K(t,\cdot)\|_{\mathrm{L}^2} &= \sup_{0<t\leq T}\|\,\mathrm{u}_{\text{TimeInj}}(t,\cdot)-u_K(t,\cdot)\|_{\mathrm{L}^2}\\ &\leq \sup_{0<t\leq T}\|\,\mathrm{u}_{\text{TimeInt}}-u_K(t,\cdot)\|_{\mathrm{L}^2}+\sup_{0<t\leq T}\|\,\mathrm{u}_{\text{TimeInj}}-\mathrm{u}_{\text{TimeInt}}\,\|_{\mathrm{L}^2}\\ &\leq e^{\left(1+C_{\mathrm{op}}^4\right)T}\left(\|\mathcal{I}_n\mathcal{P}_nu_0-u_0\|_{\mathrm{L}^2}+C\cdot T_7^{1/2}\right)\\ &+\tau\left(C+TC+\|K\|_{\mathrm{L}^\infty}(C+T^{p-1}C)\right)\end{aligned} \tag{48}$$

where we used (47) and (35) for (48).

Let us now assume that $u_0 \in \mathrm{L}^\infty(\Omega)$ and $\mathcal{A}^*\ell \in \mathrm{L}^\infty(\infty)$. We will slightly change the estimates for T_6 and T_3.

$$\begin{aligned}T_6^2 &\leq C\|K\|_{\mathrm{L}^\infty}^2\int_\Omega\Big|\int_\Omega|\,\mathrm{u}_{\text{TimeInj}}(y)-\mathrm{u}_{\text{TimeInj}}(x)-u_K(y)+u_K(x)|\\ &\qquad\times(|\,\mathrm{u}_{\text{TimeInj}}(y)-\mathrm{u}_{\text{TimeInj}}(x)|+|u_K(y)-u_K(x)|)^{p-2}\,\mathrm{d}y\Big|^2\mathrm{d}x\end{aligned} \tag{49}$$

$$\begin{aligned}&\leq C\|K\|_{\mathrm{L}^\infty}^2\left(\|u_0\|_{\mathrm{L}^\infty}+T\|\mathcal{A}^*\ell\|_{\mathrm{L}^\infty}\right)^{2(p-2)}\\ &\times\int_\Omega\int_\Omega|\,\mathrm{u}_{\text{TimeInj}}(y)-\mathrm{u}_{\text{TimeInj}}(x)-u_K(y)+u_K(x)|^2\,\mathrm{d}y\mathrm{d}x\end{aligned} \tag{50}$$

$$\leq (C+T^{2(p-2)}C)\|K\|_{\mathrm{L}^\infty}^2\|\,\mathrm{u}_{\text{TimeInj}}(t,\cdot)-u_K(t,\cdot)\|_{\mathrm{L}^2}^2 \tag{51}$$

where we used [33, Lemma 4.1 (ii)] with $\alpha = 1$ and [33, Lemma 2.1] for (49), (10) and (22) for (50) and Assumption **S.1** for (51). We can then return to T_2 and estimate as follows:

(52)

$$|T_2| \leq \tau\|K\|_{\mathrm{L}^\infty}\left(C + TC + \|K\|_{\mathrm{L}^\infty}(C + T^{p-1}C)\right)(C + T^{(p-2)}C)\|\,\mathrm{u}_{\mathrm{TimeInj}}(t,\cdot) - u_K(t,\cdot)\|_{\mathrm{L}^2}$$

$$\leq \tau\|K\|_{\mathrm{L}^\infty}\left(C + TC + \|K\|_{\mathrm{L}^\infty}(C + T^{p-1}C)\right)(C + T^{(p-2)}C)\Big[\|\,\mathrm{u}_{\mathrm{TimeInt}}(t,\cdot) - u_K(t,\cdot)\|_{\mathrm{L}^2}$$

(53)

$$+ \tau(C + TC + \|K\|_{\mathrm{L}^\infty}(C + T^{p-1}C))\Big]$$

$$\leq \alpha_2\|\,\mathrm{u}_{\mathrm{TimeInt}}(t,\cdot) - u_K(t,\cdot)\|^2_{\mathrm{L}^2} + \left[\tau\|K\|_{\mathrm{L}^\infty}\left(C + TC + \|K\|_{\mathrm{L}^\infty}(C + T^{p-1}C)\right)(C + T^{(p-2)}C)\right]^2$$

(54)

$$+ \left[\left(\|K\|_{\mathrm{L}^\infty}(C + T^{(p-2)}C)\right)^{1/2}\tau\left(C + TC + \|K\|_{\mathrm{L}^\infty}(C + T^{p-1}C)\right)\right]^2$$

where we used (35) and (51) for (52), (35) again for (53) and Young's inequality for products for (54).

For T_3, we proceed as in Proposition 4.6 in order to obtain (11):

$$|T_3| \leq C(C + T^{p-1}C)\|\mathcal{I}_n\bar{K} - \tilde{K}\|_{\mathrm{L}^2(\Omega\times\Omega)}\|\zeta_{\mathrm{TimeInt}}(t)\|_{\mathrm{L}^2}$$

$$\leq (C + T^{2(p-1)}C)\|\mathcal{I}_n\bar{K} - \tilde{K}\|^2_{\mathrm{L}^2(\Omega\times\Omega)} + \alpha_3\|\zeta_{\mathrm{TimeInt}}(t)\|^2_{\mathrm{L}^2} \tag{55}$$

where we used Young's inequality for products for (55). We proceed as above to conclude. □

4.2.3 Discrete-to-continuum local rates

In this section we will derive rates between the fully discrete problem and the continuum problem. In particular, by combining the results of Proposition 4.13 and Theorem 4.10, one obtains general rates for fixed $T > 0$ and several classes of u_0, $\mathcal{A}^*\ell$.

Of greater interest is the question of the possibility of letting $T \to \infty$, as the solution of the gradient flow (7) solved for large T converges to the minimizer of the original regularization problem (1). Theorem 3.7 answers that question in a positive way: indeed, this can be achieved by correctly choosing the partition, the functions u_0, $\mathcal{A}^*\ell$ and the kernel K as well as indexing the time $T(n)$ in a meaningful way and imposing conditions between the time and space discretization parameters τ_n and ε_n.

Proof of Theorem 3.7. In the proof $C > 0$ will denote a constant that can be arbitrarily large, (which might be) dependent on Ω, u_0 or/and $\mathcal{A}^*\ell$, that may change from line to line. We also briefly comment on some notation: u_{ε_n} in (26) is a short-hand for $u_{K_{\varepsilon_n}}$ appearing in (28) when using the kernel K_{ε_n}.

We have

$$\|\mathcal{I}_n\bar{K}_{\varepsilon_n}\|_{\mathrm{L}^\infty} \leq \|\tilde{K}_{\varepsilon_n}\|_{\mathrm{L}^\infty} \leq \varepsilon_n^{-(d+p)}\|K\|_{\mathrm{L}^\infty} \tag{56}$$

by [33, Lemma 2.1]. The latter is finite by Assumption **K.1** and the existence claims then follow from Proposition 4.13 and Theorem 4.10.

When combining (28) and (26), we note that the constant is independent of T and therefore, we have that for large $T \geq 1$ and small $\varepsilon_n \leq e^{-1/\kappa} \leq 1$:

$$\begin{aligned}\sup_{1\leq k\leq N}\sup_{t\in(t^{k-1},t^k]}\|\mathcal{I}_n\bar{u}_n^k - u(t,\cdot)\|_{\mathrm{L}^2(\Omega')} &\leq Ce^{(1+C_{\mathrm{op}}^4)T}\Bigg(\|\mathcal{I}_n\mathcal{P}_nu_0 - u_0\|_{\mathrm{L}^2}\\ &+ \|\mathcal{I}_n\mathcal{P}_n\mathcal{A}^*\ell - \mathcal{A}^*\ell\|_2 + \frac{T^{(p-1)}}{\varepsilon_n^{d+p}}\|\mathcal{I}_n\mathcal{P}_n\tilde{K}(\cdot/\varepsilon_n) - \tilde{K}(\cdot/\varepsilon_n)\|_{\mathrm{L}^2(\Omega\times\Omega)}\\ &+ \tau\frac{T^{p-1}}{\varepsilon_n^{d+p}}\left[\frac{T^{p-2}}{\varepsilon_n^{d+p}} + \frac{T^{(p-2)/2}}{\varepsilon_n^{(d+p)/2}}\right]\Bigg) + CT\varepsilon_n\end{aligned} \tag{57}$$

$$\le Ce^{\left(1+C_{\mathrm{op}}^4\right)T}\left(n^{-\alpha_1}+n^{-\alpha_2}+\frac{T^{(p-1)}}{\varepsilon_n^{d+p+\alpha_3}n^{\alpha_3}}+\tau\frac{T^{(2p-3)}}{\varepsilon_n^{2(d+p)}}\right)+CT\varepsilon_n \tag{58}$$

$$=: C(T_1+T_2+T_3+T_4+T_5) \tag{59}$$

where we used Assumption **K.1** and Assumption **R.3** for (57), Lemma 2.14 and the fact that for $p \ge 3$,

$$2p-3 \ge \begin{cases} p-1 \\ p/2 \\ 3p/2-2 \end{cases}$$

for (58).

We now want all the terms in (59) to tend to 0. By choosing $T = T(n) = \left(1+C_{\mathrm{op}}^4\right)^{-1}\log(\varepsilon^{-\kappa})$ for some $\kappa > 0$, we now derive sufficient conditions on ε_n and $\tau = \tau_n$ such that $T_i \to 0$ for $1 \le i \le 5$.

We have that $\varepsilon_n \gg n^{-\alpha_1/\kappa}$ implies that $T_1 \to 0$. Indeed, from the latter, $1 \gg \varepsilon_n^{-\kappa}n^{-\alpha_1} = e^{\left(1+C_{\mathrm{op}}^4\right)T}n^{-\alpha_1}$. Analogously, $\varepsilon_n \gg n^{-\alpha_2/\kappa}$ implies that $T_2 \to 0$.

We have that $\varepsilon_n \gg \left[\mathrm{expW}\left(n^{\frac{\alpha_3}{\max(1+(d+p+\alpha_3)/\kappa,p-1)}}\right)\right]^{-1/\kappa}$ implies that $T_3 \to 0$. Indeed, similarly to the above, the latter is equivalent to

$$\begin{aligned} 1 &\gg \left[\varepsilon_n^{-\kappa}\log(\varepsilon_n^{-\kappa})\right]^{\max(1+(d+p+\alpha_3)/\kappa,p-1)}\left(1+C_{\mathrm{op}}^4\right)^{(1-p)}n^{-\alpha_3} \\ &\ge \frac{1}{\varepsilon_n^{\kappa+d+p+\alpha_3}}\log(\varepsilon_n^{-\kappa})^{(p-1)}\left(1+C_{\mathrm{op}}^4\right)^{(1-p)}n^{-\alpha_3} \\ &= T_3. \end{aligned}$$

We have that

$$\tau \ll \left(\frac{1}{1+C_{\mathrm{op}}^4}\right)^{(3-2p)}\frac{\varepsilon_n^{2(d+p)+\kappa}}{\log(\varepsilon_n^{-\kappa})^{(2p-3)}} = e^{-\left(1+C_{\mathrm{op}}^4\right)T}\frac{\varepsilon_n^{2(d+p)}}{T^{(2p-3)}}$$

which implies that $T_4 \to 0$.

Lastly, we have that $\lim_{n\to\infty}\varepsilon_n\log(\varepsilon_n^{-\kappa}) = 0$ implying that $T_5 \to 0$ which concludes the proof. □

Remark 4.14 (Generality of Theorem 3.7). The choice of $T(n)$ in Theorem 3.7 is arbitrary: by considering the general rates obtained by combining the results of Proposition 4.13 and Theorem 4.10, one could derive similar results to (12) with other $T(n)$. Furthermore, by combining (27) with (26), one could extend the results to more general u_0 and $\mathcal{A}^*\ell$.

Proof of Theorem 3.10. In the proof $C > 0$ will denote a constant that can be arbitrarily large, (which might be) dependent on Ω, u_0 or/and $\mathcal{A}^*\ell$, that may change from line to line.

We have:

$$\begin{aligned} \|\mathcal{I}_n\bar{u}_n^N - u_\infty\|_{\mathrm{L}^2(\Omega')} &\le \|\mathcal{I}_n\bar{u}_n^N - u(T,\cdot)\|_{\mathrm{L}^2(\Omega')} + \|u(T,\cdot)-u_\infty\|_{\mathrm{L}^2} \\ &\le \sup_{1\le k\le N}\sup_{t\in(t^{k-1},t^k]}\|\mathcal{I}_n\bar{u}_n^k - u(t,\cdot)\|_{\mathrm{L}^2(\Omega')} + \|u(T,\cdot)-u_\infty\|_{\mathrm{L}^2} \\ &=: T_1 + T_2. \end{aligned}$$

We note that $C_{\mathrm{op}} = 1$ since $\mathcal{A} = \mathrm{Id}$. For the T_1 term, we can apply Theorem 3.7. For the T_2 term, we note from Lemma I.3 and first order conditions that:

$$C\|u(T,\cdot)-u_\infty\|_{\mathrm{L}^2}^2 + \langle\nabla\mathcal{F}(u_\infty), u(T,\cdot)-u_\infty\rangle = \|u(T,\cdot)-u_\infty\|_{\mathrm{L}^2}^2 \le \mathcal{F}(u(T,\cdot)) - \mathcal{F}(u_\infty).$$

Furthermore, by standard considerations of gradient flows [59] we obtain that

$$\mathcal{F}(u(T,\cdot)) - \mathcal{F}(u_\infty) \le e^{-T}(\mathcal{F}(u_0)-\mathcal{F}(u_\infty)) = \varepsilon_n^{\kappa/2}(\mathcal{F}(u_0)-\mathcal{F}(u_\infty))$$

so that combining the latter yields the claim. □

Remark 4.15 (Asymptotic rates in Theorems 3.7 and 3.10.). The asymptotic aspect in Theorems 3.7 and 3.10 are not very restrictive: indeed, one part comes from Theorem 4.10 and discussed in Remark 4.11 while the second part comes from our estimates in Theorem 3.7. The latter are related to finding the smallest n such that $\varepsilon_n \leq e^{-1/\kappa}$ and $T \geq 1$ which for all reasonable choices of ε_n should occur for a very small n.

Remark 4.16 (g_n when $\mathcal{A} = \mathrm{Id}$). In Theorem 3.10, we pick $\mathcal{A} = \mathrm{Id}$ for the following two reasons: first, $C_{\mathrm{op}} = 1$ so that T_n can be explicitly defined; second, the Assumption **O.3** simplifies greatly as we can just pick $\bar{G}_n = \mathrm{Id}$ in this case. The latter part also works for any unitary operator $\mathcal{A}$ as discussed in Section 3.1.

4.3 Application to random graph models

Let us now consider the evolution problem (8). We recall that we will now be working on $(0, 1)$ with the uniform partition so that Assumptions **S.1** and **S.2** will always be satisfied and we have $d = 1$.

4.3.1 Well-posedness

Corollary 4.17 (Well-posedness of (8))**.** *Assume that Assumptions **S.1**, **O.1**, **O.2**, and **O.3** hold. Let $K : [0,\infty) \mapsto [0,\infty)$, $p \geq 2$, $\mu > 0$, $T > 0$, $u_0 \in \mathrm{L}^p(\Omega)$, $\ell \in \mathrm{L}^2(\Omega)$ and assume that $\mathcal{A}^*\ell \in \mathrm{L}^p(\Omega)$. Furthermore, let $n \in \mathbb{N}$ and define $\bar{K} = \mathcal{P}_n\tilde{K}$, $\tilde{K}(x,y) = K(|x-y|)$, $\bar{f} = \mathcal{P}_n\mathcal{A}^*\ell$, $\bar{u}_0 = \mathcal{P}_n u_0$. Then, $\mathbb{P}$-a.e., for any partition $0 = t^0 < t^1 < \cdots < t^N = T$, there exist a sequence $\{\bar{u}_n^k\}_{k=0}^N$ satisfying* (8) *with parameters $\bar{\Lambda}_n$, $\bar{f}$ and $\bar{u}_0$ that is well-defined and unique. We also have*

$$\|\mathcal{I}_n\bar{u}_n^k\|_{\mathrm{L}^r} \leq \|u_0\|_{\mathrm{L}^r} + T\|\mathcal{A}^*\ell\|_{\mathrm{L}^r}$$

for $1 \leq r \leq \infty$.

Proof. Since $\|\mathcal{I}_n\bar{\Lambda}_n\|_{\mathrm{L}^\infty}$ is bounded we proceed exactly as in Corollary 4.8 by considering (23) with kernel $\mathcal{I}_n\bar{\Lambda}_n$. □

4.3.2 Rates

As an intermediate step in establishing the rates between the random discrete and continuum local problems, we will have to compare random and deterministic discrete solutions which is what we discuss in the next proposition. The proof of the latter is similar to the proof of Proposition 4.13 but requires a few additional probabilistic estimates. Furthermore, on the results side, some of terms in the error bounds of Proposition 4.18 only differ from terms in (27) and (28) by having $\sup_{x\in\Omega}\|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} + \|\tilde{K}^{\varepsilon_n}\|_{\mathrm{L}^\infty}$ instead of $\|\tilde{K}^{\varepsilon_n}\|_{\mathrm{L}^\infty}$. For completeness, the full proof can be found in Section III.2.

Proposition 4.18 (Random-to-deterministic rates)**.** *Assume that Assumptions **O.1**, **O.2**, **O.3** and **K.1** hold. Let $p \geq 2$, $\mu > 0$, $T > 0$, $u_0 \in \mathrm{L}^p(\Omega)$, $\ell \in \mathrm{L}^2(\Omega)$ and assume that $\mathcal{A}^*\ell \in \mathrm{L}^p(\Omega)$. Furthermore, let $n \in \mathbb{N}$ and define $\bar{K}^{\varepsilon_n} = \mathcal{P}_n\tilde{K}^{\varepsilon_n}$, $\bar{f} = \mathcal{P}_n\mathcal{A}^*\ell$, $\bar{u}_0 = \mathcal{P}_n u_0$. We also suppose that ρ_n is a positive sequence with $\rho_n \to 0$ and $\rho_n \ll \varepsilon_n^{1+p}$. Let $\Lambda_n \in \mathbb{R}^{n\times n}$ be the weight matrix defined as in Definition 2.16 with $\bar{K} = \bar{K}^{\varepsilon_n}$.*

Then, for any partition $0 = t^0 < t^1 < \cdots < t^N = T$, there exists a sequence $\{\bar{v}_n^k\}_{k=0}^N$ satisfying (4) *with parameters $\bar{K}^{\varepsilon_n}$, $\bar{f}$ and $\bar{u}_0$. In addition, $\mathbb{P}$-a.e., there exists a sequence $\{\bar{u}_n^k\}_{k=0}^N$ solving* (8) *with parameters $\bar{\Lambda}_n$, $\bar{f}$ and $\bar{u}_0$.*

For any $\theta > 0$, we have the following rates with probability larger than $1 - \frac{(C+T^{2(p-1)}C)}{\theta^2 n\rho_n}\|\tilde{K}^{\varepsilon_n}\|_{\mathrm{L}^\infty}$ for some $C > 0$ dependent on Ω, u_0 and $\mathcal{A}^*\ell$:

1. if $u_0 \in \mathrm{L}^{2p-2/(2p-1)}(\Omega)$ and $\mathcal{A}^*\ell \in \mathrm{L}^{2p-2/(2p-1)}(\Omega)$:

$$\sup_{0\leq t\leq T}\|\,\mathrm{u}_{\mathrm{TimeInt}} - \mathrm{v}_{\mathit{TimeInt}}\,\|_{\mathrm{L}^2} \leq Ce^{\left(\frac{2+3C_{\mathrm{op}}^4}{2}\right)T}\Bigg(\theta + \tau^{p/(2p-1)}\sup_{x\in\Omega}\|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}^{p/(2p-1)}(1+T^{p-1-1/p})^{p/(2p-1)} \times (1+T^{p-1})^{p/(2p-1)}\left(1+\sup_{x\in\Omega}\|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} + \|\tilde{K}^{\varepsilon_n}\|_{\mathrm{L}^\infty}\right)^{p/(2p-1)}$$

$$+ \tau^{(p+1)/(2p)}(1+T^{p-1-1/p})^{1/2} \sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}^{1/2}$$

$$\times (1+T^{p-1})^{(p+1)/(2p)} \left(1 + \sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} + \|\tilde{K}^{\varepsilon_n}\|_{\mathrm{L}^\infty}\right)^{(p+1)/(2p)}$$

$$+ \tau(1+T^{p-1}) \left(1 + \sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} + \|\tilde{K}^{\varepsilon_n}\|_{\mathrm{L}^\infty}\right)\Bigg);$$

2. *if* $u_0 \in \mathrm{L}^\infty(\Omega)$ *and* $\mathcal{A}^*\ell \in \mathrm{L}^\infty(\Omega)$*:*

$$\sup_{0\le t\le T} \| \mathrm{u}_{\mathrm{TimeInt}} - \mathrm{v}_{\mathit{TimeInt}} \|_{\mathrm{L}^2} \le Ce^{\left(\frac{2+3C_{\mathrm{op}}^4}{2}\right)T}\Bigg(\theta$$

$$+ \tau(1+T^{p-1}) \left(1 + \sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} + \|\tilde{K}^{\varepsilon_n}\|_{\mathrm{L}^\infty}\right)$$

$$\times \left[1 + \left(\sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}(1+T^{p-2})\right)^{1/2} + \sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}(1+T^{p-2})\right]\Bigg).$$

Corollary 4.19 (Discrete random-to-continuum nonlocal rates)**.** *Assume that Assumptions **O.1**, **O.2**, **O.3** and **K.1** hold. Let* $p \ge 2$*,* $\mu > 0$*,* $T > 0$*,* $u_0 \in \mathrm{L}^p(\Omega)$*,* $\ell \in \mathrm{L}^2(\Omega)$ *and assume that* $\mathcal{A}^*\ell \in \mathrm{L}^p(\Omega)$*. Furthermore, let* $n \in \mathbb{N}$ *and define* $\bar{K}^{\varepsilon_n} = \mathcal{P}_n\tilde{K}^{\varepsilon_n}$*,* $\bar{f} = \mathcal{P}_n\mathcal{A}^*\ell$*,* $\bar{u}_0 = \mathcal{P}_n u_0$*. We also suppose that* ρ_n *is a positive sequence with* $\rho_n \to 0$ *and* $\rho_n \ll \varepsilon_n^{1+p}$*. Let* $\Lambda_n \in \mathbb{R}^{n\times n}$ *be the weight matrix defined as in Definition 2.16 with* $\bar{K} = \bar{K}^{\varepsilon_n}$*.*

Then, for any partition $0 = t^0 < t^1 < \cdots < t^N = T$*, there exists a sequence* $\{\bar{u}_n^k\}_{k=0}^N$ *solving* (8) *with parameters* $\bar{\Lambda}_n$*,* $\bar{f}$ *and* $\bar{u}_0$*, and a solution* u_{ε_n} *to* (5) *with kernel* K^{ε_n}*.*

For any $\theta > 0$*, we have the following rates with probability larger than* $1 - \frac{(C+T^{2(p-1)}C)}{\theta^2 n\rho_n}\|\tilde{K}^{\varepsilon_n}\|_{\mathrm{L}^\infty}$ *for some* $C > 0$ *dependent on* Ω*,* u_0 *and* $\mathcal{A}^*\ell$*:*

1. *if* $u_0 \in \mathrm{L}^{2p-2/(2p-1)}(\Omega)$ *and* $\mathcal{A}^*\ell \in \mathrm{L}^{2p-2/(2p-1)}(\Omega)$*, then:*

$$\sup_{1\le k\le N} \sup_{t\in(t^{k-1},t^k]} \|\mathcal{I}_n\bar{u}_n^k - u_{\varepsilon_n}\|_{\mathrm{L}^2} \le \tau C(1+T^{p-1}) \left(1 + \sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} + \frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)$$

$$+ Ce^{\left(\frac{2+3C_{\mathrm{op}}^4}{2}\right)T}\Bigg(\theta$$

$$+ \tau^{p/(2p-1)} \sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}^{p/(2p-1)}(1+T^{p-1-1/p})^{p/(2p-1)}$$

$$\times (1+T^{p-1})^{p/(2p-1)} \left(1 + \sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} + \frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)^{p/(2p-1)}$$

$$+ \tau^{(p+1)/(2p)}(1+T^{p-1-1/p})^{1/2} \sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}^{1/2}$$

$$\times (1+T^{p-1})^{(p+1)/(2p)} \left(1 + \sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} + \frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)^{(p+1)/(2p)}$$

$$+ \tau(1+T^{p-1}) \left(1 + \sup_{x\in\Omega} \|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} + \frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)\Bigg)$$

(60)

$$+ Ce^{(1+C_{\mathrm{op}}^4)T}\Bigg(\|\mathcal{I}_n\mathcal{P}_n u_0 - u_0\|_{\mathrm{L}^2} + \|\mathcal{I}_n\mathcal{P}_n\mathcal{A}^*\ell - \mathcal{A}^*\ell\|_{\mathrm{L}^2}$$

$$+ \frac{(1+T^{p-1})}{\varepsilon_n^{1+p}} \sup_{x\in\Omega} \|\mathcal{I}_n\mathcal{P}_n K(|x-\cdot|/\varepsilon_n) - K(|x-\cdot|/\varepsilon_n)\|_{\mathrm{L}^2} + \tau\left(1 + \frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)(1+T^{p-1})$$

$$+ \tau^{p/(2p-1)} \left(\frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)^{p/(2p-1)} \left(1 + \frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)^{p/(2p-1)} (1+T^{p-1})^{p/(2p-1)}(1+T^{p-1-1/p})^{p/(2p-1)}$$

$$+\tau^{(p+1)/(2p)}\left(\frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)^{1/2}\left(1+T^{p-1-1/p}\right)^{1/2}\left(1+\frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)^{(p+1)/(2p)}(1+T^{p-1})^{(p+1)/(2p)}\Bigg);$$

2. *if* $u_0 \in \mathrm{L}^\infty(\Omega)$ *and* $\mathcal{A}^*\ell \in \mathrm{L}^\infty(\Omega)$, *then:*

$$\sup_{1\le k\le N}\sup_{t\in(t^{k-1},t^k]}\|\mathcal{I}_n\bar{u}_n^k-u_{\varepsilon_n}\|_{\mathrm{L}^2}\le Ce^{\left(\frac{2+3C_{\mathrm{op}}^4}{2}\right)T}\Bigg(\theta$$
$$+\tau(1+T^{p-1})\left(1+\sup_{x\in\Omega}\|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}+\frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)$$
$$\times\left[1+\left(\sup_{x\in\Omega}\|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}(1+T^{p-2})\right)^{1/2}+\sup_{x\in\Omega}\|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}(1+T^{p-2})\right]\Bigg)$$
$$+Ce^{(1+C_{\mathrm{op}}^4)T}\Bigg(\|\mathcal{I}_n\mathcal{P}_nu_0-u_0\|_{\mathrm{L}^2}+\|\mathcal{I}_n\mathcal{P}_n\mathcal{A}^*\ell-\mathcal{A}^*\ell\|_{\mathrm{L}^2}\tag{61}$$
$$+\frac{(1+T^{p-1})}{\varepsilon_n^{d+1}}\|\mathcal{I}_n\mathcal{P}_n\tilde{K}(\cdot/\varepsilon_n)-\tilde{K}(\cdot/\varepsilon_n)\|_{\mathrm{L}^2(\Omega\times\Omega)}$$
$$+\tau\left(1+\frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)(1+T^{p-1})\left[1+\frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}(1+T^{p-2})+\left(\frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}(1+T^{p-2})\right)^{1/2}\right]\Bigg).$$

Proof. In the proof $C>0$ will denote a constant that can be arbitrarily large, (which might be) dependent on Ω, u_0 or/and $\mathcal{A}^*\ell$, that may change from line to line.

The existence of $\{\bar{u}_n^k\}_{k=0}^N$ follows from Proposition 4.18. Let $\{\bar{v}_n^k\}_{k=0}^N$ be as in Proposition 4.18 as well. The existence of u_{ε_n} follows from Theorem 3.7. We note that

$$\begin{aligned}\sup_{1\le k\le N}\sup_{t\in(t^{k-1},t^k]}\|\mathcal{I}_n\bar{u}_n^k-u_{\varepsilon_n}\|_{\mathrm{L}^2}&=\sup_{0<t\le T}\|\,\mathrm{u_{TimeInj}}-u_{\varepsilon_n}\|_{\mathrm{L}^2}\\&\le\sup_{0<t\le T}\|\,\mathrm{u_{TimeInj}}-\mathrm{u_{TimeInt}}\,\|_{\mathrm{L}^2}+\sup_{0<t\le T}\|\,\mathrm{u_{TimeInt}}-\mathrm{v_{TimeInt}}\,\|_{\mathrm{L}^2}\\&\quad+\sup_{0<t\le T}\|\,\mathrm{v_{TimeInt}}-\mathrm{v_{TimeInj}}\,\|_{\mathrm{L}^2}+\sup_{0<t\le T}\|\,\mathrm{v_{TimeInj}}-u_{\varepsilon_n}\|_{\mathrm{L}^2}\\&\le\tau\left(C+CT+(C+T^{p-1}C)\left[\sup_{x\in\Omega}\|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}+\|\tilde{K}^{\varepsilon_n}\|_{\mathrm{L}^\infty}\right]\right)\\&\quad+\sup_{0<t\le T}\|\,\mathrm{u_{TimeInt}}-\mathrm{v_{TimeInt}}\,\|_{\mathrm{L}^2}+\sup_{0<t\le T}\|\,\mathrm{v_{TimeInj}}-u_{\varepsilon_n}\|_{\mathrm{L}^2}\end{aligned}\tag{62}$$

where we used (21) for (62).

Relying on (56), it is possible to apply Proposition 4.13 with the kernel $\bar{K}^{\varepsilon_n}$ and obtain the same bounds (with the scaling factor $\varepsilon_n^{-(1+p)}$). Combining the latter with Proposition 4.18 we obtain (60) and (61). □

Similarly to the setting in Section 4.2.3, we can combine the results of Corollary 4.19 and Theorem 4.10 to obtain precise rates for fixed $T>0$. Recalling the discussion in Remark 4.14, the analogous results to Theorems 3.7 and 3.10 are Theorems 3.12 and 3.13. Before proceeding to the proof of the latter two results, we present two probabilistic lemmas that are essential for the establishment of the rates.

First, we start a variant of [33, Lemma 3.1], by explicitly writing out the estimates for the scaled kernels. In this section, we recall that we will be working on $(0,1)$ with the uniform partition so that Assumptions **S.1** and **S.2** will always be satisfied and we have $d=1$. In this setting the operators $\mathcal{I}_n$ and $\mathcal{P}_n$ can be defined

$$(\mathcal{I}_nu)(x)=\sum_{i\in[n]}u_i\chi_{\Omega_{n,i}}(x)\qquad(\mathcal{I}_nu)(x,y)=\sum_{i\in[n],j\in[n]}u_{ij}\chi_{\Omega_{n,i}}(x)\chi_{\Omega_{n,j}}(y)$$
$$(\mathcal{P}_nu)_i=n\int_{\Omega_{n,i}}u(x)\,\mathrm{d}x\qquad(\mathcal{P}_nu)_{ij}=n^2\int_{\Omega_{n,i}\times\Omega_{n,j}}u(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

where $\Omega_{n,i}=(\frac{i-1}{n},\frac{i}{n})$.

Lemma 4.20 (Convergence of the random weight matrix)**.** *Assume K satisfies Assumption **K.1**, $p>1$, $d=1$, ρ_n and ε_n are positive sequences with $\varepsilon_n \to 0$ and $\rho_n \to 0$ and define $\tilde{K}_{\varepsilon_n}(x,y) = \frac{2}{c(p,1)\varepsilon_n^{p+1}} K\left(\frac{|x-y|}{\varepsilon}\right)$, where $c(p,1)$ is defined by* (6)*. Define $\bar{K}^{\varepsilon_n} = \mathcal{P}_n \tilde{K}_{\varepsilon_n}$ and let $\Lambda_n \in \mathbb{R}^{n\times n}$ be the weight matrix defined as in Definition 2.16 with $\bar{K} = \bar{K}^{\varepsilon_n}$. Assume that*

$$\frac{\log(n)\varepsilon_n^{2p}}{n} \ll \rho_n^2 \ll \varepsilon_n^{p+1}.$$

Then, with probability one, for n large enough, we have

$$|\sup_{x\in\Omega} \|\mathcal{I}_n \bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} - \sup_{x\in\Omega} \|\mathcal{I}_n \bar{K}_{\varepsilon_n}(x,\cdot)\|_{\mathrm{L}^1}| < \varepsilon_n^{-p}. \tag{63}$$

Furthermore, with probability one, for n large enough, we have

$$\sup_{x\in\Omega} \|\mathcal{I}_n \bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} \le \frac{C}{\varepsilon_n^p}.$$

Proof. We can estimate as follows:

$$\begin{aligned}
&\mathbb{P}\left(|\sup_{x\in\Omega} \|\mathcal{I}_n \bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} - \sup_{x\in\Omega} \|\mathcal{I}_n \bar{K}^{\varepsilon_n}(x,\cdot)\|_{\mathrm{L}^1}| > \varepsilon_n^{-p}\right) \\
&\quad = \mathbb{P}\left(|\max_{i\in[n]} \frac{1}{n}\sum_{j\in[n]} \bar{\Lambda}_{n,ij} - \max_{i\in[n]} \frac{1}{n}\sum_{j\in[n]} \bar{K}_{ij}^{\varepsilon_n}| > \varepsilon_n^{-p}\right) \\
&\quad \le \mathbb{P}\left(\max_{i\in[n]} |\frac{1}{n}\sum_{j\in[n]} \left(\bar{\Lambda}_{n,ij} - \bar{K}_{ij}^{\varepsilon_n}\right)| > \varepsilon_n^{-p}\right) \\
&\quad \le \sum_{i\in[n]} \mathbb{P}\left(|\frac{1}{n}\sum_{j\in[n]} \left(\bar{\Lambda}_{n,ij} - \bar{K}_{ij}^{\varepsilon_n}\right)| > \varepsilon_n^{-p}\right) \\
&\quad \le 2\sum_{i\in[n]} \exp\left(-\frac{\frac{1}{2}n^2\varepsilon_n^{-2p}}{\sum_{j\in[n]} \frac{\bar{K}_{ij}^{\varepsilon_n}}{\rho_n}(1-\rho_n\bar{K}_{ij}^{\varepsilon_n}) + \frac{n}{3\varepsilon_n^p}(\frac{1}{\rho_n} + \frac{C}{\varepsilon_n^{p+1}})}\right) \qquad (64) \\
&\quad \le 2\sum_{i\in[n]} \exp\left(-\frac{cn^2\varepsilon_n^{-2p}}{\frac{n}{\rho_n^2} + \frac{n}{\varepsilon_n^p}(\frac{1}{\rho_n} + \frac{1}{\varepsilon_n^{p+1}})}\right) \\
&\quad \le 2\sum_{i\in[n]} \exp\left(-\frac{cn\varepsilon_n^{-2p}\rho_n^2}{1 + \frac{\rho_n}{\varepsilon_n^p} + \frac{\rho_n^2}{\varepsilon_n^{2p+1}}}\right) \\
&\quad \le 2\sum_{i\in[n]} \exp\left(-\frac{cn\rho_n^2}{\varepsilon_n^{2p}}\right)
\end{aligned}$$

where we used Bernstein's lemma for (64) after noticing that $\mathbb{E}[\bar{\Lambda}_{n,ij} - \bar{K}_{ij}^{\varepsilon_n}] = 0$, $\left|\bar{\Lambda}_{n,ij} - \bar{K}_{ij}^{\varepsilon_n}\right| \le \frac{1}{\rho_n} + \frac{C}{\varepsilon_n^{p+1}}$ and $\mathbb{E}[\bar{\Lambda}_{n,ij} - \bar{K}_{ij}^{\varepsilon_n}]^2 = \frac{\bar{K}_{ij}^{\varepsilon_n}}{\rho_n}(1-\rho_n\bar{K}_{ij}^{\varepsilon_n}) \le \frac{1}{4\rho_n^2}$. Choosing $\gamma > 2$ such that $\frac{cn\rho_n^2}{\varepsilon_n^{2p}} \ge \gamma\log(n)$ we have

$$\mathbb{P}\left(|\sup_{x\in\Omega} \|\mathcal{I}_n \bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} - \sup_{x\in\Omega} \|\mathcal{I}_n \bar{K}^{\varepsilon_n}(x,\cdot)\|_{\mathrm{L}^1}| > \varepsilon_n^{-p}\right) \le 2\sum_{i\in[n]} n^{-\gamma}$$

which is summable. By the Borel-Cantelli Lemma, with probability one, for all but finitely many n,

$$|\sup_{x\in\Omega} \|\mathcal{I}_n \bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} - \sup_{x\in\Omega} \|\mathcal{I}_n \bar{K}^{\varepsilon_n}(x,\cdot)\|_{\mathrm{L}^1}| < \varepsilon_n^{-p}.$$

For the furthermore part of the lemma we note that we can write

$$
\begin{aligned}
\left\|\mathcal{I}_n\bar{K}^{\varepsilon_n}(x,\cdot)\right\|_{\mathrm{L}^1} &= \int_\Omega \mathcal{I}_n\bar{K}^{\varepsilon_n}(x,y)\,\mathrm{d}y \\
&= \int_\Omega \sum_{i.j=1}^n (\mathcal{P}_n\tilde{K}^{\varepsilon_n})_{i,j}\chi_{\Omega_{n,i}}(x)\chi_{\Omega_{n,j}}(y)\,\mathrm{d}y \\
&= \frac{n^2C(p,1)}{\varepsilon_n^{p+1}}\int_\Omega \sum_{i,j=1}^n \int_{\Omega_{n,i}\times\Omega_{n,j}} K\left(\frac{|w-z|}{\varepsilon_n}\right)\mathrm{d}w\,\mathrm{d}z\,\chi_{\Omega_{n,i}}(x)\chi_{\Omega_{n,j}}(y)\,\mathrm{d}y \\
&= \frac{nC(p,1)}{\varepsilon_n^{p+1}}\sum_{i,j=1}^n \int_{\Omega_{n,i}\times\Omega_{n,j}} K\left(\frac{|w-z|}{\varepsilon_n}\right)\mathrm{d}w\,\mathrm{d}z\,\chi_{\Omega_{n,i}}(x).
\end{aligned}
$$

So,

$$
\sup_{x\in\Omega}\|\mathcal{I}_n\bar{K}^{\varepsilon_n}(x,\cdot)\|_{\mathrm{L}^1} \le \frac{C(p,1)}{\varepsilon_n^{p+1}}\int_{\mathbb{R}} K\left(\frac{|z|}{\varepsilon}\right)\mathrm{d}z = \frac{C(p,1)}{\varepsilon_n^{p}}\int_{\mathbb{R}} K(|z|)\,\mathrm{d}z.
$$

Combining with the first part of the lemma, this completes the proof. □

Remark 4.21. Asymptotic rates and Borel-Cantelli arguments. The asymptotic claim in (63), i.e. the existence of some N such that for all $n \ge N$ (63) holds, comes from a Borel-Cantelli argument. All we know is that, with probability one, $N < \infty$. This can be circumvented with the following trade-off: either one argues with a Borel-Cantelli Lemma and obtains an $\mathbb{P}$-a.e. statement with an asymptotic part or one does not and is then left with a claim holding with high probability.

Similarly to what was discussed in Section 4.2.3 we want to find conditions under which we will be able to take the right-hand side of (61) to 0. To that purpose, we explicit a choice of $\theta = \theta_n$ that is compatible with the conditions derived in Lemma 4.20.

Lemma 4.22 (Rates for θ_n)**.** *Let $d = 1$, $p > 1$, ρ_n and ε_n be positive sequences with $\varepsilon_n \to 0$ and $\rho_n \to 0$. For some $\kappa > 0$, let $T(n) = \left(\frac{2}{2+3C_{\mathrm{op}}^4}\right)\log(\varepsilon_n^{-\kappa})$. Assume that*

$$
\varepsilon_n \gg \left[\exp\mathrm{W}\left((n\log(n))^{1/\max(4(p-1),4+(2+4p)/\kappa)}\right)\right]^{-1/\kappa}.
$$

Then, $\frac{[\log(\varepsilon_n^{-\kappa})]^{2(p-1)}}{\varepsilon_n^{1+2p}\log(n)} \ll \varepsilon_n^{2\kappa}$. Moreover, for a positive sequence θ_n satisfying

$$
\frac{[\log(\varepsilon_n^{-\kappa})]^{2(p-1)}}{\varepsilon_n^{1+2p}\log(n)^{1/2}n^{1/2}} \ll \theta_n^2 \ll \varepsilon_n^{2\kappa},
$$

and assuming

$$
\frac{\log(n)\varepsilon_n^{2p}}{n} \ll \rho_n^2
$$

we have that $e^{\left(\frac{2+3C_{\mathrm{op}}^4}{2}\right)T}\theta_n \ll 1$ and $\frac{T^{2(p-1)}}{\theta_n^2 n\rho_n\varepsilon_n^{1+p}} \ll 1$.

Proof. We start by assuming $\frac{[\log(\varepsilon_n^{-\kappa})]^{2(p-1)}}{\varepsilon_n^{1+2p}\log(n)} \ll \varepsilon_n^{2\kappa}$ holds, take ε_n, ρ_n, $T = T(n)$ and ρ_n satisfying the appropriate assumptions and show $e^{\left(\frac{2+3C_{\mathrm{op}}^4}{2}\right)T}\theta_n \ll 1$ and $\frac{T^{2(p-1)}}{\theta_n^2 n\rho_n\varepsilon_n^{1+p}} \ll 1$. The former is equivalent to $\varepsilon_n^{-\kappa}\theta_n \ll 1$ or $\theta_n^2 \ll \varepsilon_n^{2\kappa}$.

For the latter, by assumption, we have that $\rho_n^{-1} \ll n^{1/2}\log(n)^{-1/2}\varepsilon_n^{-p}$ so that

$$
\frac{T^{2(p-1)}}{\theta_n^2 n\rho_n\varepsilon_n^{1+p}} \ll \frac{T^{2(p-1)}}{\theta_n^2 n^{1/2}\varepsilon_n^{1+2p}\log(n)^{1/2}}
$$

and therefore, the lower bound

$$\frac{T^{2(p-1)}}{n^{1/2}\varepsilon_n^{1+2p}\log(n)^{1/2}} = \frac{\left(\frac{2}{2+3C_{\mathrm{op}}^4}\right)^{2(p-1)}[\log(\varepsilon_n^{-\kappa})]^{2(p-1)}}{n^{1/2}\varepsilon_n^{1+2p}\log(n)^{1/2}} \ll \theta_n^2$$

is sufficient.

It remains to check under which conditions we have

$$\frac{\log(\varepsilon_n^{-\kappa})^{2(p-1)}}{n^{1/2}\varepsilon_n^{1+2p}\log(n)^{1/2}} \ll \varepsilon_n^{2\kappa}.$$

This is equivalent to

$$\frac{\log(\varepsilon_n^{-\kappa})^{2(p-1)}}{\varepsilon_n^{2\kappa+1+2p}} \ll n^{1/2}\log(n)$$

which, for n large enough, is implied by

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

In turn this leads to

$$\varepsilon_n \gg \left[\exp\mathrm{W}\left((n\log(n))^{1/\max(4(p-1),4+(2+4p)/\kappa)}\right)\right]^{-1/\kappa}. \qquad \square$$

Proof of Theorem 3.12. In the proof $C > 0$ will denote a constant that can be arbitrarily large, (which might be) dependent on Ω, u_0 or/and $\mathcal{A}^*\ell$, that may change from line to line.

The existence claims follow from Corollary 4.19 and Theorem 4.9.

In view of Theorem 3.7, let T_1 be the terms in the combination of (26) and (61) that are not included in the combination of (26) and (28).

Since the constant in the combination of (26) and (61) is independent of T, we have that for large $T \geq 1$ and small $\varepsilon_n \leq 1$:

$$\begin{aligned}
T_1 &\leq Ce^{\left(\frac{2+3C_{\mathrm{op}}^4}{2}\right)T}\left(\theta_n + \tau_n T^{p-1}\left(\sup_{x\in\Omega}\|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1} + \frac{\|K\|_{\mathrm{L}^\infty}}{\varepsilon_n^{1+p}}\right)\right.\\
&\qquad\left.\times\left[\left(\sup_{x\in\Omega}\|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}T^{p-2}\right)^{1/2} + \sup_{x\in\Omega}\|\mathcal{I}_n\bar{\Lambda}_n(x,\cdot)\|_{\mathrm{L}^1}T^{p-2}\right]\right)\\
&\leq Ce^{\left(\frac{2+3C_{\mathrm{op}}^4}{2}\right)T}\left(\theta_n + \tau_n\frac{T^{p-1}}{\varepsilon_n^{1+p}}\times\left[\frac{T^{(p-2)/2}}{\varepsilon_n^{p/2}} + \frac{T^{p-2}}{\varepsilon_n^{p}}\right]\right) &(65)\\
&\leq Ce^{\left(\frac{2+3C_{\mathrm{op}}^4}{2}\right)T}\left(\theta_n + \tau_n\frac{T^{2p-3}}{\varepsilon_n^{1+2p}}\right) &(66)\\
&=: T_2 + T_3
\end{aligned}$$

where we used Lemma 4.20 for (65) similar reasoning for $p \geq 3$ as in Theorem 3.7 for (66).

By assumption, we can apply Lemma 4.22 to see that $T_2 \to 0$. We have that $T_3 \to 0$ is equivalent to

$$\tau_n \ll \frac{\varepsilon_n^{1+2p+\kappa}}{\log(\varepsilon_n^{-\kappa})^{2p-3}}$$

which holds by assumption. We conclude the proof by combining (66) and (12) to obtain (14). The probability claim follows from Lemma 4.22. $\square$

Remark 4.23 (CFL condition for random-to-deterministic error convergence). Using the notation of Theorem 3.12, we see from the proof of the latter that the requirement on τ_n that ensures $T_1 \to 0$ (note that T_1 here is the error bound arising from the comparison of the random solution to the deterministic one) is slightly better than the CFL condition for our complete problem. Indeed, for $T_1 \to 0$, we only need

$$\tau_n \ll \frac{\varepsilon_n^{1+2p+\kappa}}{\log(\varepsilon_n^{-\kappa})^{2p-3}} \qquad \text{as opposed to} \qquad \tau_n \ll \frac{\varepsilon_n^{2+2p+\kappa}}{\log(\varepsilon_n^{-\kappa})^{2p-3}}$$

for the rest of the terms.

Proof of Theorem 3.13. We proceed as in the proof of Theorem 3.10 with Theorem 3.12. □

Remark 4.24 (Asymptotics in Theorems 3.12 and 3.13). The rates in Theorem 3.13 are formulated for large n. This is partly non-restrictive in practice as described in Remark 4.15 and in the proof of the theorem itself. The non-desirable part of this requirement stems from the application of Lemma 4.20 as discussed in Remark 4.21. We conclude from these results that when considering random graph models, one loses the full traceability of the asymptotic aspect of the rates as opposed to the results in Theorem 3.10.

Acknowledgements

The authors were supported by the European Research Council under the European Union's Horizon 2020 research and innovation programme grant agreement No 777826 (NoMADS). AW is particularly thankful for the warm hospitality during his stay at GREYC where the research idea behind the paper was developed.

References

[1] Fuensanta Andreu-Vaillo, Vicent Caselles, and José M. Mazón. *Parabolic Quasilinear Equations Minimizing Linear Growth Functionals*. Progress in Mathematics. Springer Basel AG, 2004.

[2] Fuensanta Andreu-Vaillo, José M. Mazón, Julio D. Rossi, and Julián Toledo. A nonlocal p-laplacian evolution equation with neumann boundary conditions. *Journal de Mathématiques Pures et Appliquées*, 90(2):201–227, 2008.

[3] Gilles Aubert and Pierre Kornprobst. Can the nonlocal characterization of sobolev spaces by bourgain et al. be useful for solving variational problems? *SIAM Journal on Numerical Analysis*, 47(2):844–860, 2009.

[4] Viorel P. Barbu. *Nonlinear semigroups and differential equations in Banach spaces*. Editura Academiei, rev. and enl. edition, 1976.

[5] John. W. Barrett and W. B. Liu. Finite element approximation of the p-laplacian. *Mathematics of Computation*, 61(204):523–537, 1993.

[6] John. W. Barrett and W. B. Liu. Finite element approximation of the parabolic p-laplacian. *SIAM Journal on Numerical Analysis*, 31(2):413–428, 1994.

[7] Mikhail Belkin and Partha Niyogi. Convergence of laplacian eigenmaps. In *Advances in Neural Information Processing Systems*, volume 19. MIT Press, 2007.

[8] Philippe Bénilan and Michael G. Crandall. Completely accretive operators. In P. Clement, editor, *Semigroup Theory and Evolution Equations: The Second International Conference*, chapter 4, pages 41–76. Taylor & Francis, 1991.

[9] Martin Benning and Martin Burger. Modern regularization methods for inverse problems. *Acta Numerica*, 27:1–111, 2018.

[10] Andrea L. Bertozzi and Arjuna Flenner. Diffuse interface models on graphs for classification of high dimensional data. *SIAM Review*, 58(2):293–328, 2016.

[11] Peter J. Bickel and Aiyou Chen. A nonparametric view of network models and Newman-Girvan and other modularities. *Proc Natl Acad Sci U S A*, 106(50):21068–21073, Dec 2009.

[12] 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. *Transactions of the American Mathematical Society*, 372(5):3019–3062, may 2019.

[13] Jean Bourgain, Haim Brezis, and Petru Mironescu. Another look at sobolev spaces. In *Optimal Control and Partial Differential Equations*, pages 439–455, 2001.

[14] Andrea Braides. *Γ-convergence for Beginners*. Oxford University Press, 2002.

[15] Haim Brezis. *Operateurs maximaux monotones: et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland mathematics studies. Elsevier Science, Burlington, MA, 1973.

[16] Leon Bungert, Jeff Calder, and Tim Roith. Uniform convergence rates for lipschitz learning on graphs, 2021.

[17] Johan Byström. Sharp constants for some inequalities connected to the p-laplace operator. *Journal of Inequalities in Pure and Applied Mathematics*, 6, 205.

[18] Jeff Calder. The game theoretic p-laplacian and semi-supervised learning with few labels. *Nonlinearity*, 32(1):301–330, dec 2018.

[19] Jeff Calder. Consistency of lipschitz learning with infinite unlabeled data and finite labeled data. *SIAM Journal on Mathematics of Data Science*, 1(4):780–812, 2019.

[20] Jeff Calder and Nicolás García Trillos. Improved spectral convergence rates for graph laplacians on ε-graphs and $k - nn$ graphs. *Applied and Computational Harmonic Analysis*, 60:123–175, 2022.

[21] Jeff Calder, Dejan Slepčev, and Matthew Thorpe. Rates of convergence for laplacian semi-supervised learning with low labeling rates, 2020.

[22] Marco Caroccia, Antonin Chambolle, and Dejan Slepčev. Mumford–shah functionals on graphs and their asymptotics. *Nonlinearity*, 33(8):3846–3888, jun 2020.

[23] Ronald R. Coifman and Stéphane Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006. Special Issue: Diffusion Maps and Wavelets.

[24] Ge Cong, Mehmet Esser, Bahram Parvin, and George Bebis. Shape metamorphism using p-laplacian equation. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, volume 4, pages 15–18 Vol.4, 2004.

[25] Richard Courant, Kurt Friedrichs, and Hans Lewy. On the partial difference equations of mathematical physics. *IBM Journal of Research and Development*, 11(2):215–234, 1967.

[26] Michael G. Crandall. An introduction to evolution governed by accretive operators. In Lamberto Cesari, Jack K. Hale, and Joseph P. LaSalle, editors, *Dynamical Systems*, pages 131–165. Academic Press, 1976.

[27] Riccardo Cristoferi and Matthew Thorpe. Large data limit for a phase transition model with the p-laplacian on point clouds. *European Journal of Applied Mathematics*, 31(2):185–231, 2020.

[28] Oleg Davydov. Algorithms and error bounds for multivariate piecewise constant approximation. In *Approximation Algorithms for Complex Systems*, pages 27–45, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[29] Carlos A. de Moura and Carlos S. Kubrusly. *The Courant–Friedrichs–Lewy (CFL) Condition: 80 Years After Its Discovery*. Springerlink: Bucher. Birkhäuser Boston, 2012.

[30] Ronald A. DeVore and George G. Lorentz. *Constructive Approximation*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 1993.

[31] Matthew Dunlop, Dejan Slepčev, Andrew Stuart, and Matthew Thorpe. Large data and zero noise limits of graph-based semi-supervised learning algorithms. *Applied and Computational Harmonic Analysis*, 49(2):655–697, September 2020.

[32] Ivar Ekeland and Roger Témam. *Convex Analysis and Variational Problems*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1999.

[33] Imad El Bouchairi, Jalal Fadili, and Abderrahim Elmoataz. Continuum limit of p-laplacian evolution problems on graphs: l^q graphons and sparse graphs. *ESAIM: Mathematical Modelling and Numerical Analysis*, 2023. arXiv 2010.08697.

[34] 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.

[35] Heinz W. Engl, Martin Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Mathematics and Its Applications. Springer Netherlands, 2000.

[36] Jalal Fadili, Nicolas Forcadel, Thi Tuyen Nguyen, and Rita Zantout. Limits and consistency of non-local and graph approximations to the eikonal equation. *IMA J. Numerical Analysis*, 2022. arXiv 2105.01977.

[37] Jan Francu. Monotone operators. a survey directed to applications to differential equations. *Applications of Mathematics*, 35:257–301, 1990.

[38] Nicolás García Trillos, Moritz Gerlach, Matthias Hein, and Dejan Slepčev. Error estimates for spectral convergence of the graph laplacian on random geometric graphs toward the laplace–beltrami operator. *Foundations of Computational Mathematics*, 20:827–887, 2020.

[39] Nicolás García Trillos and Dejan Slepčev. Continuum limit of total variation on point clouds. *Archive for Rational Mechanics and Analysis*, 220(1):193–241, 2016.

[40] Evarist Giné and Vladimir Koltschinskii. *Empirical graph Laplacian approximation of Laplace–Beltrami operators: Large sample results*, volume 51 of *IMS Lecture Notes Monographs Series*, pages 238–259. Institute of Mathematical Statistics, 2006.

[41] Roland Glowinski and A. Marroco. Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 9(R2):41–76, 1975.

[42] Charles W. Groetsch. *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*. Chapman & Hall/CRC research notes in mathematics series. Pitman Advanced Pub. Program, 1984.

[43] Yosra Hafiene, Jalal Fadili, Christophe Chesneau, and Abderrahim Elmoataz. Continuum limit of the nonlocal p-laplacian evolution problem on random inhomogeneous graphs. *ESAIM: M2AN*, 54(2):565–589, 2020.

[44] Yosra Hafiene, Jalal Fadili, and Abderrahim Elmoataz. Nonlocal p-laplacian evolution problems on graphs. *SIAM Journal on Numerical Analysis*, 56(2):1064–1090, 2018.

[45] Matthias Hein. Uniform convergence of adaptive graph-based regularization. In Gábor Lugosi and Hans Ulrich Simon, editors, *Learning Theory*, pages 50–64, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[46] Matthias Hein, Jean-Yves Audibert, and Ulrike von Luxburg. From graphs to manifolds – weak and strong pointwise consistency of graph laplacians. In Peter Auer and Ron Meir, editors, *Learning Theory*, pages 470–485, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[47] Brittney Hinds and Petronela Radu. Dirichlets principle and wellposedness of solutions for a nonlocal p-laplacian system. *Applied Mathematics and Computation*, 219(4):1411–1419, 2012.

[48] Giovanni Leoni. *A First Course in Sobolev Spaces*. Graduate studies in mathematics. American Mathematical Soc., 2009.

[49] Peter Lindqvist. *Notes on the Stationary p-Laplace Equation*. SpringerBriefs in Mathematics. Springer Cham, 2019.

[50] Jacques-Louis Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Collection études mathématiques. Dunod, 1969.

[51] Weifeng Liu, Xueqi Ma, Yicong Zhou, Dapeng Tao, and Jun Cheng. p-laplacian regularization for scene recognition. *IEEE Transactions on Cybernetics*, 49(8):2927–2940, 2018.

[52] Gianni Maso. *An Introduction to Γ-Convergence*. Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser Boston, 2012.

[53] Istvan Mezo. *The Lambert W Function: Its Generalizations and Applications*. Chapman and Hall/CRC, 2022.

[54] Nicolae H. Pavel. *Nonlinear Evolution Operators and Semigroups: Applications to Partial Differential Equations*. Lecture Notes in Mathematics. Springer, 1987.

[55] Bruno Pelletier and Pierre Pudlo. Operator norm convergence of spectral clustering on level sets. *Journal of Machine Learning Research*, 12(12):385–416, 2011.

[56] Augusto C. Ponce. A new approach to sobolev spaces and connections to \$\mathbf\gamma\$-convergence. *Calculus of Variations and Partial Differential Equations*, 19(3):229–255, 2004.

[57] Michael Reed and Barry Simon. *I: Functional Analysis*. Methods of Modern Mathematical Physics. Elsevier Science, 1981.

[58] Tim Roith and Leon Bungert. Continuum limit of lipschitz learning on graphs. *Foundations of Computational Mathematics*, 2022.

[59] Fillipo Santambrogio. *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*. Progress in Nonlinear Differential Equations and Their Applications. Springer International Publishing, 2015.

[60] Amit Singer and Hau-Tieng Wu. Spectral convergence of the connection Laplacian from random samples. *Information and Inference: A Journal of the IMA*, 6(1):58–123, 12 2016.

[61] Dejan Slepčev and Matthew Thorpe. Analysis of p-laplacian regularization in semisupervised learning. *SIAM Journal on Mathematical Analysis*, 51(3):2085–2120, 2019.

[62] Matthew Thorpe and Florian Theil. Asymptotic analysis of the ginzburg–landau functional on point clouds. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 149(2):387–427, 2019.

[63] Daniel Ting, Ling Huang, and Michael I. Jordan. An analysis of the convergence of graph laplacians. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, page 1079–1086, Madison, WI, USA, 2010. Omnipress.

[64] Hans Triebel. *Theory of Function Spaces III*. Monographs in Mathematics. Birkhäuser Basel, 2006.

[65] Nicolás García Trillos and Ryan Murray. A new analytical approach to consistency and overfitting in regularized empirical risk minimization. *European Journal of Applied Mathematics*, 28(6):886–921, 2017.

[66] Nicolás García Trillos and Dejan Slepčev. A variational approach to the consistency of spectral clustering. *Applied and Computational Harmonic Analysis*, 45(2):239–281, 2018.

[67] Nicolás García Trillos, Ryan Murray, and Matthew Thorpe. Rates of convergence for regression with the graph poly-laplacian, 2022.

[68] Yves van Gennip and Andrea Bertozzi. Gamma-convergence of graph ginzburg-landau functionals. *Advances inDifferential Equations*, 17, 04 2012.

[69] Ulrike von Luxburg, Mikhail Belkin, and Olivier Bousquet. Consistency of spectral clustering. *The Annals of Statistics*, 36(2):555 – 586, 2008.

[70] Xu Wang. Spectral convergence rate of graph laplacian, 2015.

[71] Dongming Wei. Existence, uniqueness, and numerical analysis of solutions of a quasilinear parabolic problem. *SIAM Journal on Numerical Analysis*, 29(2):484–497, 1992.

[72] Adrien Weihs and Matthew Thorpe. Consistency of fractional graph-laplacian regularization in semi-supervised learning with finite labels, 2023.

[73] Adrien Weihs and Matthew Thorpe. Consistency of fractional graph-laplacian regularization in semi-supervised learning with finite labels, 2023.

[74] Patrick J. Wolfe and Sofia C. Olhede. Nonparametric graphon estimation, 2013.

[75] Dengyong Zhou and Bernhard Schölkopf. Regularization on discrete spaces. In Walter G. Kropatsch, Robert Sablatnig, and Allan Hanbury, editors, *Pattern Recognition*, pages 361–368, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.