Nonlocal gradients within variational models

Existence theories and asymptotic analysis

Hidde Schönberger

Nonlocal gradients within variational models

Existence theories and asymptotic analysis

Katholische Universität
Eichstätt-Ingolstadt

Nonlocal gradients within variational models

Existence theories and asymptotic analysis

Kumulative Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften (Dr.rer.nat)
an der Mathematisch-Geographische Fakultät
der Katholischen Universität Eichstätt-Ingolstadt

vorgelegt von

Hidde Schönberger

aus Utrecht, Niederlande

Eichstätt, 2024

Betreuerin: Prof. Dr. Carolin Kreisbeck

Eingereicht am: 17.05.2024

Tag der mündlichen Prüfung: 18.07.2024

Prüfungskommission:

Vorsitzende:	Prof. Dr. Susanne Jochner-Oette
1. Gutachterin:	Prof. Dr. Carolin Kreisbeck
2. Gutachter:	Prof. Dr. Carlos Mora-Corral (Universidad Autónoma de Madrid)
Weitere Prüferin:	Prof. Dr. Elvira Zappale (Sapienza Università di Roma)
Ersatzprüferin:	Prof. Dr. Nadja Ray

Cover: The design of the cover is inspired by materials modeling, which is one of the main applications of the mathematical theory developed in the thesis. The yellow region represents an elastic material made up of individual particles. A characteristic feature of the nonlocal models is that the particles that are within a certain distance interact with each other, which is illustrated on the front. The back of the cover is dedicated to one of the most important mathematical tools in the thesis: The translation mechanism. This is a method that enables one to transform the nonlocal gradient into a classical gradient and vice versa. The two operators that are needed for the translation mechanism are visually represented in the diagrams on the top and bottom.

Summary

The study of nonlocal energy-based models has proliferated in the last decades due to their relevance in applications and the novel mathematical challenges that they provide. Within this context, the thesis addresses a wide variety of aspects concerning variational problems involving nonlocal gradients. These have been proposed for modeling the emergence of cracks and cavitations in materials, and can be used in image denoising applications to preserve sharp features. The contributions of the thesis are concerned with several families of nonlocal gradients and can be divided into the following overarching themes:

- (i) Development of nonlocal Sobolev spaces;
- (ii) existence theories for minimizers of nonlocal integral functionals;
- (iii) asymptotic analysis of parameter-dependent problems.

In the context of (i), we extend and unify the results established for the Riesz and finite-horizon fractional gradient by considering more general nonlocal gradients with radial kernels. Precisely, based on a delicate analysis of the Fourier symbol of the nonlocal gradient, we present minimal assumptions on the kernel function such that Poincaré inequalities and compact embeddings hold. Moreover, we provide sharp embeddings into Orlicz spaces and spaces with prescribed modulus of continuity, that refine the fractional Sobolev and Morrey inequalities.

With these tools at hand, the goal in (ii) is to establish rigorous existence results for minimizers of integral functionals depending on nonlocal gradients. We develop a translation mechanism that connects the nonlocal gradients with their local counterpart and use this to characterize the lower semicontinuity of the nonlocal functionals in terms of the classical notion of quasiconvexity. Based on this, we establish the existence of minimizers via the direct method under Dirichlet or novel Neumann-type boundary conditions. The latter conditions necessitate the study of functions with zero nonlocal gradient, which surprisingly constitute an infinite-dimensional vector space, and we employ recent results regarding pseudo-differential boundary-value problems to characterize them. We also tackle linear growth functionals, which lack the usual coercivity requirement, and provide an explicit formula for their relaxation in the space of functions with bounded fractional variation. This reveals insights about the behavior of minimizing sequences and is based on a careful analysis of the concentration effects of the fractional gradient by using tools from Young measure theory.

Building upon the existence results, the focus in (iii) lies on studying the dependence of the minimizers on parameters in the problem, specifically, the fractional parameter s and the interaction range δ. We establish using Γ-convergence that the minimizers depend continuously on s and converge to the local solutions as $s \to 1$. In parallel, we prove that localization also occurs in the vanishing horizon limit $\delta \to 0$, while the diverging horizon regime $\delta \to \infty$ leads to the models based on the Riesz fractional gradient. A crucial ingredient for the proofs of these results are uniform compactness statements, which are established by studying the dependence of the Fourier symbols on the respective parameters. In terms of applications, we develop an abstract framework for the well-posedness of bi-level optimization problems for parameter learning in image denoising. The results of the thesis show that the models involving nonlocal gradients fit into this framework, where the fractional parameter is tuned in order to obtain the optimal degree of smoothing.

Zusammenfassung

Die Betrachtung von nichtlokalen energiebasierten Modellen hat in den letzten Jahrzehnten aufgrund ihrer Relevanz für Anwendungen und der damit einhergebenden neuen mathematischen Herausforderungen erheblich zugenommen. In diesem Zusammenhang befasst sich diese Arbeit mit einer Vielzahl Aspekten von Variationsproblemen mit nichtlokalen Gradienten. Diese wurden für die Modellierung der Entstehung von Rissen und Kavitation in Materialien vorgeschlagen und können verwendet werden, um scharfe Merkmale bei der Rauschentfernung in Bildern zu behalten. Die Beiträge dieser Arbeit beziehen sich auf mehrere Familien von nichtlokalen Gradienten und können in die folgenden übergreifenden Themen unterteilt werden:

(i) Entwicklung von nichtlokalen Sobolev-Räumen;

(ii) Existenztheorien für Minimierer von nichtlokalen Integralfunktionalen;

(iii) Asymptotische Analyse von parameterabhängigen Problemen.

Im Kontext von (i) vereinheitlichen wir die Ergebnisse für den Riesz-Gradienten und den fraktionalen Gradienten mit endlichem Horizont, indem wir allgemeinere nichtlokale Gradienten mit radialsymmetrischen Kernen betrachten. Basierend auf einer detaillierten Analyse des Fourier-Symbols des nichtlokalen Gradienten stellen wir minimale Annahmen für die Kernfunktion so auf, dass Poincaré-Ungleichungen und kompakte Einbettungen gelten. Außerdem zeigen wir scharfe Einbettungen in Orlicz-Räume und Räume mit vorgeschriebenem Stetigkeitsmodul, die die fraktionalen Sobolev- und Morrey-Ungleichungen verfeinern.

Unter Verwendung dieser verfügbaren Werkzeugen ist das Ziel in (ii), die rigorose Herleitung von Existenzresultaten für Minimierer von Integralfunktionalen, die von nichtlokalen Gradienten abhängen. Wir entwickeln dazu einen Translationsmechanismus, der nichtlokale Gradienten mit ihrem lokalen Gegenstück verbindet, und eine Charakterisierung der Unterhalbstetigkeit nichtlokaler Funktionale durch den klassischen Begriff der Quasikonvexität ermöglicht. Auf dieser Grundlage können wir die Existenz von Minimierern mit Hilfe der direkten Methode unter Dirichlet- oder neuen Neumann-Randbedingungen nachweisen. Die zuletztgenannten Bedingungen erfordern die Analyse von Funktionen mit verschwindendem nichtlokalen Gradienten, die überraschenderweise einen unendlich-dimensionalen Vektorraum bilden, und charackterisieren diese mittels moderner Ergebnisse zu Randwertproblemen mit Pseudodifferentialoperatoren. Wir behandeln außerdem Funktionale mit linearem Wachstum, die die übliche Koerzitivitätsannahme nicht erfüllen, und geben eine explizite Formel für ihre Relaxierung in dem Raum der Funktionen mit beschränkter fraktionaler Variation. Dies liefert neue Erkenntnisse über das Verhalten von Minimalfolgen und basiert auf einer sorgfältigen Analyse der Konzentrationseffekte fraktionaler Gradienten unter Verwendung der Young-Maß-Theorie.

Aufbauend auf den Existenzresultaten liegt der Schwerpunkt in (iii) auf der Betrachtung der Parameterabhängigkeit der Minimierer, insbesondere bezüglich des fraktionalen Parameters s und des Interaktionsradius δ. Mittels Γ-Konvergenz zeigen wir, dass die Minimierer stetig von s abhängen und, im Fall $s \to 1$, zu lokalen Lösungen konvergieren. Parallel beweisen wir, dass die Lokalisierung auch bei einem verschwindenden Horizont $\delta \to 0$ auftritt, während das Regime des divergierenden Horizonts $\delta \to \infty$ zu den Modellen führt, die auf dem fraktionalen Riesz-Gradienten basieren. Ein entscheidender Bestandteil zum Beweis dieser Ergebnisse sind gleichmäßige Kompaktheitsaussagen, die durch die Untersuchung der Parameterabhängigkeit der Fourier-Symbole gezeigt werden. In Bezug auf Anwendungen entwickeln wir einen abstrakten Rahmen für die Lösbarkeit von Bilevel-Optimierungsproblemen für das Parameterlernen bei der Bildentrauschung. Die Ergebnisse der Arbeit zeigen, dass die Modelle mit nichtlokalen Gradienten in diesen Rahmen passen, wobei der fraktionale Parameter abgestimmt wird, um den optimalen Grad der Entrauschung zu erhalten.

Acknowledgements

To begin, I would like to thank my supervisor Carolin Kreisbeck for her constant support throughout this endeavor. You always went above and beyond to provide guidance on any issue, which was very reassuring in the face of some new and unknown experiences. I am deeply grateful for all the time and effort you dedicated to helping me develop in all aspects of my mathematical and professional career.

I am also thankful to Carlos Mora-Corral and José Carlos Bellido who kindly hosted me in Madrid and Ciudad Real during a two month period in the spring of 2023. I appreciate all the discussions we had and the opportunity of meeting new people at Universidad Autónoma de Madrid and Universidad de Castilla-La Mancha. It was a valuable experience and our collaboration resulted in an important chapter of the thesis. Beyond that, I wish to express appreciation to Carlos for taking the time to be the second reviewer of my thesis. The research visit was made possible with financial support from the Bayerische Forschungsallianz (BayIntAn_KUEI_2023_16) and the Agencia Estatal de Investigación (PID2021-124195NB-C32, CEX2019-000904-S), which is gratefully acknowledged.

My gratitude goes out to Javier Cueto, Elisa Davoli, Rita Ferreira, Anna Doležalová, Dominik Engl and Anastasia Molchanova, with whom I collaborated during my doctorate. I thoroughly enjoyed all the meetings that we had during this time, both online and in-person. I hope that our joint efforts can continue in the future.

I also wish to thank Marco Bresciani, Samuele Riccò, Marvin Weidner, Adolfo Arroyo-Rabasa and many others for the interesting discussions at various conferences or (online) meetings during this three year period.

Many thanks to the analysis group in Eichstätt - Antonella, Carolin, Dominik and also Christoph and Fabian - for the experiences at the university and especially the fun group activities outside of it; those were a great way for me to get my mind off the thesis. Besides, I am grateful to Dominik for helping with the German translation of the summary, and for his feedback on the introduction.

Finally, I could not have done this without the support of my family and friends. A special thanks to my parents Marc & Inge, my brother Daan, my grandparents, and my friends back in the Netherlands. Even though I could not always convey what I was precisely working on, you still showed interest. I am glad I could visit all of you often, despite the distance, which helped to balance all the challenges of the PhD.

Hidde Schönberger

Contents

Chapter 1

Introduction

Modeling and predicting the behavior of complex phenomena in the real world is one of the major applications of the abstract field of mathematics. Although there are a vast number of ways to design mathematical models, an exceedingly versatile and successful approach is based on the minimization of a given intrinsic quantity; the minimized quantities will often have a physically relevant meaning in the application, such as energy, time or distance. The mathematical area that is devoted to these problems is called the calculus of variations, and the thesis will consider problems within this framework.

In general, the calculus of variations is concerned with so-called variational problems that consist of the minimization of a given functional

$$\mathcal{F} : X \to \mathbb{R}_\infty := \mathbb{R} \cup \{\infty\},$$

with X a suitable space of functions. The key challenge that distinguishes these problems from more elementary optimization problems on $\mathbb{R}^n$, is that X is an infinite-dimensional space, and this necessitates the use of vastly different and more advanced tools. On the flip side, the formulation in terms of minimization problems is quite convenient for modeling purposes and many applications can be translated into this framework. Prominent examples include, among many others, Fermat's principle of least time in optics, Hamilton's principle of least action in classical mechanics, the study of geodesics and minimal surfaces in geometry, optimal control theory, and variational methods in image processing (cf. [20, 150, 182, 206]).

The example that initiated the calculus of variations, however, is commonly attributed to the well-known Brachistochrone problem, which was posed as a challenge by Johann Bernoulli to his peers in 1696 [43]. The problem consisted of finding a continuous path between two fixed points such that the time it takes for a frictionless ball to roll down this path, solely driven by gravity, is minimized. It turns out that the optimal curve is a cycloid, which is a shape that can be traced out by a circle rolling along a straight line. Multiple prominent figures in the scientific world at the time solved the problem, such as Leibniz, Newton and both Bernoulli brothers [121, 206], but it was the contributions of Euler and Lagrange in the next century that provided the first systematic method for solving variational problems.

Their approach, which is nowadays called the classical or indirect method in the calculus of variations, consists of deriving necessary conditions for minimizers, i.e., functions that minimize the functional, by setting the first variation of $\mathcal{F}$ equal to zero. In essence, one generalizes the concept of stationary points to functionals, and this leads, in the case of the commonly studied integral functionals in (1.1), to solving a partial differential equation known as the Euler-Lagrange equation. This important observation allows one to solve these equations in order to find minimizers, and gave shape to the field for years to come. One major drawback, however, is that not every stationary point of the functional needs to be a minimizer, and therefore, solving the equations does not immediately solve the minimization problem at hand. In fact, a surprising discovery made by Weierstrass in 1870 [209] showed that, even in the class of integral functionals, there are examples that do not admit any minimizers.

The urge to resolve these shortcomings gave way to the modern calculus of variations, whose development was set in motion by Hilbert when he presented his famous 23 problems in 1900 [131]. Out of the three problems related to the calculus of variations, it was the 20th problem that pertained to the existence of minimizers for variational problems involving integral functionals. Significant contributions to this problem were made by Hilbert, Noether, Tonelli, Lebesgue and Hadamard, which led to the conception of the direct method, see Section 1.1. As opposed to arguing through stationary points of the functional, one directly works with the functional by verifying that it is coercive and lower semicontinuous, after which the existence of minimizers is guaranteed. This method finally put the calculus of variations on a rigorous mathematical foundation, and is still immensely important nowadays.

The generality of the direct method allows it to be applied to any kind of functional, but the most prevalent class of functionals considered in the calculus of variations are integral functionals involving first order derivatives, that is

$$\mathcal{F}(u) = \int_\Omega f(x, u, Du)\, dx, \tag{1.1}$$

with $\Omega \subset \mathbb{R}^n$ and $u : \Omega \to \mathbb{R}^m$ a suitable function with Du its derivative, see Section 1.1 for more details. As a result, a lot of attention in the 20th century went to finding conditions under which these functionals satisfied the requirements of the direct method. Coercivity can be achieved when one works on Sobolev spaces by imposing certain growth bounds on the integrand, but only with respect to the weak topology. Hence, the lower semicontinuity has to be verified with respect to this same topology, which is a delicate issue. It turns out that convexity of the integrand plays a key role in this regard; in fact, for integrands with standard growth the weak lower semicontinuity of vectorial integral functionals can be characterized in terms of quasiconvexity, which is a generalized convexity notion introduced in the seminal work of Morrey in 1952 [165].

Further significant aspects of the modern calculus of variations include the topics of relaxation and Γ-convergence. The former is an approach for studying functionals that are not lower semicontinuous, and therefore, not tractable by the direct method. One considers a related relaxed problem, by determining the lower semicontinuous envelope of the functional. Under suitable coercivity assumptions, this new functional will admit minimizers that are related to the original problem. Explicitly, the minimizing sequences of the functional converge up to subsequence to a minimizer of the relaxation. In the case of classical integral functionals, it has been shown by Dacorogna [74] that the relaxation is again an integral functional, with the integrand replaced by its quasiconvex envelope.

Introduced by De Giorgi and Franzoni in 1975 [85], Γ-convergence is a method to study sequences of variational problems and their asymptotic behavior. It provides a systematic way to verify whether the sequence converges to a limit variational problem, in the sense that the minima and minimizers converge. In this way, one can relate the sequence of functionals to its limit functional, or study the continuous dependence of minimizers on certain parameters in the model. An overview of the direct method in the calculus of variations and the results regarding integral functionals involving gradients is given in Section 1.1.

While functionals as in (1.1) have proven very effective in applications, a recent development has seen the interest rise in different classes of functionals as well. This is because they can possess new features that are desirable in given applications. Especially models that are of a nonlocal nature have been intensively studied due to their ability to incorporate long-range effects; this is in contrast to models based on purely local notions like the gradient. Applications of nonlocal models range from continuum mechanics [195–197], to fractional models in physics, chemistry and biology [132, 177], and image processing [13, 21, 117, 134]. Additionally, in nonlocal models derivatives are often not needed, which makes it possible to work with less regular functions that can exhibit discontinuities. This can be particularly useful when one is interested in the formation of cracks and cavities in elastic materials or in recovering sharp features in noisy images, see also Section 1.3.

There are different ways in which nonlocal effects can be incorporated into variational models, but in the thesis we will mostly focus on integral functionals similar to (1.1) where the derivative is replaced by a nonlocal gradient of the form

$$D_\rho u(x) := \int_{\mathbb{R}^n} \frac{u(x)-u(y)}{|x-y|}\frac{x-y}{|x-y|}\rho(x-y)\,dy, \tag{1.2}$$

for $u : \mathbb{R}^n \to \mathbb{R}$ and with $\rho : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ a suitable kernel function. Intuitively, this gradient can be seen as an average of discrete difference quotients weighted by the function ρ. Despite the natural way of incorporating nonlocality by replacing the gradient with D_ρ in (1.1), such models have only been considered in the last few years. The first existence results for minimizers appeared in 2015 in the paper by Shieh & Spector [193] for the so-called Riesz fractional gradient, which arises by choosing

$$\rho(x) = c_{n,s}\frac{1}{|x|^{n+s-1}}$$

in (1.2) with $s \in (0,1)$ and $c_{n,s}$ a normalizing constant. They developed the suitable function spaces and applied the direct method to obtain the well-posedness of integral functionals depending on the fractional gradient. Since then, the interest in this gradient and in related variational problems has grown tremendously, see e.g. [28, 31, 66, 92, 93, 140, 208] and also Section 1.2 for a more broad overview of these and related fractional and nonlocal models.

Since the study of variational problems involving nonlocal gradients is still in its inception, there are a lot of fundamental tasks in the context of the direct method left open, which we consider in this thesis. They include the development of the appropriate function spaces and Poincaré inequalities, characterizing weak lower semicontinuity, proving relaxation and localization results, and investigating different types of boundary conditions. The gradients we consider range from the Riesz fractional gradient, to a truncated fractional gradient with finite interaction range, and to even more general nonlocal gradients of the type (1.2).

Although theoretical in nature, the models we consider are motivated by applications of which we highlight two in Section 1.3. The first is in peridynamics, which is a nonlocal formulation of continuum mechanics introduced by Silling [196]. The advantage of this formulation is that peridynamic models can incorporate interactions at a distance in the material and allow discontinuities like cracks and cavities to emerge through the use of nonlocal terms instead of derivatives. The original formulation of bond-based peridynamics is quite limited though, since the linear elastic models are restricted to materials with Poisson ratio equal to $\frac{1}{4}$ [195, 197], and only a small class of nonlinear models can be recovered in the localization limit [27, 160]. Therefore, a more general state-based theory was developed (cf. [195,197]), which overcame these shortcomings. We point out that the integral functionals depending on nonlocal gradients constitute a mathematically rigorous class that fits into the state-based framework.

Secondly, the nonlocal models can be used as regularizers in applications of image denoising. This is because the freedom of choosing the kernel ρ, for example, through the fractional parameter $s \in (0,1)$, allows one to change the amount of regularity that is imposed on functions. Therefore, one can tune the regularization in order to find a balance between the amount of smoothing of the noise and the retention of sharp features in the image. The learning of the optimal regularizer with respect to a given data set can be done with a bi-level training scheme, and this topic will also appear in Chapter 7 again.

We now turn to the main body of the thesis, which consists of six chapters that correspond to the articles [36, 72, 73, 82, 141, 189]. Each chapter is introduced and highlighted in Section 1.4 below. To summarize, Chapter 2 deals with linear growth integral functionals involving the Riesz fractional gradient. Such functionals lack the coercivity properties to apply the direct method, and, therefore, we characterize their relaxation to the space of functions with bounded fractional variation, see also Section 1.1.2 for the theory around classical linear growth functionals. In Chapter 3, we further the development around integral functionals involving the finite-horizon fractional gradient introduced in [31], providing a broad theory for the weak lower semicontinuity, Γ-convergence and localization of these functionals subject to Dirichlet conditions. Chapter 4 addresses the same gradients, where we investigate problems with Neumann-type boundary conditions instead of Dirichlet conditions. This necessitates the careful study of the functions with vanishing finite-horizon fractional gradient, which are characterized using the recent existence and regularity theory for pseudo-differential operators [2, 125]. The general gradients of the type (1.2) and associated function spaces are studied in Chapter 5, where almost minimal conditions on ρ are derived in order for Poincaré inequalities and compact embeddings to hold. This paves the way for the study of integral functionals involving this more general class of nonlocal gradients, which is taken up in Chapter 6. Moreover, it is shown via a vanishing and diverging horizon limit that these models are consistent with their local and fractional counterpart, respectively. Finally, in Chapter 7, an abstract framework is built around the learning of optimal regularizers in applications of image denoising. Several examples with regularizers of nonlocal and fractional type related to the ones in the thesis are shown to fit within the general framework. In fact, we show in Section 1.3.2 that also the integral functionals depending on nonlocal gradients can be used in these applications.

1.1 The direct method and classical integral functionals

As mentioned in the introduction, the direct method was developed in the 20th century to provide a rigorous mathematical framework for the calculus of variations and existence of minimizers for variational problems. The goal of this section is to detail the general abstract theory of the direct method in its modern form, which is used repeatedly in the main body of the thesis. Subsequently, we will highlight the results regarding integral functionals depending on derivatives and how the direct method specifies to this setting. We refer to e.g. [19, 49, 75, 80, 112, 173, 182] for more details on this topic.

1.1.1 The abstract framework

We begin with the abstract direct method, which is used to establish the existence of minimizers for a functional $\mathcal{F} : X \to [-\infty, \infty]$ on a topological space X. To obtain a clean and comprehensive theory in which all the results of the thesis fit in, we assume that we are in one of the two situations:

(A1) X is a first countable topological space;

(A2) X is a reflexive separable Banach space endowed with the weak topology.

Although these assumptions are not crucial for the direct method, they will become important later when talking about relaxation and Γ-convergence. Let us now introduce the two main properties of functionals which enable the use of the direct method.

Definition 1.1.1 (Sequential lower semicontinuity). *A functional $\mathcal{F} : X \to [-\infty, \infty]$ is sequentially lower semicontinuous if for every sequence $(u_j)_j \subset X$ with $u_j \to u$ in X it holds that*

$$\mathcal{F}(u) \leq \liminf_{j\to\infty} \mathcal{F}(u_j).$$

Definition 1.1.2 (Coercivity). *A functional $\mathcal{F} : X \to [-\infty, \infty]$ is coercive if the closure of the sub-level set $L_c(\mathcal{F}) := \{u \in X : \mathcal{F}(u) \leq c\}$ is sequentially compact for every $c \in \mathbb{R}$.*

Remark 1.1.3. a) There is also the notion of lower semicontinuity, which states that the sub-level sets $L_c(\mathcal{F})$ are closed in X for all $c \in [-\infty, \infty]$; sequential lower semicontinuity only requires them to be sequentially closed. Under the assumption (A1) of first countability, the two notions agree. Moreover, in the setting of (A2) and the additional assumption that $\mathcal{F}$ is coercive, the sequential and topological notion also coincide, see Proposition 1.1.6 below. In view of this and since we will mostly only work with the sequential notion, we often refer to sequential lower semicontinuity simply as lower semicontinuity.

b) There are some alternative notions of coercivity available in the literature, most notably the weaker assumption that the closure of the sub-level sets $\overline{L_c(\mathcal{F})}$ are countably compact for all $c \in \mathbb{R}$ [80, Definition 1.12]. However, this coincides with our definition in the setting of (A1) or (A2). This is because countable compactness agrees with sequential compactness in first countable spaces and in the weak topology of Banach spaces by the Eberlein-Šmulian theorem. Finally, in the setting of (A2), coercivity is also equivalent to the simpler assumption

$$\mathcal{F}(u) \to \infty \quad \text{as} \quad \|u\| \to \infty,$$

with $\|\cdot\|$ the norm on the Banach space X, see [80, Example 1.14].

c) If we are in the case (A2) but $\mathcal{F}$ is only defined on a subset $\mathcal{A} \subset X$, then as long as $\mathcal{A}$ is closed, the coercivity and sequential lower semicontinuity of $\mathcal{F}$ on $\mathcal{A}$ remain preserved if we extend $\mathcal{F}$ to X as identically ∞. Therefore, working on the full space X as opposed to $\mathcal{A}$ loses no generality. △

We now state the result on the existence of solutions of the variational problem

$$\inf\{\mathcal{F}(u) : u \in X\}, \tag{1.3}$$

whose proof we detail for the reader's convenience.

Theorem 1.1.4 (Direct method). *Let $\mathcal{F} : X \to [-\infty, \infty]$ be sequentially lower semicontinuous and coercive. Then, $\mathcal{F}$ admits a minimizer $u_0 \in X$, that is,*

$$\mathcal{F}(u_0) = \inf\{\mathcal{F}(u) : u \in X\}.$$

Proof. If $\mathcal{F} \equiv \infty$, then any point in X is a minimizer of $\mathcal{F}$. Otherwise, we may take a minimizing sequence $(u_j)_j \subset X$ that satisfies

$$\lim_{j\to\infty} \mathcal{F}(u_j) = \inf\{\mathcal{F}(u) : u \in X\} < \infty.$$

Therefore, there is a $c \in \mathbb{R}$ such that $u_j \in L_c(\mathcal{F})$ for all $j \in \mathbb{N}$ large enough. By coercivity of $\mathcal{F}$, there is a $u_0 \in X$ so that up to a non-relabeled subsequence $u_j \to u_0$. The sequential lower semicontinuity then yields

$$\mathcal{F}(u_0) \leq \liminf_{j\to\infty} \mathcal{F}(u_j) = \inf\{\mathcal{F}(u) : u \in X\}.$$

□

In the case that lower semicontinuity is not satisfied, the direct method is not applicable and (1.3) might not have any solutions. A common strategy is to resort to the relaxation of the functional.

Definition 1.1.5 (Relaxation). *Let $\mathcal{F} : X \to [-\infty, \infty]$, then we define its relaxation $\mathcal{F}^{\mathrm{rel}} : X \to [-\infty, \infty]$ as*

$$\mathcal{F}^{\mathrm{rel}}(u) := \inf \left\{ \liminf_{j\to\infty} \mathcal{F}(u_j) \ : \ u_j \to u \text{ in } X \right\}.$$

We also introduce the lower semicontinuous and sequentially lower semicontinuous envelopes

$$\operatorname{lsc} \mathcal{F}(u) := \sup \{\mathcal{G}(u) \ : \ \mathcal{G} \text{ lower semicontinuous}, \ \mathcal{G} \leq \mathcal{F}\}$$

and

$$\operatorname{slsc} \mathcal{F}(u) := \sup \{\mathcal{G}(u) \ : \ \mathcal{G} \text{ sequentially lower semicontinuous}, \ \mathcal{G} \leq \mathcal{F}\}$$

see [112, Definition 3.8], which are the largest lower semicontinuous and sequentially lower semicontinuous functionals below $\mathcal{F}$, respectively. In general, it follows from the definitions that $\operatorname{lsc} \mathcal{F} \leq \operatorname{slsc} \mathcal{F} \leq \mathcal{F}^{\mathrm{rel}} \leq \mathcal{F}$. In light of [112, Proposition 3.12 and 3.16], where the latter uses the local metrizability of the weak topology in the setting of (A2), we also obtain the following.

Proposition 1.1.6. *Let $\mathcal{F} : X \to [-\infty, \infty]$ and assume (A1) is satisfied, or (A2) is satisfied and $\mathcal{F}$ is coercive. Then, it holds that $\mathcal{F}^{\mathrm{rel}} = \operatorname{lsc} \mathcal{F} = \operatorname{slsc} \mathcal{F}$.*

We can now state the main result regarding relaxation and how it relates to the original minimization problem (1.3), see [80, Theorem 3.8].

Theorem 1.1.7. *Let $\mathcal{F} : X \to [-\infty, \infty]$ be coercive, then $\mathcal{F}^{\mathrm{rel}}$ is coercive and admits a minimizer with*

$$\min\{\mathcal{F}^{\mathrm{rel}}(u) \ : \ u \in X\} = \inf\{\mathcal{F}(u) \ : \ u \in X\}.$$

If $\mathcal{F} \not\equiv \infty$, then any minimizing sequence $(u_j)_j$ of $\mathcal{F}$, i.e.,

$$\lim_{j\to\infty} \mathcal{F}(u_j) = \inf\{\mathcal{F}(u) \ : \ u \in X\},$$

converges up to subsequence to a minimizer of $\mathcal{F}^{\mathrm{rel}}$.

We now proceed with the concept of Γ-convergence introduced by De Giorgi and Franzoni in 1975 [85], which is related to the convergence of a sequence of variational problems.

Definition 1.1.8 (Γ-convergence). *Let $(\mathcal{F}_j)_j, \mathcal{F}_\infty : X \to [-\infty, \infty]$, then we say that $(\mathcal{F}_j)_j$ (sequentially) Γ-converges to $\mathcal{F}_\infty$, and write $\mathcal{F}_\infty = \Gamma\text{-}\lim_{j\to\infty} \mathcal{F}_j$, if:*

(*i*) *For every sequence $(u_j)_j \subset X$ with $u_j \to u$ in X, it holds that*

$$\mathcal{F}_\infty(u) \leq \liminf_{j\to\infty} \mathcal{F}_j(u_j).$$

(*ii*) *For each $u \in X$, there exists a sequence $(u_j)_j \subset X$ with $u_j \to u$ in X such that*

$$\mathcal{F}_\infty(u) \geq \limsup_{j\to\infty} \mathcal{F}_j(u_j).$$

Property (*i*) is called the liminf-inequality, while (*ii*) is referred to as the limsup-inequality. In fact, if (*i*) is satisfied, then any sequence with the property as in (*ii*) satisfies

$$\lim_{j\to\infty} \mathcal{F}_j(u_j) = \mathcal{F}_\infty(u).$$

Therefore, the sequence in property (*ii*) is called a recovery sequence. We also introduce a uniform notion of coercivity, that is used alongside Γ-convergence.

Definition 1.1.9 (Equi-coercivity). *A sequence $(\mathcal{F}_j)_j : X \to [-\infty, \infty]$ is called equi-coercive, if there exists a coercive functional $\mathcal{G} : X \to [-\infty, \infty]$ such that $\mathcal{F}_j \geq \mathcal{G}$ for all $j \in \mathbb{N}$.*

Remark 1.1.10. a) The definition of equi-coercivity is equivalent to the one in [80, Definition 7.6] in light of [80, Proposition 7.7]; indeed, we do not need to assume that $\mathcal{G}$ is lower semicontinuous since we can replace it by $\operatorname{lsc} \mathcal{G}$ which is also coercive, cf. Theorem 1.1.7 and Proposition 1.1.6.

b) We note that there is also a topological notion of Γ-convergence, see [80, Definition 4.1]. However, this agrees with the sequential notion in the case (A1), or in the case (A2) with the additional assumption of equi-coercivity [80, Proposition 8.1 and 8.10].

c) In the setting of (A2), there exists a metric that induces the weak topology on bounded sets [80, Proposition 8.7]. For any such metric, the Γ-limit of an equi-coercive sequence with respect to this metric coincides with the weak Γ-limit [80, Proposition 8.10]. In particular, if X is a type of Sobolev space that is compactly embedded into L^p for some $p \in (1, \infty)$, then the metric given by L^p-convergence can be used instead of the weak topology on X cf. [80, Example 8.9]. Because of this, we often state our Γ-convergence results with respect to the L^p-convergence. △

Suppose now that $(\mathcal{F}_j)_j : X \to [-\infty, \infty]$, and we are in the setting (A1), or in the setting (A2) with the additional assumption of equi-coercivity so that Remark 1.1.10 b) applies. If $(\mathcal{F}_j)_j$ has a Γ-limit $\mathcal{F}_\infty$, then this functional is automatically lower semicontinuous [80, Proposition 6.8]. Intuitively, this means that Γ-convergence induces a relaxation process in addition to the convergence of variational problems. In fact, this can be made precise as it holds that

$$\Gamma\text{-}\lim_{j\to\infty} \mathcal{F}_j = \mathcal{F}_\infty \quad \text{if and only if} \quad \Gamma\text{-}\lim_{j\to\infty} \mathcal{F}_j^{\mathrm{rel}} = \mathcal{F}_\infty,$$

see [80, Proposition 6.11]. Moreover, if $\mathcal{F} : X \to [-\infty, \infty]$ is a functional in the setting of Proposition 1.1.6, then $\mathcal{F}^{\mathrm{rel}}$ coincides with the Γ-limit of the constant sequence $\mathcal{F}_j \equiv \mathcal{F}$ for all $j \in \mathbb{N}$, i.e.,

$$\mathcal{F}^{\mathrm{rel}} = \Gamma\text{-}\lim_{j\to\infty} \mathcal{F}, \tag{1.4}$$

cf. [80, Remark 4.5].

We now phrase the main result about Γ-convergence, which justifies its usage in studying the convergence of variational problems, see [80, Theorem 7.8 and Corollary 7.20].

Theorem 1.1.11. *Let $(\mathcal{F}_j)_j : X \to [-\infty, \infty]$ be an equi-coercive sequence that Γ-converges to $\mathcal{F}_\infty : X \to [-\infty, \infty]$. Then, $\mathcal{F}_\infty$ is coercive and admits a minimizer with*

$$\min\{\mathcal{F}_\infty(u) \,:\, u \in X\} = \lim_{j\to\infty} \inf\{\mathcal{F}_j(u) \,:\, u \in X\}.$$

If $\mathcal{F}_\infty \not\equiv \infty$, then every sequence $(u_j)_j \subset X$ of almost minimizers, i.e.,

$$\lim_{j\to\infty} \mathcal{F}_j(u_j) = \lim_{j\to\infty} \inf\{\mathcal{F}_j(u) \,:\, u \in X\},$$

converges up to subsequence to a minimizer of $\mathcal{F}_\infty$.

1.1.2 Classical integral functionals

With the abstract theory established, we now consider the application to the specific class of integral functionals involving derivatives, which is the most prominent class studied in the modern calculus of variations with applications in classical mechanics, biology, chemistry, finance, optimal control theory and image processing [20, 150, 182, 206]. The natural space to work on is the Sobolev space on a bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$, which consists of the functions in L^p with distributional derivative in L^p, i.e.,

$$W^{1,p}(\Omega;\mathbb{R}^m) := \{u \in L^p(\Omega;\mathbb{R}^m) \,:\, Du \in L^p(\Omega;\mathbb{R}^{m\times n})\},$$

endowed with the norm

$$\|u\|_{W^{1,p}(\Omega;\mathbb{R}^m)} := \|u\|_{L^p(\Omega;\mathbb{R}^m)} + \|Du\|_{L^p(\Omega;\mathbb{R}^{m\times n})}$$

with $p \in [1,\infty]$, see [5, 52, 94, 108, 149] for a detailed account on Sobolev spaces. Let $W_0^{1,p}(\Omega;\mathbb{R}^m)$ be the collection of functions with zero trace on $\partial\Omega$ and $W_g^{1,p}(\Omega;\mathbb{R}^m) := g + W_0^{1,p}(\Omega;\mathbb{R}^m)$ for a boundary condition $g \in W^{1,p}(\Omega;\mathbb{R}^m)$. We now consider the integral functionals of the form

$$\mathcal{F}(u) := \int_\Omega f(x,u,Du)\,dx \quad \text{for } u \in W_g^{1,p}(\Omega;\mathbb{R}^m), \tag{1.5}$$

with $f : \Omega\times\mathbb{R}^m\times\mathbb{R}^{m\times n} \to \mathbb{R}$ a Carathéodory integrand, i.e., $f(\cdot,z,A)$ is measurable for all $(z,A) \in \mathbb{R}^m \times \mathbb{R}^{m\times n}$ and $f(x,\cdot,\cdot)$ is continuous for a.e. $x \in \Omega$. We first study the reflexive case $p \in (1,\infty)$, where we are in the setting of (A2) given the weak closedness of $W_g^{1,p}(\Omega;\mathbb{R}^m)$, cf. Remark 1.1.3 c). Therefore, in order to apply the direct method one needs to show that $\mathcal{F}$ is coercive and lower semicontinuous in the weak topology.

We first focus on the lower semicontinuity, which is the most involved of the two, and dates back to the work of Morrey in 1952 [165]. He introduced the following generalized convexity notion, which plays a key role in the characterization of lower semicontinuity.

Definition 1.1.12 (Quasiconvexity). *A Borel-measurable locally bounded function $h : \mathbb{R}^{m\times n} \to \mathbb{R}$ is called quasiconvex, if for every $A \in \mathbb{R}^{m\times n}$*

$$f(A) \le \int_Y f(A + D\varphi(y))\,dy \quad \textit{for all } \varphi \in W_0^{1,\infty}(Y;\mathbb{R}^m),$$

with $Y := (0,1)^n$.

Remark 1.1.13. a) The inequality in the definition of quasiconvexity resembles Jensen's inequality, but it only needs to holds for gradient fields. This shows that quasiconvexity is a weaker notion than convexity, cf. [75, Theorem 5.3]. In fact, when either $n = 1$ or $m = 1$, then quasiconvexity coincides with convexity [75, Theorem 5.3].

b) One can replace the class of test functions by either $C_0^\infty(Y;\mathbb{R}^m)$ or $W_{\mathrm{per}}^{1,\infty}(Y;\mathbb{R}^m)$ without changing the definition of quasiconvexity [75, Remark 5.2 and Proposition 5.13], where $W_{\mathrm{per}}^{1,\infty}(Y;\mathbb{R}^m)$ denotes the Y-periodic functions in $W^{1,\infty}(\mathbb{R}^n;\mathbb{R}^m)$. Additionally, if f is quasiconvex, then it holds for any Lipschitz domain $O \subset \mathbb{R}^n$ by [75, Proposition 5.11] that

$$f(A) \le \frac{1}{|O|}\int_O f(A + D\varphi(y))\,dy \quad \text{for all } \varphi \in W_0^{1,\infty}(O;\mathbb{R}^m).$$

△

It turns out that under suitable growth conditions, quasiconvexity can be used to characterize the (sequential) weak lower semicontinuity of functionals of the form (1.5). The following statement can be found in e.g. [75, Theorem 8.1 and 8.11], with the addition of boundary data being due to [75, Remark 3.19 (ii)].

Theorem 1.1.14 (Characterization of lower semicontinuity). *Let $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to [0,\infty)$ be a Carathéodory integrand that satisfies for $C > 0$*

$$f(x,z,A) \leq C(1+|z|^p+|A|^p) \quad \text{for a.e. } x \in \Omega \text{ and all } (z,A) \in \mathbb{R}^m \times \mathbb{R}^{m\times n}. \tag{1.6}$$

Then, $\mathcal{F}$ in (1.5) *is sequentially weakly lower semicontinuous on $W_g^{1,p}(\Omega;\mathbb{R}^m)$ if and only if $A \mapsto f(x,z,A)$ is quasiconvex for a.e. $x \in \Omega$ and all $z \in \mathbb{R}^m$.*

Remark 1.1.15 (Growth bounds). The non-negativity of the integrand and the upper bound can be further generalized as in e.g. [75, Definition 8.10]. Additionally, we mention that there is the notion of polyconvexity introduced by Ball [23], which is stronger than quasiconvexity, but also provides a sufficient condition for lower semicontinuity when the integrand f takes values in $\mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$, see [75, Theorem 8.16]. This is useful in applications of hyperelasticity, since one can impose the orientation-preservation condition $\det Du > 0$, by setting $f(x,z,A) = \infty$ when $\det A \leq 0$. Well-known examples of such integrands include the ones that are used for neo-Hookean, Mooney-Rivlin and Ogden materials [182, Examples 6.2-6.4]. △

With the lower semicontinuity resolved, it remains to consider the topic of coercivity, which relies on the well-known Poincaré inequality: There is a $C = C(\Omega,n,p) > 0$ such that

$$\|u\|_{L^p(\Omega;\mathbb{R}^m)} \leq C\|Du\|_{L^p(\Omega;\mathbb{R}^{m\times n})} \quad \text{for all } u \in W_0^{1,p}(\Omega;\mathbb{R}^m). \tag{1.7}$$

Now, if f satisfies a growth bound from below of the form

$$\mu|A|^p \leq f(x,z,A) \quad \text{for a.e. } x \in \Omega \text{ and all } (z,A) \in \mathbb{R}^m \times \mathbb{R}^{m\times n}, \tag{1.8}$$

with $\mu > 0$, then it holds that $\mathcal{F}(u) \geq \mu\|Du\|^p_{L^p(\Omega;\mathbb{R}^{m\times n})}$. In particular, if $(u_j)_j \subset W_g^{1,p}(\Omega;\mathbb{R}^m)$ with

$$\|u_j\|_{W^{1,p}(\Omega;\mathbb{R}^m)} \to \infty,$$

then the Poincaré inequality shows that $\mathcal{F}(u_j) \to \infty$. In view of Remark 1.1.3 b), we conclude that $\mathcal{F}$ is coercive. An application of the direct method (Theorem 1.1.4) shows the following.

Theorem 1.1.16 (Existence of minimizers). *Let $p \in (1,\infty)$, $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to [0,\infty)$ be a Carathéodory integrand that satisfies* (1.6) *and* (1.8)*, and assume $A \mapsto f(x,z,A)$ is quasiconvex for a.e. $x \in \Omega$ and all $z \in \mathbb{R}^m$. Then, $\mathcal{F}$ in* (1.5) *admits a minimizer over $W_g^{1,p}(\Omega;\mathbb{R}^m)$.*

Now that the main goal of existence of minimizers is achieved, we will deduce some information about these minimizers inspired by the classical methods in the calculus of variations. Indeed, these methods consisted of showing that minimizers are zeroes of the so-called first variation of $\mathcal{F}$, which is a generalization of the directional derivative of $\mathcal{F}$. In the case of integral functionals, this first variation can be explicitly computed, which leads to a system of PDEs known as the (weak) Euler-Lagrange equations [182, Theorem 3.1]. By $D_z f$ and $D_A f$ we denote the derivative of f with respect to its second and third argument, respectively.

Proposition 1.1.17 (Euler-Lagrange equations). *Let $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory integrand that is continuously differentiable in its second and third argument with*

$$|f(x,z,A)| + |D_z f(x,z,A)| + |D_A f(x,z,A)| \leq C(1+|z|^p+|A|^p)$$

for a.e. $x \in \Omega$ and all $(z,A) \in \mathbb{R}^m \times \mathbb{R}^{m\times n}$. Then, any minimizer of $\mathcal{F}$ in (1.5) *over $W_g^{1,p}(\Omega;\mathbb{R}^m)$ satisfies*

$$\begin{cases} \operatorname{div}[D_A f(x,u,Du)] = D_z f(x,u,Du) & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega, \end{cases} \tag{1.9}$$

in a distributional sense, that is,

$$\int_\Omega D_A f(x,u,Du)\cdot D\varphi + D_z f(x,u,Du)\cdot \varphi \, dx = 0 \quad \textit{for all } \varphi \in C_c^\infty(\Omega;\mathbb{R}^m).$$

Remark 1.1.18. a) Under the assumptions of Theorem 1.1.16 and Proposition 1.1.17, one can combine these statements to establish the existence of weak solutions to the Euler-Lagrange equations (1.9). This provides the solvability for a general class of nonlinear systems of PDEs with Dirichlet boundary conditions.

b) Instead of using the Poincaré inequality in (1.7), one can use the Poincaré-Wirtinger inequality

$$\|u\|_{L^p(\Omega;\mathbb{R}^m)} \le C\|Du\|_{L^p(\Omega;\mathbb{R}^{m\times n})}, \tag{1.10}$$

for all $u \in X := \left\{v \in W^{1,p}(\Omega;\mathbb{R}^m) \,:\, \int_\Omega v\,dx = 0\right\}$. One can then deduce under the same assumptions as Theorem 1.1.16, that $\mathcal{F}$ admits a minimizer on X. In fact, if we assume that $\mathcal{F}$ is invariant under translations by constants, that is, $\mathcal{F}(u+c) = \mathcal{F}(u)$ for $u \in W^{1,p}(\Omega;\mathbb{R}^m)$ and $c \in \mathbb{R}$, then $\mathcal{F}$ admits a minimizer over $W^{1,p}(\Omega;\mathbb{R}^m)$. Under the assumptions of Proposition 1.1.17, the Euler-Lagrange equations for this minimizer then become

$$\begin{cases} \operatorname{div}[D_A f(x,u,Du)] = D_z f(x,u,Du) & \text{in } \Omega, \\ D_A f(x,u,Du)\cdot \nu = 0 & \text{on } \partial\Omega, \end{cases} \tag{1.11}$$

with ν an outward pointing unit normal to $\partial\Omega$. These boundary conditions are called natural boundary conditions, since they are not imposed but arise via the boundary term coming from integration by parts. △

Next, we would like to have an explicit expression of the relaxation of integral functionals, which will be useful to study the behavior of minimizing sequences when lower semicontinuity is not granted, cf. Theorem 1.1.7. Since quasiconvexity is intrinsically linked to lower semicontinuity due to Theorem 1.1.14, it is not surprising that the following concept of quasiconvex envelope is crucial in this regard.

Definition 1.1.19 (Quasiconvex envelope). *Let $h : \mathbb{R}^{m\times n} \to \mathbb{R}$ be locally bounded and Borel-measurable, then its quasiconvex envelope is defined as*

$$h^{\mathrm{qc}}(A) = \sup\left\{\tilde{h}(A) \,:\, \tilde{h} : \mathbb{R}^{m\times n} \to \mathbb{R} \textit{ is quasiconvex, } \tilde{h} \le h\right\}.$$

In case $h : \mathbb{R}^{m\times n} \to \mathbb{R}$ is bounded from below by a quasiconvex function, it holds that h^{qc} is the largest quasiconvex function below h, and the following characterization is valid

$$h^{\mathrm{qc}}(A) = \inf\left\{\int_Y h(A + D\varphi(y))\,dy \,:\, \varphi \in W_0^{1,\infty}(Y;\mathbb{R}^m)\right\},$$

see [75, Theorem 6.9]. We can now state the relaxation result, see e.g. [75, Proposition 9.5 and Theorem 9.8] or [182, Theorem 7.6], where we specify to the simpler case of

$$\mathcal{F}(u) = \int_\Omega f(x,Du)\,dx \quad \text{for } u \in W_g^{1,p}(\Omega;\mathbb{R}^m), \tag{1.12}$$

with $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ a Carathéodory integrand.

Theorem 1.1.20 (Relaxation formula). *Suppose $p \in (1,\infty)$ and $f : \Omega \times \mathbb{R}^{m\times n} \to [0,\infty)$ is a Carathéodory integrand with*

$$\mu|A|^p \leq f(x,A) \leq C(1+|A|^p) \quad \text{for a.e. } x \in \Omega \text{ and all } A \in \mathbb{R}^{m\times n},$$

for some $\mu, C > 0$. Then, the relaxation of $\mathcal{F}$ in (1.12) with respect to the weak convergence in $W^{1,p}_g(\Omega;\mathbb{R}^m)$ is given by

$$\mathcal{F}^{\mathrm{rel}}(u) = \int_\Omega f^{\mathrm{qc}}(x,Du)\,dx \quad \text{for } u \in W^{1,p}_g(\Omega;\mathbb{R}^m),$$

with $f^{\mathrm{qc}}(x,\cdot)$ the quasiconvex envelope of $f(x,\cdot)$.

Note that due to the weak closedness of $W^{1,p}_g(\Omega;\mathbb{R}^m)$, the relaxation would not change if we first extend $\mathcal{F}$ to $W^{1,p}(\Omega;\mathbb{R}^m)$ as ∞. Therefore, in view of the coercivity of $\mathcal{F}$, the benefits of relaxation in Theorem 1.1.7 apply to $\mathcal{F}^{\mathrm{rel}}$.

Linear growth functionals. With the theory in the reflexive case of $p \in (1,\infty)$ rounded off, we now focus our attention to the case $p = 1$, which has a rich history as well (cf. [10, 79, 113, 120, 144, 181, 183]). The leading motivating example for considering this case, is the famous Plateau problem of finding a surface with minimal area and given boundary conditions [118]. Namely, this corresponds to minimizing the area functional

$$\mathcal{F}(u) = \int_\Omega \sqrt{1+|Du|^2}\,dx \quad \text{for } u \in W^{1,1}_g(\Omega),$$

which has an integrand $f(x,A) = \sqrt{1+|A|^2}$ with linear growth. More generally, we are interested in functionals of the form

$$\mathcal{F}(u) = \int_\Omega f(x,Du)\,dx \quad \text{for } u \in W^{1,1}_g(\Omega;\mathbb{R}^m),$$

with $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ a Carathéodory integrand satisfying for $\mu, C > 0$

$$\mu|A| \leq f(x,A) \leq C(1+|A|) \quad \text{for a.e. } x \in \Omega \text{ and all } A \in \mathbb{R}^{m\times n}. \tag{1.13}$$

The immediate difficulty is that $W^{1,1}(\Omega;\mathbb{R}^m)$ is not reflexive, so bounded sequences might not admit weakly convergent subsequences. Hence, the functional $\mathcal{F}$ is not coercive with respect to the weak topology. A way to resolve this is to consider a weaker topology, since bounded sequences $(u_j)_j \subset W^{1,1}(\Omega;\mathbb{R}^m)$ converge up to subsequence to a function $u \in BV(\Omega;\mathbb{R}^m)$ in the weak* topology, or equivalently, the strong L^1-topology; here the space of functions with bounded variation $BV(\Omega;\mathbb{R}^m)$ denotes the L^1-functions with their distributional gradient being a finite Radon-measure, that is,

$$BV(\Omega;\mathbb{R}^m) = \{u \in L^1(\Omega;\mathbb{R}^m) \,:\, Du \in \mathcal{M}(\Omega;\mathbb{R}^{m\times n})\},$$

see e.g. [11, 109, 110, 210]. Therefore, if we extend $\mathcal{F}$ in the following sense

$$\mathcal{F}(u) = \begin{cases} \int_\Omega f(x,Du)\,dx & \text{for } u \in W^{1,1}_g(\Omega;\mathbb{R}^m), \\ \infty & \text{for } u \in BV(\Omega;\mathbb{R}^m) \setminus W^{1,1}_g(\Omega;\mathbb{R}^m), \end{cases}$$

we deduce by the Poincaré inequality and growth bounds in (1.13), that $\mathcal{F}$ is coercive with respect to the strong L^1-topology; note that we could also take the weak* convergence, but the L^1-topology fits into the setting (A1). Unfortunately, by changing the topology we do not have lower semicontinuity anymore, since the domain $W^{1,1}_g(\Omega;\mathbb{R}^m)$ is not even closed in L^1 for bounded sequences.

Therefore, to obtain suitable generalized minimizers that relate to the original problem, we consider the relaxation of $\mathcal{F}$ in the L^1-topology, that is,

$$\mathcal{F}^{\mathrm{rel}}(u) = \inf\left\{\liminf_{j\to\infty}\mathcal{F}(u_j)\ :\ (u_j)_j \subset W^{1,1}_g(\Omega;\mathbb{R}^m),\ u_j \to u \text{ in } L^1(\Omega;\mathbb{R}^m)\right\}, \tag{1.14}$$

for $u \in BV(\Omega;\mathbb{R}^m)$. By Theorem 1.1.7, the functional $\mathcal{F}^{\mathrm{rel}}$ will admit minimizers that are the limit of minimizing sequences of $\mathcal{F}$.

To determine the relaxation, we need to somehow be able to capture the concentration effects that gradients in $W^{1,1}$ can exhibit. To this aim, we denote for $u \in BV(\Omega;\mathbb{R}^m)$ the Radon-Nykodým decomposition

$$Du = \nabla u\,dx + D_* u,$$

where $\nabla u \in L^1(\Omega;\mathbb{R}^{m\times n})$ is the absolutely continuous part of Du with respect to the Lebesgue measure, and $D_* u \in \mathcal{M}(\Omega;\mathbb{R}^{m\times n})$ is the singular part. This singular part is related to these concentration effects and cannot be simply plugged into the integral functional. Instead, one introduces a notion that captures the behavior of the integrand of f at infinity, the so-called recession function, which enables one to account for concentration effects. Explicitly, the recession function $f^\infty : \overline{\Omega}\times\mathbb{R}^{m\times n} \to \mathbb{R}$ is defined as

$$f^\infty(x,A) = \lim_{\substack{(x',A')\to(x,A)\\ t\to\infty}} \frac{f(x,tA)}{t}, \tag{1.15}$$

which is continuous whenever it exists and positively 1-homogeneous in its second argument. With the help of this function, the relaxation of $\mathcal{F}$ can be explicitly determined, see e.g. [182, Lemma 11.1, Proposition 12.24 and Theorem 12.25].

Theorem 1.1.21 (Linear growth relaxation). *Let $f : \Omega\times\mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory integrand that satisfies* (1.13)*, and suppose that f^∞ exists and $A \mapsto f(x,A)$ is quasiconvex for all $x \in \Omega$. Then, the relaxation of $\mathcal{F}$ in* (1.14) *is given by*

$$\mathcal{F}^{\mathrm{rel}}(u) = \int_\Omega f(x,\nabla u)\,dx + \int_\Omega f^\infty\left(x, \frac{dD_* u}{d|D_* u|}\right) d|D_* u| + \int_{\partial\Omega} f^\infty\left(x,(g-u)\otimes\nu\right)\,d\mathcal{H}^{n-1} \quad \textit{for } u \in BV(\Omega;\mathbb{R}^m),$$

with $\frac{dD_* u}{d|D_* u|}$ the Radon-Nykodým derivative of $D_* u$ with respect to its total variation $|D_* u|$, ν an outward normal to $\partial\Omega$ and $\mathcal{H}^{n-1}$ the $(n-1)$-dimensional Hausdorff measure.

Remark 1.1.22. a) The second term in the relaxation is exactly what accounts for the concentration effects that can occur. Additionally, the third term is related to the boundary condition. Indeed, since BV-functions can admit jumps on hypersurfaces, the trace values are not preserved in the relaxation process and are instead penalized by the amount that they deviate from g. Another way to see this, is by extending the functional to a larger domain $\Omega \Subset \Omega'$ and setting all functions equal to g outside Ω (for any extension of g). Then, for $u \in BV(\Omega';\mathbb{R}^m)$ with $u = g$ a.e. in $\Omega' \setminus \Omega$, the variation Du will contain the jump part on the boundary, i.e., the measure $(g-u)\otimes\nu\,\mathcal{H}^{n-1}|_{\partial\Omega}$, which exactly accounts for the penalty term.

b) In the case of the area-integrand $f(x,A) = \sqrt{1+|A|^2}$, which is convex, its recession function is equal to $f^\infty(x,A) = |A|$ and the relaxation becomes

$$\mathcal{F}^{\mathrm{rel}}(u) = \int_\Omega \sqrt{1+|\nabla u|^2}\,dx + |D_* u|(\Omega) + \int_{\partial\Omega} |g-u|\,d\mathcal{H}^{n-1}(x) \quad \text{for } u \in BV(\Omega).$$

c) Under some additional conditions, one can remove the assumption of quasiconvexity, in which case the relaxation will feature the quasiconvex envelope of f and its recession function, see [17]. In this way, it combines the relaxation effect of extending to the space of BV-functions and the quasiconvexification as in Theorem 1.1.20. △

1.2 Fractional calculus and nonlocalities

Dating back to around the same time of the Brachistochrone problem, the first reported mention of fractional calculus is in a private letter from l'Hospital addressed to Leibniz in 1695 [185]. Within this letter, he inquires about Leibniz' notation for the nth derivative of a function $d^n f/dx^n$, wondering what would happen if one formally plugs in $n = \frac{1}{2}$. In the subsequent years, many prominent mathematicians were interested in the concept of non-integer order derivatives, such as Fourier, Euler, Laplace and Lacroix, but it took until 1823 for Abel to publish the first genuine application of fractional calculus [1]. He used it to come up with an elegant solution to the Tautochrone problem, which consisted of finding a curve such that the time it takes for a ball to roll down it is independent of the starting position. This curve is actually closely related to the Brachistochrone curve, since they both are (different) parts of a cycloid, and this draws an unexpected connection between the origins of fractional calculus and the calculus of variations.

After the work of Abel, there were several approaches to rigorously defining fractional derivatives in one dimension, which led to a collection of different possible definitions; the most famous examples include the Riemann-Liouville, Grünwald-Letnikov and Caputo fractional derivative [177, 188]. At the heart of these definitions lies the idea that the fractional derivative should interpolate between the function and its derivative as the fractional order varies from 0 to 1, see Figure 1.1, and that the order of the fractional derivative behaves additively under composition.

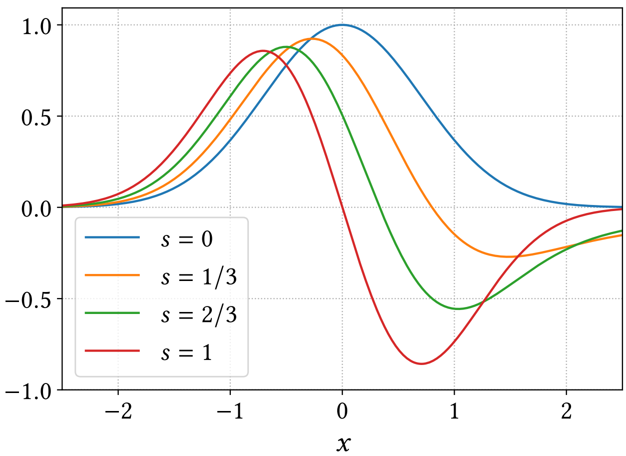

Figure 1.1: The Riemann-Liouville fractional derivative of $x \mapsto e^{-x^2}$ with the fractional order s ranging from 0 to 1.

We note, however, that, depending on the given definition of fractional derivative and admissible class of functions, these properties might not be completely satisfied. Regardless, a common denominator among all fractional derivatives is that they are nonlocal operators defined through integrals, which makes it so that the value of the fractional derivative depends on the values of the function in a larger neighborhood. This is in stark contrast to the local behavior of classical derivatives, and is one of the main reasons for the recent interest in using fractional derivatives in applications. Indeed, fractional derivatives can be used to model long-range interactions or memory-effects with applications in physics, biology, chemistry and control theory [132, 177].

Fractional Laplacian. Turning to the multidimensional setting, it becomes apparent that the literature has almost exclusively focused on the fractional Laplacian as the central fractional differential object [146, 151]. For $s \in (0,2)$ and $\varphi \in C_c^\infty(\mathbb{R}^n)$ it is defined as a principal value integral

$$(-\Delta)^{s/2}\varphi(x) := \nu_{n,s} \lim_{r \downarrow 0} \int_{B_r(x)^c} \frac{\varphi(x)-\varphi(y)}{|x-y|^{n+s}}\, dy \quad x \in \mathbb{R}^n,$$

with $B_r(x)$ the ball with radius $r > 0$ around x and $\nu_{n,s}$ the normalizing constant given by

$$\nu_{n,s} := \frac{2^s \Gamma((s+n)/2)}{\pi^{n/2}|\Gamma(-s/2)|},$$

with Γ the Gamma function. The fractional Laplacian interpolates between a function and its (negative) Laplacian and has been applied in fractional versions of the diffusion equation, porous medium equation, Cahn-Hilliard equation, Schrödinger equation and many more, cf. [151] and the references therein. Moreover, it plays a key role in probability theory as well since it is the generator of the s-stable Lévy jump process in the same way the Laplacian is the generator of Brownian motion [157]. Beyond the applications, the fractional Laplacian has served as the prototypical example in the development of the mathematical theory around nonlocal operators, such as in regularity theory [124, 184], the development of extension techniques [57] and the study of nonlocal Neumann boundary conditions [100]. Additionally, the related fractional p-Laplacian plays an analogous role in the nonlinear theory.

The study of the fractional Laplacian and related (nonlinear) operators is carried out with the help of the Gagliardo fractional Sobolev spaces defined as

$$W^{s,p}(\mathbb{R}^n) := \left\{ u \in L^p(\mathbb{R}^n) \;:\; [u]_{W^{s,p}(\mathbb{R}^n)} := \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(x)-u(y)|^p}{|x-y|^{n+ps}}\, dx\, dy \right)^{1/p} < \infty \right\},$$

with

$$\|u\|_{W^{s,p}(\mathbb{R}^n)} := \|u\|_{L^p(\mathbb{R}^n)} + [u]_{W^{s,p}(\mathbb{R}^n)},$$

for $s \in (0,1)$ and $p \in [1,\infty)$, see e.g., [5, 96]. These spaces originally arose as trace spaces of the classical Sobolev spaces, but have recently been more prominently applied in the context of fractional problems. The fractional spaces also possess useful properties like Sobolev inequalities and continuous and compact embeddings. Moreover, it turns out that minimizing the Gagliardo semi-norm $[\cdot]_{W^{s,p}(\mathbb{R}^n)}$ among functions satisfying a Dirichlet condition in the complement of a bounded domain, is equivalent to solving the fractional p-Laplace equation; in fact, when $p = 2$, the Gagliardo semi-norm is even equal to the L^2-norm of the fractional Laplacian up to a constant [96, Proposition 3.4]. This reformulation of the fractional Laplace equation opens up the way for applying the tools from the calculus of variations to obtain rigorous existence results, and has inspired the study of more general nonlocal functionals of double-integral type, where aspects such as coercivity and lower semicontinuity [33,34,107,166,174], relaxation [143,164], and different localizations [8,35,47,158] are considered; these models are closely related to the bond-based models in peridynamics, cf. Section 1.3.1.

1.2.1 The Riesz fractional gradient

While the models revolving around the fractional Laplacian have received a lot of attention, it is surprising that a fractional analogue of the gradient operator has largely been missing in the literature; indeed, in the classical vector calculus and calculus of variations it is the gradient that is the fundamental object, not the Laplacian. This issue was recently resolved with the introduction of the Riesz fractional gradient by Shieh & Spector in 2015 [193], which is defined for $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $s \in (0,1)$ as

$$D^s\varphi(x) := c_{n,s}\int_{\mathbb{R}^n} \frac{\varphi(x)-\varphi(y)}{|x-y|^{n+s}}\frac{x-y}{|x-y|}\,dy \quad x \in \mathbb{R}^n,$$

with the constant

$$c_{n,s} := 2^s\pi^{-n/2}\frac{\Gamma((n+s+1)/2)}{\Gamma((1-s)/2)}.$$

This fractional gradient interpolates between the Riesz transform and the classical gradient as s varies from 0 to 1, which is substantially different from the one-dimensional fractional derivatives considered at the start of the section. It was shown by Šilhavý [208] that the Riesz fractional gradient is the unique operator that satisfies the natural physical requirements of rotation and translation invariance, and is homogeneous of degree s; in this sense it can be considered as a canonical fractional derivative, as opposed to the various different definitions in one dimension. This new perspective of the fractional gradient has opened up the possibility to study new types of fractional problems, and has inspired a vast amount of works in recent years, see e.g. [28,31,66,92,93,140,153] among many others.

Since the fractional gradient is a central object in this thesis, we will expand on some of its properties and relations with other operators, and cover the associated function spaces and variational problems. Firstly, it is possible to define the fractional divergence of $\psi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$ as

$$\operatorname{div}^s\psi(x) := c_{n,s}\int_{\mathbb{R}^n} \frac{\varphi(x)-\varphi(y)}{|x-y|^{n+s}}\cdot\frac{x-y}{|x-y|}\,dy \quad x \in \mathbb{R}^n,$$

which enables one to formulate the following extension the classical formula $-\operatorname{div}\circ D = (-\Delta)$,

$$-\operatorname{div}^s\circ D^t = (-\Delta)^{(s+t)/2}$$

for $s,t \in (0,1)$ see [208, Theorem 5.3]. In fact, this and many other identities can be quickly deduced from the characterization of the fractional operators in Fourier space. With the convention of the Fourier transform given by

$$\widehat{u}(\xi) := \int_{\mathbb{R}^n} u(x)e^{-2\pi i x\cdot\xi}\,dx \quad \xi \in \mathbb{R}^n,$$

for functions $u \in L^1(\mathbb{R}^n)$ [122,123], it holds for $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\psi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$ that

$$\widehat{D^s\varphi}(\xi) = \frac{2\pi i\xi}{|2\pi\xi|^{1-s}}\widehat{\varphi}(\xi), \quad \widehat{(-\Delta)^{s/2}\varphi}(\xi) = |2\pi\xi|^s\widehat{\varphi}(\xi) \quad \text{and} \quad \widehat{\operatorname{div}^s\psi}(\xi) = \frac{2\pi i\xi}{|2\pi\xi|^{1-s}}\cdot\widehat{\psi}(\xi), \quad (1.16)$$

for $\xi \in \mathbb{R}^n\setminus\{0\}$ (cf. [194]). In light of these identities and Plancherel's theorem, it is straightforward to deduce the fractional integration by parts formula

$$\int_{\mathbb{R}^n} D^s\varphi\cdot\psi\,dx = -\int_{\mathbb{R}^n}\varphi\,\operatorname{div}^s\psi\,dx,$$

which can be used to extend the definition of the fractional gradient to L^p-spaces for $p \in [1,\infty]$.

As a result, one can naturally define a class of fractional Sobolev spaces for $s \in (0,1)$ and $p \in [1,\infty]$ as

$$H^{s,p}(\mathbb{R}^n) := \{u \in L^p(\mathbb{R}^n) : D^s u \in L^p(\mathbb{R}^n;\mathbb{R}^n)\}$$

with the norm

$$\|u\|_{H^{s,p}(\mathbb{R}^n)} := \|u\|_{L^p(\mathbb{R}^n)} + \|D^s u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)}.$$

It was shown by Shieh & Spector [193, Theorem 1.7] together with the density result in [54, Theorem A.1] (see also [140, Theorem 2.7]), that these spaces coincides with the well-known Bessel potential spaces when $p \in (1, \infty)$, that is

$$H^{s,p}(\mathbb{R}^n) = \left\{u \in L^p(\mathbb{R}^n) \,:\, (\langle\cdot\rangle^s\, \widehat{u})^\vee \in L^p(\mathbb{R}^n)\right\},$$

where $\langle\cdot\rangle = \sqrt{1 + |\cdot|^2}$ and the (inverse) Fourier transform should be understood in the sense of tempered distributions, see [5, 123, 204] for more on these spaces. The Bessel potential spaces coincide with the Gagliardo spaces when $p = 2$, but are different in general, and they appear naturally in harmonic analysis [136] and the regularity theory of the fractional Laplacian [124]. We remark that the Gagliardo and Bessel potential spaces can also be unified within the scale of Triebel-Lizorkin spaces, with the identities $W^{s,p}(\mathbb{R}^n) = F^s_{p,p}(\mathbb{R}^n)$ and $H^{s,p}(\mathbb{R}^n) = F^s_{p,2}(\mathbb{R}^n)$ for $p \in (1, \infty)$, see [187, 204].

For variational problems involving the fractional gradient, it is common to consider a bounded open domain $\Omega \subset \mathbb{R}^n$, and work with functions that are zero in the complement, that is,

$$H^{s,p}_0(\Omega) := \{u \in H^{s,p}(\mathbb{R}^n) \,:\, u = 0 \text{ a.e. in } \Omega^c\},$$

or the more general affine spaces $H^{s,p}_g(\Omega) := g + H^{s,p}_0(\Omega)$ for a given $g \in H^{s,p}(\mathbb{R}^n)$. Shieh & Spector [193, 194] (see also [28, 140]) developed the essential technical tools for utilizing these spaces, such as continuous and compact embeddings, and a fractional Poincaré inequality of the form

$$\|u\|_{L^p(\mathbb{R}^n)} \leq C\|D^s u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \quad \text{for all } u \in H^{s,p}_0(\Omega), \tag{1.17}$$

with a constant $C = C(\Omega, n, p, s) > 0$ (cf. [193, Theorem 3.3]). They proved this inequality in an elegant manner by using that

$$\varphi(x) = c_{n,-s} \int_{\mathbb{R}^n} D^s\varphi(y) \cdot \frac{x - y}{|x - y|^{n-s+1}}\, dy \quad x \in \mathbb{R}^n, \tag{1.18}$$

for all $\varphi \in C^\infty_c(\mathbb{R}^n)$, which is known as the fractional fundamental theorem of calculus [193, Theorem 1.12], [179, Proposition 15.8].

With these technical tools at hand, one can consider the minimization of functionals of the form

$$\mathcal{F}_s(u) := \int_{\mathbb{R}^n} f(x, u, D^s u)\, dx \quad \text{for } u \in H^{s,p}_g(\Omega; \mathbb{R}^m), \tag{1.19}$$

with $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ a Carathéodory integrand and $p \in (1, \infty)$. In this case, we are in the setting (A2) of a weakly closed subset of a separable reflexive Banach space. We note the similarity in structure with the classical integral functionals in (1.5), although $\mathcal{F}_s$ is inherently nonlocal. Coercivity of the functional $\mathcal{F}_s$ follows in the same manner as the classical case by utilizing the fractional Poincaré inequality in (1.17) and imposing the bound

$$\mu|A|^p \leq f(x, z, A) \quad \text{for a.e. } x \in \mathbb{R}^n \text{ and all } (z, A) \in \mathbb{R}^m \times \mathbb{R}^{m\times n} \tag{1.20}$$

with $\mu > 0$; here, it is important that $\mathcal{F}_s$ consists of an integral over the full space and not just over Ω. It remains to study the weak lower semicontinuity of the functional to obtain the existence of minimizers. Shieh & Spector have shown that convexity gives a sufficient condition for lower semicontinuity in the scalar case $m = 1$ [193, 194], whereas Bellido, Cueto & Mora-Corral considered the vectorial case and proved that polyconvexity also provides a sufficient condition [28], cf. Remark 1.1.15. A complete characterization was first given in [140, Theorem 1.1] by Kreisbeck and the author, and surprisingly identifies quasiconvexity as the crucial notion in the fractional case as well. Throughout this section, we assume the technical condition $|\partial\Omega| = 0$.

Theorem 1.2.1 (Characterization of lower semicontinuity). *Let $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to [0,\infty)$ be a Carathéodory integrand that satisfies for $C > 0$ and $a \in L^1(\mathbb{R}^n)$*

$$f(x,z,A) \leq C(a(x) + |z|^p + |A|^p) \quad \text{for a.e. } x \in \Omega \text{ and all } (z,A) \in \mathbb{R}^m \times \mathbb{R}^{m\times n}. \tag{1.21}$$

Then, $\mathcal{F}_s$ in (1.19) *is sequentially weakly lower semicontinuous on $H^{s,p}_g(\Omega;\mathbb{R}^m)$ if and only if $A \mapsto f(x,z,A)$ is quasiconvex for a.e. $x \in \Omega$ and all $z \in \mathbb{R}^m$.*

This theorem constitutes a fractional analogue to Theorem 1.1.14 and it is quite remarkable that quasiconvexity appears again in this context. Indeed, a more natural convexity notion in the fractional case would be the following.

Definition 1.2.2 (D^s-quasiconvexity). *A Borel-measurable locally bounded function $h : \mathbb{R}^{m\times n} \to \mathbb{R}$ is called D^s-quasiconvex if for all $A \in \mathbb{R}^{m\times n}$*

$$h(A) \leq \int_Y h(A + D^s\varphi(y))\,dy \quad \text{for all } \varphi \in H^{s,\infty}_{\text{per}}(Y;\mathbb{R}^m),$$

with $H^{s,\infty}_{\text{per}}(Y;\mathbb{R}^m)$ the space of Y-periodic functions in $H^{s,\infty}(\mathbb{R}^n;\mathbb{R}^m)$.

It turns out that this seemingly new notion is equivalent to classical quasiconvexity (Definition 1.1.12), and, therefore, can also be used in the characterization of the weak lower semicontinuity of $\mathcal{F}_s$, [140, Corollary 4.8]. This fact and the proof of Theorem 1.2.1 both rely on the following identities

$$D^s\varphi = D(I_{1-s} * \varphi) \quad \text{and} \quad D\varphi = D^s(-\Delta)^{\frac{1-s}{2}}\varphi, \tag{1.22}$$

for $\varphi \in C_c^\infty(\mathbb{R}^n)$, where $I_t \in L^1_{\text{loc}}(\mathbb{R}^n)$ denotes the Riesz potential

$$I_t := \nu_{n,-t}|\cdot|^{t-n},$$

for $t \in (0,n)$, and arises as an inverse to the fractional Laplacian. One can utilize (1.22) to transform the fractional gradient into a classical gradient and back, making it possible to move between the two settings. This provides a systematic method for proving results in the fractional case by reducing them to the well-established results concerning the local gradient and explains why quasiconvexity appears in Theorem 1.2.1. However, care should be taken since the procedure involves nonlocal operations, which do not preserve complementary values. Beyond technical difficulties, the complementary values also account for the fact that in Theorem 1.2.1 quasiconvexity is only required in Ω. Namely, for weakly converging sequences in $H^{s,p}_g(\Omega;\mathbb{R}^m)$ it holds that the fractional gradients converge strongly in the complement of Ω [140, Lemma 2.12], which eliminates the need for any convexity notion there to get lower semicontinuity.

Combining the coercivity and lower semicontinuity of the fractional functionals, yields the existence of minimizers via the direct method (Theorem 1.1.4).

Theorem 1.2.3 (Existence of minimizers). *Let $p \in (1,\infty)$, $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to [0,\infty)$ be a Carathéodory integrand that satisfies* (1.20) *and* (1.21)*, and assume $A \mapsto f(x,z,A)$ is quasiconvex for a.e. $x \in \Omega$ and all $z \in \mathbb{R}^m$. Then, $\mathcal{F}_s$ in* (1.19) *admits a minimizer over $H^{s,p}_g(\Omega;\mathbb{R}^m)$.*

These minimizers also satisfy certain Euler-Lagrange equations, which consist of a system of fractional partial differential equations subject to complementary-value conditions, see [28, 193].

Proposition 1.2.4 (Euler-Lagrange equations). *Let $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory integrand that is continuously differentiable in its second and third argument with*

$$|f(x,z,A)| + |D_z f(x,z,A)| + |D_A f(x,z,A)| \leq C(a(x) + |z|^p + |A|^p),$$

for a.e. $x \in \mathbb{R}^n$ and all $(z, A) \in \mathbb{R}^m \times \mathbb{R}^{m\times n}$ with $C > 0$ and $a \in L^1(\mathbb{R}^n)$. Then, any minimizer of $\mathcal{F}_s$ in (1.19) *over $H^{s,p}_g(\Omega;\mathbb{R}^m)$ satisfies*

$$\begin{cases} \operatorname{div}^s[D_A f(x,u,D^s u)] = D_z f(x,u,D^s u) & \text{in } \Omega, \\ u = g & \text{in } \Omega^c, \end{cases} \tag{1.23}$$

in a distributional sense, that is,

$$\int_{\mathbb{R}^n} D_A f(x,u,D^s u)\cdot D^s\varphi + D_z f(x,u,D^s u)\cdot\varphi\,dx = 0 \quad \text{for all } \varphi \in C_c^\infty(\Omega;\mathbb{R}^m).$$

Remark 1.2.5. a) In the case with $m = 1$, $p = 2$, and $f(x,z,A) = |A|^2$, we can use $\operatorname{div}^s \circ D^s = -(-\Delta)^s$ to deduce that the Euler-Lagrange equation becomes

$$\begin{cases} (-\Delta)^s u = 0 & \text{in } \Omega, \\ u = g & \text{in } \Omega^c, \end{cases}$$

which is the fractional Laplace equation. This is not surprising since (1.16) and Plancherel's identity show that $\|D^s u\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^n)} = \|(-\Delta)^{s/2}u\|^2_{L^2(\mathbb{R}^n)}$, which means that the minimization of both of these quantities is equivalent. As a result, we also deduce that the L^2-norm of the fractional gradient is proportional to the Gagliardo semi-norm, which illuminates the identity $W^{s,2}(\mathbb{R}^n) = H^{s,2}(\mathbb{R}^n)$.

b) When $m = 1$, $p \in (1,\infty)$ and we minimize the L^p-norm of the fractional gradient, that is, $f(x,z,A) = |A|^p$, then the Euler-Lagrange equation becomes

$$\begin{cases} \operatorname{div}^s\left(|D^s u|^{p-2}D^s u\right) = 0 & \text{in } \Omega, \\ u = g & \text{in } \Omega^c, \end{cases}$$

which can be seen as a fractional analogue of the p-Laplace equation. We note, however, that this fractional p-Laplace operator is different from the one that arises by minimizing the Gagliardo semi-norm when $p \neq 2$. △

The previous two propositions can be combined to deduce the existence of weak solutions to the general system of fractional PDEs in (1.23). In the case where lower semicontinuity is not satisfied, one can use relaxation (cf. Definition 1.1.5) to recover the existence of minimizers. We restrict to the simpler case without (x,u)-dependence

$$\mathcal{F}_s(u) = \int_{\mathbb{R}^n} f(D^s u)\,dx \quad \text{for } u \in H^{s,p}_g(\Omega;\mathbb{R}^m), \tag{1.24}$$

with $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ continuous, which is covered in [140, Theorem 1.2].

Theorem 1.2.6 (Relaxation formula). *Let $p \in (1,\infty)$ and $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ be continuous with*

$$\mu|A|^p \leq f(A) \leq C|A|^p \quad \text{for all } A \in \mathbb{R}^{m\times n},$$

with $\mu, C > 0$. Then, the relaxation of $\mathcal{F}_s$ in (1.24) *with respect to the weak convergence in $H^{s,p}_g(\Omega;\mathbb{R}^m)$ is given by*

$$\mathcal{F}_s^{\mathrm{rel}}(u) = \int_\Omega f^{\mathrm{qc}}(D^s u)\,dx + \int_{\Omega^c} f(D^s u)\,dx \quad \text{for } u \in H^{s,p}_g(\Omega;\mathbb{R}^m).$$

It is natural that the relaxation features the quasiconvex envelope, since we know that quasiconvexity is also the correct notion in the fractional case in view of Theorem 1.2.1 and (1.22). Additionally, the integral over Ω^c remains unchanged in the relaxation process, precisely because of the property that fractional gradients of weakly converging sequences converge strongly there [140, Lemma 2.3]. Overall, this rounds off a collection of results regarding the fractional functionals that is quite comprehensive with regard to the direct method in the reflexive range $p \in (1, \infty)$, and the similarities to the classical case in Section 1.1 are noteworthy. A natural next step, inspired by the results on classical linear growth functionals in Section 1.1, is to consider the fractional functionals in the case $p = 1$. This is exactly the topic of Chapter 2, where the functionals are suitably extended via relaxation to the space of functions with bounded fractional variation, see Section 1.4 below for a more in-depth description.

1.2.2 Finite-horizon and general nonlocal gradients

Beyond the models based on the Riesz fractional gradient, there are promising research directions related to variational problems involving more general nonlocal gradients. Among these, are the so-called finite-horizon fractional gradients introduced by Bellido, Cueto & Mora-Corral in [31], which are considered for applications in continuum mechanics and peridynamic models. In contrast to the fractional operators that require integration over the full space, the values of the finite-horizon fractional gradient only depend on a radial neighborhood of size $\delta > 0$, which is called the horizon. This makes the gradient more realistic in applications and more tractable for numerical simulations.

Explicitly, for $s \in (0, 1)$ the finite-horizon fractional gradient of $\varphi \in C^\infty(\mathbb{R}^n)$ is defined as

$$D_\delta^s \varphi(x) := c_{n,s,\delta} \int_{\mathbb{R}^n} \frac{\varphi(x) - \varphi(y)}{|x - y|^{n+s}} \frac{x - y}{|x - y|} w_\delta(x - y)\, dy \quad x \in \mathbb{R}^n, \tag{1.25}$$

where $w_\delta : \mathbb{R}^n \longrightarrow [0, \infty)$ is a cut-off function and $c_{n,s,\delta} > 0$ a normalizing constant satisfying:

(H1) w_δ is radial, radially decreasing, smooth and has support in $B_\delta(0)$;

(H2) there is a $\lambda \in (0, 1)$ such that $w_\delta = 1$ on $B_{\lambda\delta}(0)$;

(H3) it holds that
$$c_{n,s,\delta} \int_{B_\delta(0)} \frac{w_\delta(x)}{|x|^{n+s-1}}\, dx = n.$$

The finite-horizon divergence $\operatorname{div}_\delta^s$ can be defined analogously to the fractional divergence and acts as a dual operator in the sense that

$$\int_\Omega D_\delta^s \varphi \cdot \psi \, dx = -\int_{\Omega_\delta} \varphi \, \operatorname{div}_\delta^s \psi \, dx,$$

for all $\varphi \in C^\infty(\mathbb{R}^n)$ and $\psi \in C_c^\infty(\Omega; \mathbb{R}^n)$ with Ω_δ the nonlocal closure of Ω given by $\Omega_\delta := \Omega + B_\delta(0)$. This integration by parts formula can be used to define a distributional version of the finite-horizon fractional gradient and shows that in order to make sense of $D_\delta^s \varphi$ on Ω, it is only required to know the values of φ in the nonlocal closure Ω_δ. The naturally associated Sobolev spaces for an open subset $\Omega \subset \mathbb{R}^n$ become

$$H^{s,p,\delta}(\Omega) := \left\{ u \in L^p(\Omega_\delta) \; : \; D_\delta^s u \in L^p(\Omega; \mathbb{R}^n) \right\}, \tag{1.26}$$

with the norm
$$\|u\|_{H^{s,p,\delta}(\Omega)} := \|u\|_{L^p(\Omega_\delta)} + \|D_\delta^s u\|_{L^p(\Omega;\mathbb{R}^n)}.$$

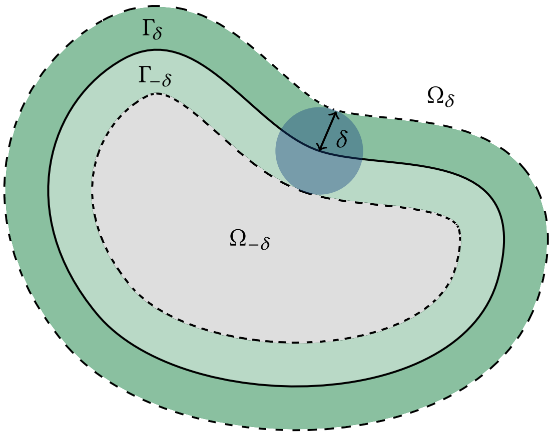

Figure 1.2: Illustration of a set $\Omega \subset \mathbb{R}^n$ with its expansion Ω_δ, the outer and inner collar regions Γ_δ (green) and $\Gamma_{-\delta}$ (light green), and the reduced set $\Omega_{-\delta}$ (gray).

Additionally, the relevant boundary conditions for these problems, consist of prescribed values in a tubular neighborhood around the domain; precisely, we consider the spaces

$$H_0^{s,p,\delta}(\Omega) := \left\{u \in H^{s,p,\delta}(\Omega) \,:\, u = 0 \text{ a.e. in } \Gamma_{\pm\delta} := \Omega_\delta \setminus \overline{\Omega_{-\delta}}\right\},$$

where $\Omega_{-\delta} := \{x \in \Omega \,:\, \operatorname{dist}(x, \partial\Omega) > \delta\}$, and, for $g \in H^{s,p,\delta}(\Omega)$, the affine spaces $H_g^{s,p,\delta}(\Omega) = g + H_0^{s,p,\delta}(\Omega)$. We also introduce the notation

$$\Gamma_{-\delta} := \Omega \setminus \overline{\Omega_{-\delta}} \quad \text{and} \quad \Gamma_\delta := \Omega_\delta \setminus \overline{\Omega},$$

which yields $\Gamma_{\pm\delta} = \Gamma_{-\delta} \cup \Gamma_\delta \cup \partial\Omega$, see Figure 1.2 for a visual representation. The reason why one imposes a condition on the layer $\Gamma_{\pm\delta}$ of thickness 2δ, is to ensure that the finite-horizon fractional gradient is supported in Ω; moreover, it is also natural from the point of view of the Euler-Lagrange equations in (1.30), since it features a composition of two nonlocal operators with horizon δ.

This set-up facilitates the study of integral functionals of the form

$$\mathcal{F}_\delta^s(u) := \int_\Omega f(x, u, D_\delta^s u)\, dx \quad \text{for } u \in H_g^{s,p,\delta}(\Omega; \mathbb{R}^m), \tag{1.27}$$

with $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ a Carathéodory integrand and $p \in (1, \infty)$. Bellido, Cueto & Mora-Corral introduced this class in [31] and developed the necessary tools regarding the function spaces, such as embedding results and the following Poincaré inequality

$$\|u\|_{L^p(\Omega)} \leq C\|D_\delta^s u\|_{L^p(\Omega;\mathbb{R}^n)} \quad \text{for all } u \in H_0^{s,p,\delta}(\Omega). \tag{1.28}$$

To prove these results, they carried out a delicate analysis in Fourier space to deduce that there is a function $V_\delta^s \in C^\infty(\mathbb{R}^n \setminus \{0\}; \mathbb{R}^n)$ with suitable properties and such that

$$\varphi(x) = \int_{\mathbb{R}^n} D_\delta^s \varphi(y) \cdot V_\delta^s(x - y)\, dy \quad x \in \mathbb{R}^n, \tag{1.29}$$

for all $\varphi \in C_c^\infty(\mathbb{R}^n)$; this identity mirrors the fractional fundamental theorem of calculus in (1.18). As a consequence, they proved the existence of minimizers of $\mathcal{F}_\delta^s$ by providing sufficient conditions for the lower semicontinuity of the functionals in (1.27) based on convexity [31] and polyconvexity [30]. Furthermore, the minimizers of the functional satisfy a nonlocal PDE of the form

$$\begin{cases} \operatorname{div}_\delta^s[D_A f(x,u,D_\delta^s u)] = D_z f(x,u,D_\delta^s u) & \text{in } \Omega_{-\delta}, \\ u = g & \text{in } \Gamma_{\pm\delta}, \end{cases} \tag{1.30}$$

under suitable assumptions [31, Theorem 8.2]. Further aspects of the finite-horizon models will be considered in Chapters 3 and 4, including a full characterization of lower semicontinuity, Γ-convergence results, localization and asymptotics with respect to the fractional order, a characterization of functions with $D_\delta^s u = 0$, and a novel nonlocal Neumann-type problem, see Section 1.4 for more details.

General nonlocal gradients. We close this section by discussing a more general class of nonlocal gradients, unifying both settings of the Riesz and finite-horizon fractional gradient. They are of the form

$$D_\rho \varphi(x) := \int_{\mathbb{R}^n} \frac{\varphi(x)-\varphi(y)}{|x-y|}\frac{x-y}{|x-y|}\rho(x-y)\,dy, \tag{1.31}$$

for $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\rho \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ a non-negative radial function that satisfies

$$\inf_{B_\varepsilon(0)} \rho > 0 \text{ for some } \varepsilon > 0 \quad \text{and} \quad \int_{\mathbb{R}^n} \min\{1,|x|^{-1}\}\rho(x)\,dx < \infty. \tag{1.32}$$

The cases

$$\rho^s := c_{n,s}\frac{1}{|\cdot|^{n+s-1}} \quad \text{and} \quad \rho_\delta^s := c_{n,s,\delta}\frac{w_\delta}{|\cdot|^{n+s-1}},$$

recover the gradients $D_{\rho^s} = D^s$ and $D_{\rho_\delta^s} = D_\delta^s$. Operators similar to D_ρ have been considered in [16, 92, 93, 102, 156, 161], but, apart from the case of the Riesz and finite-horizon fractional gradient, the functional analytic tools for studying variational problems involving D_ρ are still absent. In Chapter 5, we fill this gap by investigating the function spaces associated to the nonlocal gradient D_ρ and determining almost minimal conditions on ρ in order to have Poincaré inequalities and suitable embedding results, see Section 1.4. This enables one to prove the existence of minimizers for variational problems involving more general nonlocal gradients. Furthermore, we build upon these results in Chapter 6, by proving rigorous localization results for the nonlocal energies as we scale the kernel ρ in an isotropic way. This establishes the compatibility of a class of state-based peridynamics models with classical models in the vanishing horizon limit, see Section 1.3.1 and 1.4 for more context. To fully understand the horizon-dependence, we also consider in Chapter 6 the case of diverging horizons, which leads to the models based on the Riesz fractional gradient.

1.3 Motivation and applications

Beyond their theoretical interest, we present in this section two applications of variational models involving nonlocal gradients in the field of peridynamics and image denoising. We give an overview of each setting and draw connections with the theoretical results of the thesis.

1.3.1 Peridynamics

The theory of peridynamics was initiated by Silling in 2000 [196] and has led to a vast literature including many applied contributions, see, e.g., the books [44, 91, 116, 154] and the references therein. It constitutes a nonlocal formulation of continuum mechanics, where forces are measured through the interaction of particles; only those particles that are within a certain distance, known as the horizon $\delta > 0$, are considered. In this way, the models make sense over bounded domains, and, as the horizon vanishes, lead to localized models. Besides the ability to incorporate long-range interactions, one of the key advantages of the peridynamic framework is that it does not make use of (full) derivatives and allows for less regular deformations. Therefore, the appearance of discontinuities such as cavities and fractures in elastic materials is naturally possible via this approach as opposed to the classical models based on gradients. We will present the theory from an energy-based perspective, but one can also study the associated time-dependent equations of motion, see e.g., [196, 197] and the overview in [71, Section 1.3].

Bond-based peridynamics. The original formulation of peridynamics consisted of the bond-based framework [196], where the forces between the individual bonds of particles are averaged. Precisely, one considers an interior domain $\Omega \subset \mathbb{R}^n$ and its nonlocal closure $\Omega_\delta = \Omega + B_\delta(0)$, where all the interactions take place. Then, the bond-based energy associated to a deformation $u : \Omega_\delta \to \mathbb{R}^m$ is given by

$$\mathcal{E}_{\mathrm{b}}(u) := \int_{\Omega_\delta} \int_{\Omega_\delta \cap B_\delta(x)} w(x-y, u(x)-u(y))\, dy\, dx, \tag{1.33}$$

with $w : B_\delta(0) \times \mathbb{R}^m \to [0,\infty)$ a suitable density that assigns an energy to each bond. Due to Fubini's theorem, one may always assume that $w(-h,-\eta) = w(h,\eta)$, and, if we set $w(h,\eta) = 0$ for $|h| \geq \delta$, the energy can be rewritten as

$$\mathcal{E}_{\mathrm{b}}(u) = \int_{\Omega_\delta} \int_{\Omega_\delta} w(x-y, u(x)-u(y))\, dy\, dx.$$

The bond-based formulation has been incredibly successful, as it is a fairly simple model that is able capture discontinuity effects that are relevant in the modeling of elastic materials. From a mathematical viewpoint, the bond-based energies give rise to double-integral functionals and various aspects such as lower semicontinuity and coercivity [33, 34, 107, 166, 174], relaxation [143, 164] and localization [8, 35, 47, 158] have been addressed for them. Despite its success, the bond-based formulation faces a few shortcomings. The first is that this approach can only model materials with certain constraints, for example, those with Poisson ratio equal to $\frac{1}{4}$ in the linear case [195, 197]. Secondly, in the vanishing horizon limit, the bond-based functionals recover a rather restrictive class of local models, and not the ones that are commonly used in nonlinear hyperelasticity as in Remark 1.1.15, cf. [27, 160].

State-based peridynamics. Due to these drawbacks, a more general framework was introduced in [197], which is called state-based peridynamics. Instead of assigning an energy to each individual bond, one considers all interactions within the horizon distance, which is called the deformation state, and assigns an energy to this. Precisely, for any deformation $u : \mathbb{R}^n \to \mathbb{R}^m$, the state at $x \in \mathbb{R}^n$ is given by the function

$$S_x^u : B_\delta(0) \to \mathbb{R}^m, \quad S_x^u(h) := u(x+h) - u(x).$$

The space of all possible states, which is a suitable subspace of all functions from $B_\delta(0)$ to $\mathbb{R}^m$, is denoted by $\mathcal{V}$. The state-based energies are then defined as

$$\mathcal{E}_{\mathrm{s}}(u) := \int_{\Omega_\delta} \mathcal{W}(x, S_x^u)\, dx, \tag{1.34}$$

where $\mathcal{W} : \Omega_\delta \times \mathcal{V} \to [0,\infty)$ is a density that assigns an energy to each state. Strictly speaking, $\mathcal{W}$ is a functional, since it takes functions as arguments, and this makes the state-based formulation extremely general. For example, if we take

$$\mathcal{W}(x,S) := \int_{(\Omega_\delta - x)\cap B_\delta(0)} w(h, S(h))\, dh \qquad \text{for } (x,S) \in \Omega_\delta \times \mathcal{V},$$

then, after the substitution $h = y - x$, (1.34) reduces to the bond-based energy in (1.33). If, instead, we make the choice

$$\mathcal{W}(x,S) := W\left(x, \int_{B_\delta(0)} \frac{S(h)}{|h|} \otimes \frac{h}{|h|}\rho(h)\,dh\right) \qquad \text{for } (x,S) \in \Omega_\delta \times \mathcal{V},$$

with $W : \Omega_\delta \times \mathbb{R}^{m\times n} \to [0,\infty)$ a suitable Carathéodory integrand, and ρ a radial kernel that defines a nonlocal gradient as in (1.31), then (1.34) turns into

$$\mathcal{E}_s(u) = \int_{\Omega_\delta} W(x, D_\rho u)\,dx. \tag{1.35}$$

Note that for these models the functions need to be defined on $\Omega_{2\delta}$ to maintain the symmetric integration domain $B_\delta(0)$ in the nonlocal gradient; this is also the case for the finite-horizon fractional models in (1.27), where a slightly different, yet equivalent, convention is used with $\Omega_{-\delta}$ being the interior domain. It actually turns out that the models based on nonlocal gradients were already proposed in [197, Section 18], albeit with a different interpretation.

These observations show that the models based on nonlocal gradients provide a mathematically rigorous example of state-based peridynamics, which are different from the bond-based models. In addition, they are compatible with fracture and cavitation in materials, for example, in the setting with the fractional spaces $H^{s,p}$ when $sp < 1$ or $sp < n$, respectively [29, Lemma 2.4 and 2.5]. Beyond that, the models involving gradients overcome some of the shortcomings of the bond-based setting. As we establish in Chapter 3 for finite-horizon fractional gradients and in Chapter 6 for general nonlocal gradients, one can use densities W that are quasiconvex in their second argument to obtain the existence of minimizers. In fact, also the densities from classical nonlinear hyperelasticity theory are admissible [30]. The connection with the classical case can be strengthened further, by considering vanishing horizon limits for the functionals in (1.35). Chapter 6 is concerned with this topic, and we prove via a rigorous Γ-limit that one recovers a local integral functional with the same density W (see Theorem 1.4.10). This shows that the state-based models with nonlocal gradients, as opposed to the bond-based models, are consistent with their local counterparts.

Nonlocal boundary conditions. To close this section, we discuss some of the different types of boundary conditions that are used for peridynamics and the related Euler-Lagrange equations. Starting with the bond-based model in (1.33), a natural condition is the Dirichlet condition $u = 0$ in the nonlocal boundary $\Omega_\delta \setminus \Omega$. Indeed, under suitable assumptions as in [33, Theorem 8.3], the minimizers of (1.33) with Dirichlet boundary conditions satisfy

$$\begin{cases} \mathcal{L}_w u = 0 & \text{a.e. in } \Omega, \\ u = 0 & \text{a.e. in } \Omega_\delta \setminus \Omega \end{cases} \qquad \text{with} \quad \mathcal{L}_w u(x) := \int_{B_\delta(x)} D_\eta w(x-y, u(x)-u(y))\,dy.$$

In essence, $\mathcal{L}_w$ is the natural finite-horizon differential operator associated to (1.33), and with the choice $w(h,\eta) = |\eta|^2/|h|^{n+2s}$, the operator $\mathcal{L}_w$ reduces to a finite-horizon equivalent of the fractional Laplacian (cf. [37]). In the absence of Dirichlet conditions, [33, Theorem 8.3] shows that the minimizers satisfy

$$\int_{\Omega_\delta \cap B_\delta(x)} D_\eta w(x-y, u(x)-u(y))\,dy = 0 \quad \text{for a.e. } x \in \Omega_\delta.$$

On Ω, the equation is simply $\mathcal{L}_w u = 0$, but on $\Omega_\delta \setminus \Omega$, the integration domain is not symmetric and the geometry of Ω plays a role in the operator. One can interpret this as a natural Neumann boundary operator. For the state-based models involving gradients in (1.35), the natural Dirichlet condition is imposed in the double layer $\Omega_{2\delta} \setminus \Omega$, and the minimizers weakly satisfy

$$\begin{cases} \operatorname{div}_\rho [D_A W(x, D_\rho u)] = 0 & \text{in } \Omega, \\ u = 0 & \text{in } \Omega_{2\delta} \setminus \Omega, \end{cases}$$

under suitable assumptions on W, cf. (1.30). The expression $\operatorname{div}_\rho[D_A W(\cdot, D_\rho u)]$ features a composition of two operators with horizon δ, which explains the need for boundary conditions in the double layer. The case of Neumann-type boundary conditions is not straightforward and is explored in Chapter 4 for finite-horizon fractional gradients, cf. (1.51). It requires a careful study of the functions with zero nonlocal gradient, which surprisingly constitute an infinite-dimensional space.

Recently, there has also been much interest in coupling the nonlocal models from peridynamics with local boundary conditions via a heterogeneous horizon function $\delta : \Omega \to [0, \infty)$, see [191,192] and also [103, 115, 201, 203]. Indeed, the idea is that the horizon $\delta(x)$ is positive inside Ω, thus modeling nonlocal interactions, but localizes near the boundary, that is, $\delta(x) \to 0$ as $x \to \partial\Omega$. In this way, one can combine the advantages of peridynamics, while still having the local boundary conditions that are the most realistic from a practical perspective. In [191, 192], this concept is treated for double-integral functionals related to the bond-based models in (1.33). The case with nonlocal gradients has not appeared in the literature, and will be tackled in a forthcoming work. One of the main difficulties with this new setting is that the nonlocal gradients with heterogeneous horizon are not Fourier multipliers anymore, and their study necessitates the use of tools from the more technical theory of pseudo-differential operators.

1.3.2 Image denoising

The second application we highlight is in the area of image processing, specifically, the removal of noise in measured images. Our set-up consists of an open and bounded domain $\Omega \subset \mathbb{R}^n$ (generically $n = 2$), and a pair $u^c, u^\eta \in L^2(\Omega)$ of clean and noisy images, respectively. Here, the aim is to reconstruct the clean image u^c, given the noisy measurement u^η. We focus on the variational regularization method, which relies on tools from the calculus of variations and is very popular and versatile. The method consists of minimizing a functional that features a fidelity term and a regularization term; the former is used to make the reconstruction fit the measurement, while the latter uses prior information on the clean image to smooth out the noise. We restrict ourselves to the case where the fidelity term is given by the L^2-distance to the noisy image, which yields the formulation:

$$\text{Minimize } \mathcal{J}(u) := \|u - u^\eta\|^2_{L^2(\Omega)} + \mathcal{R}(u) \quad \text{over } u \in L^2(\Omega),$$

with $\mathcal{R} : L^2(\Omega) \to [0, \infty]$ the regularization functional. In view of the direct method, it is sufficient for $\mathcal{R}$ to be sequentially lower semicontinuous with respect to the weak L^2-topology in order for $\mathcal{J}$ to admit minimizers; indeed, the coercivity is automatically satisfied thanks to the fidelity term. As a result, there is a lot of freedom in choosing the regularizer, while still having a well-posed problem formulation.

The most well-known regularizer is the TV-regularizer proposed by Rudin, Osher & Fatemi [186], which penalizes the total variation norm of the image. This causes the image to have fewer oscillations, but does not rule out jumps or sharp features in the reconstructed image. More complex regularization terms are also considered in the literature such as infimal-convolution total variation [60], total generalized convolution [51], and many others [20]. Moreover, in recent years there has been increased interest in using nonlocal regularizers for image processing, see e.g. [12,14,15,21,46, 53,117]. They have the advantage of allowing less regular functions in the reconstruction, and can be very useful for preserving the sharp features that most images possess. In fact, the functionals involving nonlocal gradients from Section 1.2 are also perfect for this purpose. For example, one can consider the regularizer (on the domain Ω_δ instead of Ω)

$$\mathcal{R}_s(u) : L^2(\Omega_\delta) \to [0,\infty], \quad \mathcal{R}_s(u) := \begin{cases} \displaystyle\int_\Omega f(x, D^s_\delta u)\, dx & \text{for } u \in H^{s,2,\delta}_0(\Omega), \\ \infty & \text{otherwise,} \end{cases} \tag{1.36}$$

where $f : \Omega \times \mathbb{R}^n \to [0,\infty]$ is a Carathéodory integrand with

$$\mu|A|^p \leq f(x,A) \leq C(1+|A|^p) \quad \text{for a.e. } x \in \Omega \text{ and all } A \in \mathbb{R}^n.$$

If f is convex in its second argument, then $\mathcal{R}_s$ will be weakly lower semicontinuous on $L^2(\Omega_\delta)$ (cf. [31]), which gives the existence of minimizers of the functional

$$\mathcal{J}_s : L^2(\Omega_\delta) \to [0,\infty], \quad \mathcal{J}_s(u) = \|u - u^\eta\|^2_{L^2(\Omega_\delta)} + \mathcal{R}_s(u).$$

Moreover, there is the freedom of choosing $s \in (0,1)$ and also f, allowing for more or less regular functions in the reconstruction model depending on the given situation.

Bi-level optimization. Given the great variety of possible regularizers, there has been a growing interest in finding a systematic way for designing them. One method, known as bi-level parameter learning [95], consists of optimization over a parameter dependent family of regularizers in a supervised learning scheme. One starts with a collection of noisy and clean data images and fits the regularizer to this data using a given cost functional. This results in a nested variational problem, and the existence of optimal parameters and derivation of optimality conditions are important theoretical considerations.

Precisely, consider a family of regularizers $\{\mathcal{R}_\lambda\}_\lambda : L^2(\Omega) \to [0,\infty]$, where $\lambda \in \Lambda$ ranges over a subset of a first countable topological space X. In the case of a simple L^2-cost functional, and single data point $(u^c, u^\eta) \in L^2(\Omega \times \Omega)$, the bi-level problem reads:

$$\begin{aligned} &\text{(Upper-level)} && \text{Minimize } \mathcal{I}(\lambda) := \inf_{w \in K_\lambda} \|w - u^c\|^2_{L^2(\Omega)} \text{ over } \lambda \in \Lambda, \\ &\text{(Lower-level)} && K_\lambda := \operatorname*{arg\,min}_{u \in L^2(\Omega)} \mathcal{J}_\lambda(u), \end{aligned} \tag{1.37}$$

where $\mathcal{J}_\lambda(u) := \|u - u^\eta\|^2_{L^2(\Omega)} + \mathcal{R}_\lambda(u)$ is the reconstruction functional. In this model we try to minimize the distance between the best reconstruction for a given parameter and the clean image over all possible parameters $\lambda \in \Lambda$, thus leading to the best possible regularizer within the family $\{\mathcal{R}_\lambda\}_\lambda$. The existence of minimizers to such bi-level problems usually requires Λ to be a compact set in order for $\mathcal{I}$ to be coercive; for example, box-constraints are used in [13, 25, 135], to force real-valued parameters to lie within a closed and bounded interval. However, it is not clear how to restrict the parameter values, and, depending on the given data, better parameters may lie outside the domain.

Because of this, several recent references have considered bi-level problems with open parameter sets Λ and determined a suitable extension of the bi-level problem to the closure $\overline{\Lambda}$ in order to recover the existence of optimal parameters, see e.g. [83, 84, 87, 152]. This usually involves reconstruction models at the boundary of the domain $\partial\Lambda$ where the regularizers have a completely different structure compared to the original family of regularizers. Since the analyses in [83, 84, 87, 152] are adapted to specific families of regularizers, we consider in Chapter 7 a general abstract framework for bi-level problems that is applicable to non-closed sets Λ. The main result consists of a characterization of the relaxation of $\mathcal{I}$ to the closure $\overline{\Lambda}$, which provides the most natural extension from the point of view of minimization. Moreover, we consider four different examples of families of regularizers and determine the relaxation of $\mathcal{I}$ explicitly, see Section 1.4 for more details.

While the regularizers depending on nonlocal gradients do not appear in Chapter 7, we want to close this section by showing that the family in (1.36) does fit into the general framework, enabling us to determine an optimal fractional order. To this aim, we consider the family $\{\mathcal{R}_s\}_s$ and associated reconstruction functionals $\{\mathcal{J}_s\}_s$ as in (1.36) with $s \in \Lambda := (0,1)$, and the bi-level optimization scheme

$$\begin{aligned}
&\text{(Upper-level)} \qquad \text{Minimize } \mathcal{I}(s) := \inf_{w \in K_s} \|w - u^c\|^2_{L^2(\Omega_\delta)} \text{ over } s \in (0,1),\\
&\text{(Lower-level)} \qquad K_s := \operatorname*{arg\,min}_{u \in L^2(\Omega_\delta)} \mathcal{J}_s(u).
\end{aligned}$$

Note that there is a unique minimizer of $\mathcal{J}_s$, i.e., $K_s = \{w^{(s)}\}$, since $\mathcal{J}_s$ is strictly convex as the sum of a strictly convex and convex functional. To find the relaxation of $\mathcal{I}$ according to Theorem 1.4.12, we need to determine the L^2-Mosco-limits of the family $\{\mathcal{R}_s\}_s$, which are Γ-limits with respect to the weak and strong topology in $L^2(\Omega_\delta)$ simultaneously. To phrase the result, we introduce the functionals $\mathcal{R}_0, \mathcal{R}_1 : L^2(\Omega_\delta) \to [0,\infty]$ as

$$\mathcal{R}_0(u) = \begin{cases} \displaystyle\int_\Omega f(x, D^0_\delta u)\,dx & \text{for } u \in L^2(\Omega_\delta) \text{ with } u = 0 \text{ a.e. in } \Omega_\delta \setminus \Omega_{-\delta},\\ \infty & \text{otherwise,} \end{cases}$$

and

$$\mathcal{R}_1(u) = \begin{cases} \displaystyle\int_\Omega f(x, Du)\,dx & \text{for } u \in W^{1,2}(\Omega_\delta) \text{ with } u = 0 \text{ a.e. in } \Omega_\delta \setminus \Omega_{-\delta},\\ \infty & \text{otherwise.} \end{cases}$$

Here, D^0_δ is the zero-order nonlocal gradient, and it is bounded from $L^2(\Omega_\delta)$ to $L^2(\Omega;\mathbb{R}^n)$. It now follows from Theorem 1.4.4 that the Mosco-limits are given by

$$\overline{\mathcal{R}}_s(u) := \mathrm{Mosc}(L^2)\text{–}\lim_{s' \to s} \mathcal{R}_s(u) = \begin{cases} \mathcal{R}_s & \text{for } s \in (0,1),\\ \mathcal{R}_0 & \text{for } s = 0,\\ \mathcal{R}_1 & \text{for } s = 1, \end{cases} \qquad \text{for } s \in [0,1].$$

Since $\mathcal{R}_0$ and $\mathcal{R}_1$ are also convex, (1.55) is satisfied and Theorem 1.4.12 states that the relaxed bi-level problem is given by

$$\begin{aligned}
&\text{(Upper-level)} \qquad \text{Minimize } \overline{\mathcal{I}}(s) := \inf_{w \in \overline{K_s}} \|w - u^c\|^2_{L^2(\Omega_\delta)} \text{ over } s \in [0,1],\\
&\text{(Lower-level)} \qquad \overline{K_s} := \operatorname*{arg\,min}_{u \in L^2(\Omega_\delta)} \|u - u^\eta\|^2_{L^2(\Omega_\delta)} + \overline{\mathcal{R}}_s(u).
\end{aligned}$$

By the abstract theory of relaxation, this new problem admits an optimal parameter $\bar{s} \in [0,1]$ that relates back to the minimizing parameter sequences of the original problem. The extended bi-level scheme incorporates the full range of fractional exponents from 0 to 1, interpolating between the Lebesgue space $L^2(\Omega_\delta)$ and the Sobolev space $W^{1,2}(\Omega_\delta)$. This makes it possible to tune the amount of regularity exactly as the data requires.

1.4 Contributions of the thesis

In this section we give an overview of the articles that make up the chapters of the thesis. While doing so, we present some of the main ideas behind the papers and highlight a selection of the most important results. We start off with Chapter 2, which is concerned with the relaxation of linear growth functionals depending on Riesz fractional gradients and corresponds with the published paper:

[189] H. Schönberger. Extending linear growth functionals to functions of bounded fractional variation. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 154(1):304–327, 2024. https://doi.org/10.1017/prm.2023.14.

The main objective of this paper is to extend the existence theory for minimizers of variational problems involving the Riesz fractional gradient (cf. Theorem 1.2.3) to the linear growth case $p = 1$. Explicitly, we consider the functional

$$\mathcal{F}_s(u) = \int_{\mathbb{R}^n} f(x, D^s u)\, dx \quad \text{for } u \in H^{s,1}_g(\Omega;\mathbb{R}^m), \tag{1.38}$$

where $\Omega \subset \mathbb{R}^n$ is a Lipschitz domain, $g \in H^{s,1}(\mathbb{R}^n;\mathbb{R}^m)$, and $f : \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a Carathéodory integrand that satisfies the growth bound

$$|f(x, A)| \leq M|A| + a(x) \quad \text{for all } x \in \mathbb{R}^n \text{ and } A \in \mathbb{R}^{m\times n}, \tag{1.39}$$

with $M > 0$ and $a \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$, and the coercivity bound

$$\mu|A| - c \leq f(x, A) \quad \text{for all } x \in \mathbb{R}^n \text{ and } A \in \mathbb{R}^{m\times n}, \tag{1.40}$$

with $\mu, c > 0$. Since $H^{s,1}(\mathbb{R}^n;\mathbb{R}^m)$ is not reflexive, we cannot prove coercivity of $\mathcal{F}_s$ with respect to the weak topology. However, as for classical linear growth functionals, we may weaken the topology to that of L^1-convergence, and use that bounded sequences $(u_j)_j \subset H^{s,1}_g(\Omega;\mathbb{R}^m)$ converge up to subsequence to a function $u \in BV^s_g(\Omega;\mathbb{R}^m)$ with respect to the L^1-convergence; here, the space of functions with bounded fractional variation is defined as

$$BV^s(\mathbb{R}^n;\mathbb{R}^m) := \left\{u \in L^1(\mathbb{R}^n;\mathbb{R}^m) \,:\, D^s u \in \mathcal{M}(\mathbb{R}^n;\mathbb{R}^{m\times n})\right\},$$

with $BV^s_g(\Omega;\mathbb{R}^m) = \{u \in BV^s(\mathbb{R}^n;\mathbb{R}^m) \,:\, u = g \text{ a.e. in } \Omega^c\}$, see [54,66,68] for more on these spaces. Hence, we can prove using (1.40) and the improved Poincaré inequality in Proposition 2.3.6, that the extended functional

$$\mathcal{F}_s(u) = \begin{cases} \displaystyle\int_{\mathbb{R}^n} f(x, D^s u)\, dx & \text{for } u \in H^{s,1}_g(\Omega;\mathbb{R}^m), \\ \infty & \text{for } u \in BV^s_g(\Omega;\mathbb{R}^m) \setminus H^{s,1}_g(\Omega;\mathbb{R}^m), \end{cases}$$

is coercive with respect to the L^1-convergence. However, it is not anymore lower semicontinuous with respect to this weaker convergence, so we use the concept of relaxation (cf. Definition 1.1.5), noting that we are in the setting of (A1), to recover the existence of minimizers; that is, we consider the functional

$$\mathcal{F}^{\mathrm{rel}}_s(u) := \left\{\liminf_{j\to\infty} \mathcal{F}_s(u_j) \,:\, (u_j)_j \subset H^{s,1}_g(\Omega;\mathbb{R}^m),\ u_j \to u \text{ in } L^1(\mathbb{R}^n;\mathbb{R}^m)\right\}, \tag{1.41}$$

for $u \in BV^s_g(\Omega;\mathbb{R}^m)$. This relaxation will admit minimizers and minimizing sequences of $\mathcal{F}_s$ converge up to subsequence to minimizers of $\mathcal{F}^{\mathrm{rel}}_s$, see Theorem 1.1.7. The main result of the chapter is an explicit representation of the relaxation without any convexity assumptions on the integrand f. To phrase the result, we introduce the upper recession function

$$f^{\#}(x, A) = \limsup_{\substack{(x',A')\to(x,A)\\ t\to\infty}} \frac{f(x, tA)}{t},$$

and recall the formulas for the recession function in (1.15) and the quasiconvex envelope in Definition 1.1.19. Moreover, we denote for $u \in BV^s(\mathbb{R}^n;\mathbb{R}^m)$ the decomposition

$$D^s u = \nabla^s u\, dx + D^s_* u,$$

where $\nabla^s u \in L^1(\mathbb{R}^n;\mathbb{R}^{m\times n})$ is the absolutely continuous part of $D^s u \in \mathcal{M}(\mathbb{R}^n;\mathbb{R}^{m\times n})$ with respect to the Lebesgue measure, and $D^s_* u \in \mathcal{M}(\mathbb{R}^n;\mathbb{R}^{m\times n})$ is the singular part. The main result is the following, where we remark that in the paper the relaxation is taken with respect to the weak*-convergence in $BV^s_g(\Omega;\mathbb{R}^m)$, which is equivalent to L^1-convergence given the coercivity bound (1.40).

Theorem 1.4.1 (Linear growth relaxation). *Let $s \in (0,1)$ and assume $f : \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a Carathéodory integrand that satisfies* (1.39) *and* (1.40)*, and that*

$$f^\infty(x,A) \text{ exists and } (f^{\mathrm{qc}})^\#(x,A) = \limsup_{\substack{A'\to A\\ t\to\infty}} \frac{f^{\mathrm{qc}}(x,tA')}{t} \text{ for all } (x,A) \in \overline{\Omega}\times\mathbb{R}^{m\times n}, \tag{1.42}$$

with f^{qc} the quasiconvex envelope of f with respect to its second argument. Then, the relaxation of $\mathcal{F}_s$ in (1.38) *given by* (1.41) *can be represented as*

$$\mathcal{F}_s^{\mathrm{rel}}(u) = \int_\Omega f^{\mathrm{qc}}(x,\nabla^s u)\,dx + \int_{\overline{\Omega}} (f^{\mathrm{qc}})^\# \left(x, \frac{dD^s_* u}{d|D^s_* u|}\right) d|D^s_* u| + \int_{\Omega^c} f(x,\nabla^s u)\,dx, \tag{1.43}$$

for $u \in BV^s_g(\Omega;\mathbb{R}^m)$.

Remark 1.4.2. a) In the case where f is already quasiconvex, the second condition in (1.42) is not necessary and the relaxation simplifies to

$$\mathcal{F}_s^{\mathrm{rel}}(u) = \int_{\mathbb{R}^n} f(x,\nabla^s u)\,dx + \int_{\overline{\Omega}} f^\infty\left(x, \frac{dD^s_* u}{d|D^s_* u|}\right) d|D^s_* u| \quad \text{for } u \in BV^s_g(\Omega;\mathbb{R}^m).$$

In general, the second condition in (1.42) should be seen as a continuity condition in the first argument of f^{qc} at infinity. In fact, it can be replaced by the stronger continuity condition on f itself

$$|f(x,A) - f(y,A)| \leq \omega(|x-y|)(1+|A|) \quad \text{for all } x,y \in \overline{\Omega} \text{ and } A \in \mathbb{R}^{m\times n},$$

where $\omega : [0,\infty) \to [0,\infty)$ is a continuous and increasing function with $\omega(0) = 0$, see Remark 2.5.1 c).

b) In the case of the convex area-integrand $f(x,A) = \sqrt{1+|A|^2} - 1$ (adjusted for unbounded domains) with $m = 1$, it holds that $f^\infty(x,A) = |A|$, so the relaxation becomes

$$\mathcal{F}_s^{\mathrm{rel}}(u) = \int_{\mathbb{R}^n} \sqrt{1+|\nabla^s u|^2} - 1\,dx + |D^s_* u|(\overline{\Omega}) \quad \text{for } u \in BV^s_g(\Omega).$$

The theorem provides the existence of minimizers for this extended fractional Plateau problem. △

Theorem 1.4.1 can be seen as an extension of the fractional relaxation result in Theorem 1.2.6 to the case $p = 1$, and, at the same time, as a fractional analogue of the classical linear growth relaxation in Theorem 1.1.21. Moreover, since there is no convexity assumption on f, the effect of quasiconvexification and the extension to BV^s_g are combined in the relaxation procedure. We note that similarly to the case $p \in (1,\infty)$, the quasiconvexification occurs only in $\overline{\Omega}$, which is due to the strong convergence of fractional variations in the complement of Ω, see Lemma 2.3.5. In addition, the reason why the singular part of the fractional variation is only integrated in $\overline{\Omega}$ is because of the fact that this singular part behaves in a local way, cf. Remark 2.3.4. Indeed, this makes it so that $D^s_* u = D^s_* g = 0$ in $\overline{\Omega}^c$ for all $u \in BV^s_g(\Omega;\mathbb{R}^m)$.

The proof of the relaxation result relies on an extension of the identities in (1.22) to functions in BV^s, see [68], which paves the way for comparing the fractional variation with the classical variation. This is essentially the reason for the appearance of quasiconvexity in the fractional setting. To deal with the singular part of the fractional variation that appears through concentration effects in L^1, we utilize the tool of generalized Young measures, which makes it possible to capture the oscillation and concentration behavior of sequences of measures. Finally, in order to construct a recovery sequence for the right hand side of (1.43), we carefully separate the quasiconvexification process and the extension to BV^s in order to produce different sequences for each. The construction of the latter sequence hinges on a novel density result, which allows the approximation of functions in BV^s_g by sequences in $H^{s,1}_g$ with respect to a suitable fractional notion of area-strict convergence.

Chapter 3 focuses on the models involving the finite-horizon fractional gradient in (1.25), and is based on the article:

[72] J. Cueto, C. Kreisbeck and H. Schönberger. A variational theory for integral functionals involving finite-horizon fractional gradients. *Fractional Calculus and Applied Analysis*, 26(5):2001–2056, 2023. https://doi.org/10.1007/s13540-023-00196-7.

This chapter revolves around establishing a comprehensive variational theory for integral functionals of the type (1.27) involving the finite-horizon fractional gradient and subject to a Dirichlet condition in the collar $\Gamma_{\pm\delta} := \Omega_\delta \setminus \Omega_{-\delta}$. The main conceptual idea that lies behind the results is a translation mechanism similarly to (1.22), that makes it possible to compare the finite-horizon fractional gradient with the classical gradient, and also with the Riesz fractional gradient.

To establish a connection between the finite-horizon and classical gradient, we consider an analogue to the identities in (1.22); in fact, Bellido, Cueto & Mora-Corral had already shown in [31] that the finite-horizon counterpart of the Riesz potential defined by

$$Q^s_\delta : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}, \quad Q^s_\delta(x) = c_{n,s,\delta} \int_{|x|}^{\delta} \frac{\bar{w}_\delta(t)}{t^{n+s}}\,dt$$

with $\bar{w}_\delta$ the radial representation of w_δ, enjoys the property

$$D^s_\delta \varphi = D(Q^s_\delta * \varphi) \quad \text{for all } \varphi \in C^\infty(\mathbb{R}^n). \tag{1.44}$$

This mirrors the first identity in (1.22) and is actually more convenient since Q^s_δ is integrable as opposed to the Riesz potential I_{1-s}. In order to go the other way, we heuristically invert the convolution with Q^s_δ in Fourier space, and define the operator

$$\mathcal{P}^s_\delta \varphi = \left(\frac{\widehat{\varphi}}{\widehat{Q^s_\delta}}\right)^{\vee}.$$

We prove that the convolution with Q^s_δ and the operator $\mathcal{P}^s_\delta$ can be extended in a bounded way to the Sobolev spaces $W^{1,p}(\mathbb{R}^n)$ and $H^{s,p,\delta}(\mathbb{R}^n)$ such that they constitute each other's inverse. As a consequence, there is a isomorphism between $W^{1,p}(\mathbb{R}^n)$ and $H^{s,p,\delta}(\mathbb{R}^n)$ and it holds that

$$D^s_\delta u = D(Q^s_\delta * u) \quad \text{and} \quad Dv = D^s_\delta \mathcal{P}^s_\delta v, \tag{1.45}$$

for all $u \in H^{s,p,\delta}(\mathbb{R}^n)$ and $v \in W^{1,p}(\mathbb{R}^n)$ with $p \in [1,\infty]$.

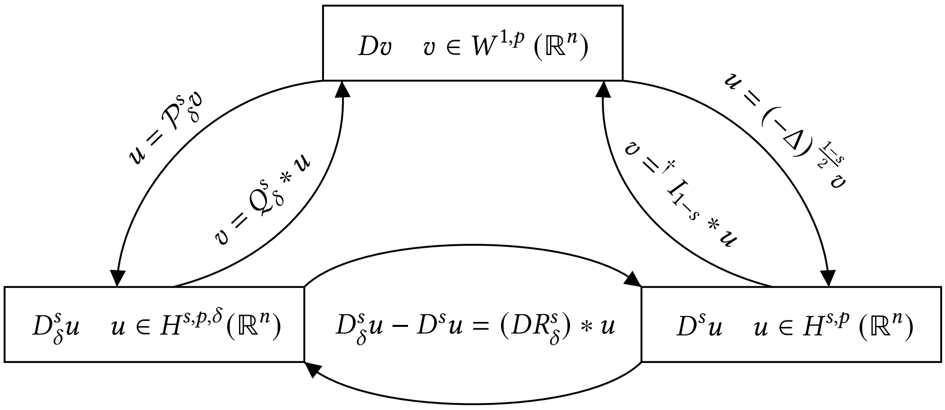

Figure 1.3: Illustration of the relations between classical, fractional, and finite-horizon fractional gradients. † When $I_{1-s} * u$ is well-defined.

We also establish a connection between the finite-horizon fractional gradient and Riesz fractional gradient. Indeed, if we define $R^s_\delta = Q^s_\delta - I_{1-s}$, then it can be shown that $DR^s_\delta \in L^1(\mathbb{R}^n; \mathbb{R}^n) \cap C^\infty(\mathbb{R}^n; \mathbb{R}^n)$. Moreover, in light of the first two identities in (1.22) and (1.45), it holds that

$$D^s_\delta \varphi = D^s \varphi + (DR^s_\delta) * \varphi \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^n). \tag{1.46}$$

We use this identity to prove that $H^{s,p}(\mathbb{R}^n) = H^{s,p,\delta}(\mathbb{R}^n)$ with equivalent norms, after which the identity in (1.46) can be extended to $u \in H^{s,p}(\mathbb{R}^n) = H^{s,p,\delta}(\mathbb{R}^n)$. A visual summary of all the connections is given in Figure 1.3.

The full translation mechanism enables us to prove results regarding the variational problems involving the finite-horizon fractional by relating it to known results in the classical case. There are three main results in the chapter, of which the first is a complete characterization of the lower semicontinuity of functionals as in (1.27).

Theorem 1.4.3 (Characterization of weak lower semicontinuity). *Let $s \in (0,1)$, $p \in (1,\infty)$, $\Omega \subset \mathbb{R}^n$ be open and bounded with $|\partial\Omega_{-\delta}| = 0$ and $g \in H^{s,p,\delta}(\Omega;\mathbb{R}^m)$. Further, let $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory function satisfying*

$$-C(1 + |z|^p + |A|^q) \le f(x,z,A) \le C(1 + |z|^p + |A|^p)$$

for a.e. $x \in \Omega$ and all $(z,A) \in \mathbb{R}^m \times \mathbb{R}^{m\times n}$ with $C > 0$ and $q \in [1,p)$. Then, $\mathcal{F}^s_\delta$ from (1.27) is sequentially weakly lower semicontinuous on $H^{s,p,\delta}_g(\Omega;\mathbb{R}^m)$ if and only if $A \mapsto f(x,z,A)$ is quasiconvex for a.e. $x \in \Omega_{-\delta}$ and all $z \in \mathbb{R}^m$.

As in the classical and fractional case, the notion of quasiconvexity is again crucial in this setting. In fact, the proof is also very reminiscent of the one of Theorem 1.2.1 except for using the new translation mechanism of (1.45). In addition, quasiconvexity is not required in the collar $\Omega \setminus \Omega_{-\delta}$, because of the strong convergence of the nonlocal gradients there, cf. Lemma 3.2.12.

The second main result pushes the translation method one step further, and states that any Γ-convergence result that holds for a sequence of integral functionals involving classical derivatives can be carried over to the setting of the finite-horizon gradients, see Theorem 3.5.1 for the precise statement. As two specific examples, we show that we can obtain a homogenization result and relaxation formula in the finite-horizon case.

The chapter finishes with a study of the dependence of the variational problems on the fractional parameter s. To phrase the result, we introduce the functionals $\mathcal{F}^s_\delta : L^p(\Omega_\delta;\mathbb{R}^m) \to \mathbb{R}_\infty$ for $s \in$

(0, 1) as

$$\mathcal{F}^s_\delta(u) = \begin{cases} \displaystyle\int_\Omega f(x, D^s_\delta u(x))\,dx & \text{for } u \in H^{s,p,\delta}_0(\Omega;\mathbb{R}^m), \\ \infty & \text{otherwise,} \end{cases} \tag{1.47}$$

where $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a Carathéodory integrand satisfying for $\mu, C > 0$

$$\mu|A|^p - C \le f(x,A) \le C(1 + |A|^p) \qquad \text{for a.e. } x \in \Omega \text{ and all } A \in \mathbb{R}^{m\times n}.$$

Additionally, we define $\mathcal{F}^0_\delta, \mathcal{F}^1_\delta : L^p(\Omega_\delta;\mathbb{R}^m) \to \mathbb{R}_\infty$ as

$$\mathcal{F}^0_\delta(u) = \begin{cases} \displaystyle\int_\Omega f(x, D^0_\delta u(x))\,dx & \text{for } u \in L^p(\Omega_\delta;\mathbb{R}^m) \text{ with } u = 0 \text{ a.e. in } \Omega_\delta \setminus \Omega_{-\delta}, \\ \infty & \text{otherwise,} \end{cases} \tag{1.48}$$

and

$$\mathcal{F}^1_\delta(u) = \begin{cases} \displaystyle\int_\Omega f(x, Du(x))\,dx & \text{for } u \in W^{1,p}(\Omega_\delta;\mathbb{R}^m) \text{ with } u = 0 \text{ a.e. in } \Omega_\delta \setminus \Omega_{-\delta}, \\ \infty & \text{otherwise.} \end{cases} \tag{1.49}$$

We note that the zero order gradient D^0_δ is defined by simply plugging in $s = 0$ in the definition, and it can be extended to a bounded operator on L^p for $p \in (1,\infty)$. The next result establishes the continuous dependence on the fractional parameter s in the framework of Γ-convergence, where we assume that $\Omega_{-\delta}$ is a Lipschitz domain and $p \in (1,\infty)$.

Theorem 1.4.4 (Γ-limits for $s_j \to s \in [0,1]$). *Let $\mathcal{F}^s_\delta$ be as in* (1.47), (1.48) *and* (1.49) *for $s \in [0,1]$ and suppose that $f(x,\cdot)$ is quasiconvex for a.e. $x \in \Omega_{-\delta}$. If $(s_j)_j \subset [0,1]$ satisfies $s_j \to s \in [0,1]$, then $(\mathcal{F}_{s_j,\delta})_j$ Γ-converges to $\mathcal{F}^s_\delta$ with respect to the weak and strong topology in $L^p(\Omega_\delta;\mathbb{R}^m)$, that is,*

$$\Gamma(L^p)\text{-}\lim_{j\to\infty} \mathcal{F}_{s_j,\delta} = \mathcal{F}^s_\delta = \Gamma(w\text{-}L^p)\text{-}\lim_{j\to\infty} \mathcal{F}_{s_j,\delta}.$$

Moreover, $(\mathcal{F}_{s_j,\delta})_j$ is equi-coercive with respect to the strong topology in $L^p(\Omega_\delta;\mathbb{R}^m)$ if $\inf_j s_j > 0$ and the weak topology in $L^p(\Omega_\delta;\mathbb{R}^m)$ otherwise.

Given the Γ-convergence and equi-coercivity of Theorem 1.4.4, it follows that the minimizers of $\mathcal{F}_{s_j,\delta}$ converge up to subsequence to minimizers of $\mathcal{F}^s_\delta$, see Theorem 1.1.11. Beyond the translation mechanism and localization of the nonlocal gradient $D^s_\delta \to D$ as $s \uparrow 1$, the proof of Theorem 1.4.4 relies crucially on an improved version of the Poincaré inequality in (1.28) where the constant $C > 0$ does not depend on $s \in [0,1)$. To establish it, we derive technical estimates on the Fourier transform of Q^s_δ uniformly in s in order to apply the Mihlin-Hörmander multiplier theorem.

We continue the study of the finite-horizon fractional gradients in Chapter 4 and it is based on the preprint:

[141] C. Kreisbeck and H. Schönberger. Non-constant functions with zero nonlocal gradients and their role in nonlocal Neumann-type problems. Preprint arXiv:2402.11308, 2024.

The paper is concerned with a natural question, that turns out to have a surprising answer with applications to Neumann-type problems involving the finite-horizon fractional gradients. Namely, we consider whether the property that for $v \in W^{1,p}(\Omega)$

$$\nabla v = 0 \text{ a.e. on } \Omega \quad \text{if and only if} \quad v \text{ is constant a.e. on } \Omega,$$

carries over when the gradient is replaced by a nonlocal gradient. In the case $\Omega = \mathbb{R}^n$ and with the Riesz fractional gradient or finite-horizon fractional gradient, this is true since the Fourier symbols of D^s and D^s_δ are positive on $\mathbb{R}^n \setminus \{0\}$; in particular, this also holds for the complementary-value spaces $H_0^{s,p,\delta}(\Omega)$, as its elements can be extended to $\mathbb{R}^n$ as zero. However, when one considers $H^{s,p,\delta}(\Omega)$ on a bounded domain Ω (cf. (1.26)), the situation changes completely. Precisely, we consider the subset of $H^{s,p,\delta}(\Omega)$ given by

$$N^{s,p,\delta}(\Omega) := \{h \in H^{s,p,\delta}(\Omega) \,:\, D^s_\delta h = 0 \text{ a.e. in } \Omega\},$$

and show that it constitutes an infinite-dimensional vector space (Proposition 4.3.3), and hence, contains more than just constant functions. Since any $u \in H^{s,p,\delta}(\Omega)$ is defined on Ω_δ and $D^s_\delta u$ on the smaller set Ω, one might think that $D^s_\delta u = 0$ on Ω at least implies that u is constant on Ω or $\Omega_{-\delta}$. Even this is not true, since for any subset of Ω, there is a $h \in N^{s,p,\delta}(\Omega)$ that is non-constant on this subset (Proposition 4.3.1).

These results motivate the further study of the set $N^{s,p,\delta}(\Omega)$ and its applications, which is taken up in Chapter 4. The first is a characterization of the space $N^{s,p,\delta}(\Omega)$ in terms of the values in the single collar $\Gamma_\delta := \Omega_\delta \setminus \overline{\Omega}$ and a mean-value condition. It is based on the observation that all $h \in N^{s,p,\delta}(\Omega)$ satisfy

$$Q^s_\delta * h = c \text{ a.e. in } \Omega \quad \text{and} \quad h = g \text{ a.e. in } \Gamma_\delta \tag{1.50}$$

for some $c \in \mathbb{R}$ and $g \in L^p(\Gamma_\delta)$ in light of (1.45). Utilizing the translation mechanism from the previous chapter, we can reformulate the convolution equation (1.50) into a Dirichlet problem involving $\mathcal{P}^s_\delta$. This reformulation enables us to exploit the recent existence, uniqueness and regularity theory for pseudo-differential equations by Abels and Grubb [2, 125], given that $\mathcal{P}^s_\delta$ fits into their setting as a perturbation of the fractional Laplacian. This leads to the following main result of the paper, characterizing the set $N^{s,p,\delta}(\Omega)$.

Theorem 1.4.5 (Characterization of $N^{s,p,\delta}(\Omega)$). *Let $\Omega \subset \mathbb{R}^n$ be a bounded $C^{1,1}$-domain, then, the following holds:*

(*i*) *If $p \in (1, \frac{2}{1-s})$, then for each $c \in \mathbb{R}$ and $g \in L^p(\Gamma_\delta)$, there exists a unique solution $h \in N^{s,p,\delta}(\Omega)$ to* (1.50). *In particular, the following map is bijective*

$$\Phi^s_\delta : N^{s,p,\delta}(\Omega) \to \mathbb{R} \times L^p(\Gamma_\delta) \quad h \mapsto \left(\int_\Omega Q^s_\delta * h \, dx, h|_{\Gamma_\delta} \right).$$

(*ii*) *If $p \in [\frac{2}{1-s}, \infty)$, then for each $c \in \mathbb{R}$ and $g \in L^p(\Gamma_\delta)$, there exists at most one solution $h \in N^{s,p,\delta}(\Omega)$ to* (1.50). *In particular, the map Φ^s_δ is only injective.*

The result shows that the set $N^{s,p,\delta}(\Omega)$ is isomorphic to $\mathbb{R} \times L^p(\Gamma_\delta)$ when $p \in (1, \frac{2}{1-s})$, whereas for all $p \in (1, \infty)$, a function $h \in N^{s,p,\delta}(\Omega)$ is uniquely determined by its values in Γ_δ and the quantity $\int_\Omega Q^s_\delta * h \, dx$. The reason for the dichotomy between the values of p below and above the critical value $\frac{2}{1-s}$, is due to the regularity of solutions of pseudo-differential Dirichlet problems. Indeed, these solutions decay towards the boundary with a rate that does not depend on p, and hence, their global regularity is limited to a certain range of Bessel potential spaces. This actually shows that Φ^s_δ will not be surjective, when $p \in [\frac{2}{1-s}, \infty)$ (Remark 4.3.13). Numerical approximations also indicate that the functions in $N^{s,p,\delta}(\Omega)$ generally display singularities across the boundary $\partial\Omega$, which means that they do not lie in $L^p(\Omega_\delta)$ for large p, see Figure 1.4.

Beyond the abstract interest of $N^{s,p,\delta}(\Omega)$, it turns out that this set plays an important role for the properties of the Sobolev spaces $H^{s,p,\delta}(\Omega)$. Indeed, the translation mechanism from Chapter 3 can be adapted to bounded domains after taking the quotient with $N^{s,p,\delta}(\Omega)$, i.e., there is

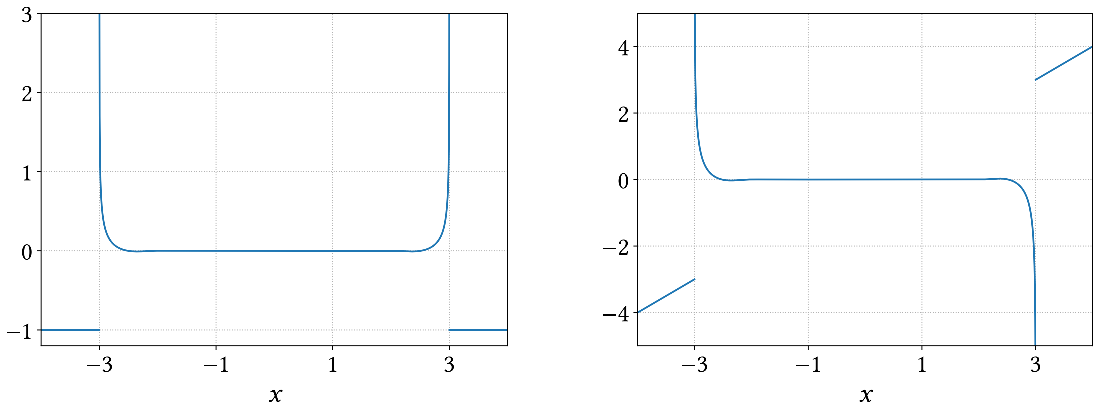

Figure 1.4: A numerical approximation of the unique solutions to (1.50) with $c = 0$ and $g(x) = -1$ and $g(x) = x$ for $x \in \Gamma_\delta$, respectively. The parameters for the computation are $n = 1$, $\Omega = (-3, 3)$, $s = \frac{1}{2}$, $\delta = 1$ and $w_\delta \in C_c^\infty(-1, 1)$ is a bump function equal to 1 on $(-\frac{1}{2}, \frac{1}{2})$.

an isomorphism between $H^{s,p,\delta}(\Omega)/N^{s,p,\delta}(\Omega)$ and $W^{1,p}(\Omega)/C$ with C the set of constant functions on Ω. Furthermore, it is possible to obtain extension operators, Poincaré inequalities and compact embeddings after suitably removing functions in $N^{s,p,\delta}(\Omega)$. In this way, we are able to carry over the results from the complementary-value spaces to the spaces $H^{s,p,\delta}(\Omega)$ without boundary conditions. Furthermore, the set $N^{s,p,\delta}(\Omega)$, which formally corresponds to the solution set of the inclusion problem $D^s_\delta u \in \{0\}$ a.e. in Ω, enables an elegant presentation of the solution theory for more general differential inclusion problems involving the gradient D^s_δ.

The main application of the new tools on the spaces $H^{s,p,\delta}(\Omega)$ is an analysis of variational problems involving the finite-horizon nonlocal gradient with Neumann-type boundary conditions. We consider the weakly closed subset of $H^{s,p,\delta}(\Omega;\mathbb{R}^m)$

$$N^{s,p,\delta}(\Omega;\mathbb{R}^m)^\perp := \left\{ u \in H^{s,p,\delta}(\Omega;\mathbb{R}^m) \ : \ \|u\|_{L^p(\Omega_\delta;\mathbb{R}^m)} = \min_{h \in N^{s,p,\delta}(\Omega;\mathbb{R}^m)} \|u - h\|_{L^p(\Omega_\delta;\mathbb{R}^m)} \right\},$$

which, when $p = 2$, agrees with the L^2-orthogonal complement of $N^{s,2,\delta}(\Omega;\mathbb{R}^m)$. This set plays the role of the functions with zero mean-value for the variational formulation of the classical Neumann problem. We obtain the following existence result.

Theorem 1.4.6 (Existence for Neumann-type problem). *Let $p \in (1,\infty)$, $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain, $F \in L^{p'}(\Omega_\delta;\mathbb{R}^m)$, and $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}_\infty$ a Carathéodory integrand such that*

$$f(x, A) \geq \mu|A|^p - C \quad \text{for a.e. } x \in \Omega \text{ and all } A \in \mathbb{R}^{m\times n}.$$

If $v \mapsto \int_\Omega f(x, Dv)\,dx$ is weakly lower semicontinuous on $W^{1,p}(\Omega;\mathbb{R}^m)$, then there exists a minimizer of

$$\mathcal{F}^s_\delta(u) := \int_\Omega f(x, D^s_\delta u)\,dx - \int_{\Omega_\delta} F \cdot u\,dx \quad \text{over all } u \in N^{s,p,\delta}(\Omega;\mathbb{R}^m)^\perp.$$

The proof relies on the new Poincaré inequality on $N^{s,p,\delta}(\Omega;\mathbb{R}^m)^\perp$ and uses the translation mechanism to exploit the weak lower semicontinuity of the classical integral functional $v \mapsto \int_\Omega f(x, Dv)\,dx$. The result is therefore valid for quasiconvex integrands with p-growth from above or polyconvex integrands without an upper bound, see Theorem 1.1.14 and Remark 1.1.15.

If the functional $\mathcal{F}^s_\delta$ is invariant under translation in $N^{s,p,\delta}(\Omega;\mathbb{R}^m)$, that is, F satisfies the nonlocal compatibility condition

$$\int_{\Omega_\delta} F\cdot h\,dx = 0 \quad \text{for all } h\in N^{s,p,\delta}(\Omega;\mathbb{R}^m),$$

then any minimizer u of $\mathcal{F}^s_\delta$ over the set $N^{s,p,\delta}(\Omega;\mathbb{R}^m)^\perp$ is also a minimizer over the full space $H^{s,p,\delta}(\Omega;\mathbb{R}^m)$. Under this assumption, and if f satisfies the conditions of Proposition 1.1.17, then the minimizer weakly satisfies

$$\begin{cases} -\operatorname{div}^s_\delta[D_A f(x,D^s_\delta u)] = F & \text{in } \Omega_{-\delta},\\ \mathcal{N}^s_\delta[D_A f(x,D^s_\delta u)] = F & \text{in } \Gamma_{\pm\delta}, \end{cases} \tag{1.51}$$

with $\mathcal{N}^s_\delta := -\operatorname{div}^s_\delta(\mathbb{1}_\Omega\,\cdot)$ the nonlocal Neumann operator associated to D^s_δ; here, $\mathbb{1}_\Omega$ is the indicator function of Ω. The Neumann operator already appeared in [26] for a concise formulation of nonlocal Green's identities related to D^s_δ. As $s\uparrow 1$, we expect the nonlocal Neumann operator $\mathcal{N}^s_\delta$ to localize on the boundary. This is made rigorous via a Γ-limit (Theorem 4.6.4), which shows that if the classical compatibility condition holds $\int_\Omega F\,dx = 0$, the minimizers of $\mathcal{F}^s_\delta$ converge up to subsequence in $L^p(\Omega;\mathbb{R}^m)$ to a minimizer of

$$u\mapsto \int_\Omega f(x,Du)\,dx - \int_\Omega F\cdot u\,dx \quad \text{over all } u\in W^{1,p}(\Omega;\mathbb{R}^m).$$

Since this functional is invariant under translation by constants, Remark 1.1.18 b) shows that its minimizers weakly satisfy

$$\begin{cases} -\operatorname{div}[D_A f(x,Du)] = F & \text{in } \Omega,\\ D_A f(x,Du)\cdot\nu = 0 & \text{on } \partial\Omega, \end{cases}$$

with ν an outward unit normal to $\partial\Omega$. This establishes the consistency of the new nonlocal Neumann-type problem in (1.51) with its classical counterpart.

With the successful treatment of finite-horizon fractional gradients, a natural next step is to investigate nonlocal gradients with more general kernels, which is addressed in Chapter 5:

[36] J. C. Bellido, C. Mora-Corral and H. Schönberger. Nonlocal gradients: Fundamental theorem of calculus, Poincaré inequalities and embeddings. Preprint arXiv:2402.16487, 2024.

We consider the gradient D_ρ as in (1.31) for a non-negative radial kernel $\rho:\mathbb{R}^n\setminus\{0\}\to[0,\infty)$ that satisfies (1.32). Such gradients have been considered before (cf. [16, 92, 93, 102, 156, 161]), but we present the first extension of the functional analytic tools that hold for D^s and D^s_δ to the more general setting involving D_ρ. Precisely, we introduce for $p\in[1,\infty]$

$$H^{\rho,p}(\mathbb{R}^n) := \{u\in L^p(\mathbb{R}^n) \,:\, D_\rho u\in L^p(\mathbb{R}^n;\mathbb{R}^n)\},$$

where the weak nonlocal gradient is defined via integration by parts. Moreover, for $\Omega\subset\mathbb{R}^n$ open, the complementary-value space is defined as

$$H^{\rho,p}_0(\Omega) := \{u\in H^{\rho,p}(\mathbb{R}^n) \,:\, u=0 \text{ a.e. in } \Omega^c\}.$$

Without additional assumptions on ρ, we prove nonlocal Leibniz rules, density results and the equivalence between spaces associated to kernels with the same behavior around the origin; in particular, we can often work, without loss of generality, with compactly supported kernels ρ. Moreover, using the locally integrable function

$$Q_\rho : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}, \quad Q_\rho(x) := \int_{|x|}^{\infty} \frac{\overline{\rho}(t)}{t}\, dt,$$

where $\overline{\rho}$ is the radial representation of ρ, an analogue to the translation method in (1.44) is proven, that is,

$$D_\rho \varphi = D(Q_\rho * \varphi) = Q_\rho * D\varphi \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^n). \tag{1.52}$$

Beyond the connection that (1.52) provides with the classical gradient, it can be used to deduce properties of D_ρ by studying the Fourier transform of Q_ρ. Indeed, we can prove estimates on $\widehat{Q}_\rho$ and its derivatives, under the following assumptions on ρ:

(H1) The function $f_\rho : (0,\infty) \to [0,\infty)$, $t \mapsto t^{n-2}\overline{\rho}(t)$ is decreasing, and there is a $\nu > 0$ such that $t^\nu f_\rho(t)$ is decreasing on $(0,\varepsilon)$;

(H2) the function f_ρ is smooth on $(0,\infty)$ and for $t \in (0,\varepsilon)$

$$\left|\frac{d^k}{dr^k} f_\rho(r)\right| \leq C_k \frac{f_\rho(r)}{r^k} \quad \text{for } k \in \mathbb{N}.$$

This enables us to prove the first main result regarding Poincaré inequalities and compact embeddings for the spaces $H_0^{\rho,p}(\Omega)$.

Theorem 1.4.7 (Poincaré inequality and compact embedding). *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, $p \in (1,\infty)$, and ρ have compact support and satisfy (H1) and, if $p \neq 2$, also (H2). Then, the following holds:*

(i) If $\liminf_{t\downarrow 0} t^{n-1}\overline{\rho}(t) > 0$, there is a $C > 0$ such that

$$\|u\|_{L^p(\Omega)} \leq C\|D_\rho u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \quad \textit{for all } u \in H_0^{\rho,p}(\Omega).$$

(ii) If $\lim_{t\downarrow 0} t^{n-1}\overline{\rho}(t) = \infty$, then $H_0^{\rho,p}(\Omega)$ is compactly embedded into $L^p(\mathbb{R}^n)$.

The proof relies on the Fourier bounds of $\widehat{Q}_\rho$ in order to invert D_ρ and show that this is a bounded or compact operation, respectively. For $p = 2$, this can be achieved with Parseval's identity, whereas the general case uses the Mihlin-Hörmander multiplier theorem and requires the bounds on the derivatives of $\widehat{Q}_\rho$ that follow from (H2). The conditions in (i) and (ii) state that ρ needs to have a singularity that is as strong or stronger than the one from the zero order fractional gradient. This is quite natural, and we show in Proposition 5.7.5 that these conditions are also essentially optimal to obtain Poincaré inequalities and compact embeddings, respectively.

The inversion of D_ρ can be strengthened in the form of a fundamental theorem of calculus. For this, we need the additional assumption that ρ is in between two fractional kernels of orders $0 < \sigma \leq \gamma < 1$:

(H3) The function $r \mapsto r^{n+\sigma-1}\overline{\rho}(r)$ is almost decreasing on $(0,\varepsilon)$;

(H4) the function $r \mapsto r^{n+\gamma-1}\overline{\rho}(t)$ is almost increasing on $(0,\varepsilon)$.

This allows us to prove the second main result, which generalizes the fundamental theorem of calculus for finite-horizon fractional gradients in (1.29).

Theorem 1.4.8 (Fundamental theorem of calculus). *Let ρ satisfy (H1)-(H4) and have compact support. Then, there exists a vector-radial function $V_\rho \in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n \setminus \{0\};\mathbb{R}^n)$ such that*

$$\varphi(x) = \int_{\mathbb{R}^n} D_\rho\varphi(y) \cdot V_\rho(x-y)\,dy \quad \text{for all } x \in \mathbb{R}^n \text{ and } \varphi \in C_c^\infty(\mathbb{R}^n).$$

Moreover, there exists a $C > 0$ such that

$$|V_\rho(x)| + |x||\nabla V_\rho(x)| \leq \frac{C}{|x|^{2n-1}\rho(x)} \quad \text{for all } x \in B_\varepsilon(0) \setminus \{0\}.$$

We give here a few examples of kernels that fit into the framework. The first example corresponds with the finite-horizon fractional gradient, but allows for a more general class of cut-off functions.

Example 1.4.9. In the following, $w \in C_c^\infty(\mathbb{R}^n)$ is a non-negative radial function with $w(0) > 0$.

a) If $s \in (0,1)$ and $w/|\cdot|^{1+s}$ is radially decreasing, then

$$\rho(x) = \frac{w(x)}{|x|^{n+s-1}} \quad x \in \mathbb{R}^n \setminus \{0\},$$

satisfies the assumptions (H1)-(H4) with $\sigma = \gamma = s$. In fact, it holds that $H^{\rho,p}(\mathbb{R}^n) = H^{s,p}(\mathbb{R}^n)$ with equivalent norms.

b) Let $s \in (0,1)$, $\kappa \in \{-1,1\}$ and assume that $\mathrm{supp}(w) \subset \overline{B_1(0)}$ with $w\log(1/|\cdot|)^\kappa/|\cdot|^{1+s}$ radially decreasing. Then,

$$\rho(x) = \frac{w(x)\log(1/|x|)^\kappa}{|x|^{n+s-1}} \quad x \in \mathbb{R}^n \setminus \{0\},$$

satisfies the assumptions (H1)-(H4) for $\sigma = s$ and any $\gamma \in (s,1)$ if $\kappa = 1$, or for $\gamma = s$ and any $\sigma \in (0,s)$ if $\kappa = -1$.

c) Given a smooth function $s : [0,\infty) \to (0,1)$ such that $w/|\cdot|^{1+s(|\cdot|)}$ is radially decreasing, the kernel

$$\rho(x) = \frac{w(x)}{|x|^{n+s(|x|)-1}} \quad x \in \mathbb{R}^n \setminus \{0\},$$

satisfies (H1)-(H4) with $\sigma = \min_{[0,\varepsilon]} s$ and $\gamma = \max_{[0,\varepsilon]} s$ for any $\varepsilon > 0$.

The fundamental theorem of calculus and the bounds on V_ρ around the origin make it possible to prove sharp embeddings of the spaces $H_0^{\rho,p}(\Omega)$ into Orlicz spaces and spaces with prescribed modulus of continuity, see e.g., [180] for more details on Orlicz spaces. These embeddings generalize and refine the fractional Sobolev and Morrey inequalities that hold for D^s [193] and D^s_δ [30], and are not restricted to the scale of Lebesgue or Hölder spaces. With ρ as in Example 1.4.9 a), we recover the embedding into $L^q(\mathbb{R}^n)$ with $q = np/(n-sp)$ when $sp < n$, and the embedding into $C^{0,s-n/p}(\mathbb{R}^n)$ if $sp > n$. Moreover, with a logarithmic kernel as in Example 1.4.9 b) with $\kappa = 1$, we find for $sp < n$ an embedding into the Orlicz space with a Young function that behaves like $t^q\log(t)^q$ for large t with $q = np/(n-sp)$, whereas for $sp > n$, the functions in $H_0^{\rho,p}(\Omega)$ have a modulus of continuity given by $\omega(t) = t^{s-n/p}/\log(1/t)$ for small t.

We build upon these results in Chapter 6 by considering the varying horizon limits of general nonlocal gradients, and the chapter agrees with the preprint:

- [73] J. Cueto, C. Kreisbeck and H. Schönberger. Γ-convergence involving nonlocal gradients with varying horizon: Recovery of local and fractional models. Preprint arXiv:2404.18509, 2024.

Precisely, we consider a kernel $\rho = \rho_1$ as in the previous chapter that satisfies (H1)-(H4) and is normalized to

$$\operatorname{supp}\rho = \overline{B_1(0)} \quad \text{and} \quad \int_{\mathbb{R}^n} \rho\, dx = n.$$

This kernel induces a nonlocal gradient with horizon equal to 1, and we scale it for $\delta > 0$ to obtain kernels

$$\rho_\delta(x) := c_\delta \rho\left(\frac{x}{\delta}\right) \quad x \in \mathbb{R}^n \setminus \{0\},$$

where $(c_\delta)_\delta \subset (0,\infty)$ is a suitable sequence of scaling constants; the kernels ρ_δ define nonlocal gradients with horizon equal to δ. We are interested in the asymptotic behavior of these gradients and their associated energy functionals as $\delta \to 0$ and $\delta \to \infty$ in the setting of Γ-convergence. For $\delta \to 0$, we show the convergence to local models, while for $\delta \to \infty$, we surprisingly only recover models based on the Riesz fractional gradient.

The functionals $\mathcal{F}_\delta : L^p(\mathbb{R}^n;\mathbb{R}^m) \to \mathbb{R}_\infty$ that we consider are of the form

$$\mathcal{F}_\delta(u) = \begin{cases} \displaystyle\int_{\Omega_\delta} f(x, D_{\rho_\delta} u)\, dx & \text{for } u \in H_0^{\rho,p,\delta}(\Omega;\mathbb{R}^m), \\ \infty & \text{else,} \end{cases}$$

with $p \in (1,\infty)$, $\Omega \subset \mathbb{R}^n$ a bounded Lipschitz domain, $H_0^{\rho,p,\delta}(\Omega;\mathbb{R}^m) := H_0^{\rho_\delta,p}(\Omega;\mathbb{R}^m)$, and $f : \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$ a Carathéodory integrand that satisfies

$$\mu|A|^p \le f(x,A) \le C(a(x) + |A|^p) \quad \text{for a.e. } x \in \mathbb{R}^n \text{ and all } A \in \mathbb{R}^{m\times n},$$

with $C, \mu > 0$ and $a \in L^1(\mathbb{R}^n)$. Moreover, to obtain weak lower semicontinuity, we assume that $A \mapsto f(x,A)$ is quasiconvex for a.e. $x \in \Omega$. Combining the Poincaré inequality from Theorem 1.4.7 with the translation mechanism in (1.52), we find using similar methods as in Chapter 3 that $\mathcal{F}_\delta$ admit a minimizer for any $\delta > 0$. This shows the well-posedness of these models involving finite-horizon nonlocal gradients, which fit into the theory of state-based peridynamics, cf. Section 1.3.1.

For the asymptotics of the vanishing horizon limit, we consider the scaling regime $c_\delta := \delta^{-n}$, which preserves the normalization $\int_{\mathbb{R}^n} \rho_\delta\, dx = n$, and prove that the nonlocal gradient D_{ρ_δ} converges to the classical gradient as $\delta \to 0$. In fact, for smooth function the optimal convergence rate of δ^2 is identified. Moreover, using that

$$\widehat{Q}_{\rho_\delta}(\xi) = \widehat{Q}_\rho(\delta\xi) \quad \text{for } \xi \in \mathbb{R}^n,$$

we are able to use the bounds on $\widehat{Q}_\rho$ from Chapter 5 to prove compactness results for D_{ρ_δ} that hold uniformly in $\delta \in (0,1]$. This is needed to prove the equi-coercivity of the functionals $(\mathcal{F}_\delta)_\delta$ and leads to the following result.

Theorem 1.4.10 (Localization via vanishing horizon). *Let $c_\delta := \delta^{-n}$ for $\delta \in (0,1]$, then the functionals $(\mathcal{F}_\delta)_\delta$ Γ-converge as $\delta \to 0$ with respect to the strong $L^p(\mathbb{R}^n;\mathbb{R}^m)$-topology to the functional $\mathcal{F} : L^p(\mathbb{R}^n;\mathbb{R}^m) \to \mathbb{R}$ given by*

$$\mathcal{F}(u) = \begin{cases} \displaystyle\int_{\Omega} f(x, Du)\, dx & \text{for } u \in W_0^{1,p}(\Omega;\mathbb{R}^m), \\ \infty & \text{else,} \end{cases}$$

where the functions in $W_0^{1,p}(\Omega;\mathbb{R}^m)$ are extended to $\mathbb{R}^n$ as zero. Moreover, the family $(\mathcal{F}_\delta)_\delta$ is equi-coercive.

The previous result is in the setting of (A1), so from Theorem 1.1.11 we deduce that the minimizers of $\mathcal{F}_\delta$ converge up to subsequence in $L^p(\mathbb{R}^n;\mathbb{R}^m)$ as $\delta \to 0$ to a minimizer of $\mathcal{F}$. This result establishes that the state-based peridynamic models involving nonlocal gradients are compatible with the local models involving quasiconvex integrands through a vanishing horizon limit. This is not the case for the related case of bond-based peridynamics, cf. Section 1.3.1.

For the limit $\delta \to \infty$, we consider the scaling $c_\delta := \overline{\rho}(1/\delta)^{-1}$ for $\delta \in (1/\varepsilon, \infty)$ and, additionally, assume that the pointwise limit

$$\rho_\infty(x) := \lim_{\delta\to\infty} \rho_\delta(x) = \lim_{\delta\to\infty} \overline{\rho}(1/\delta)^{-1}\rho(x/\delta),$$

exists for $x \in \mathbb{R}^n \setminus \{0\}$. We note that this scaling leads to $\rho_\infty(x) = 1$ for $|x| = 1$, which is up to a constant the only relevant scaling. It turns out that whenever the limit exists, ρ_∞ must be a fractional kernel, i.e., $\rho_\infty = |\cdot|^{n+s_\infty-1}$ for $s_\infty \in (0,1)$, irrespective of whether ρ behaves like a fractional kernel at the origin. This is because the kernel ρ_∞ picks up multiplicativity through the isotropic scaling process, and, hence, must be a power function. For illustration, the choices ρ from Example 1.4.9 a) and b) recover $s_\infty = s$, while c) leads to the fractional kernel of order $s_\infty = s(0)$. Similarly to the vanishing horizon case, our analysis includes a convergence result for D_{ρ_δ} to the fractional gradient D^{s_∞} (where the normalizing constant c_{n,s_∞} is omitted for a cleaner presentation), and a compactness result uniformly in δ. These are used to obtain the following Γ-convergence statement.

Theorem 1.4.11 (Γ-convergence for diverging horizon). *Let $c_\delta := \overline{\rho}(1/\delta)^{-1}$ for $\delta \in (1/\varepsilon,\infty)$, then the functionals $(\mathcal{F}_\delta)_\delta$ Γ-converge as $\delta \to \infty$ with respect to the strong $L^p(\mathbb{R}^n;\mathbb{R}^m)$-topology to the functional $\mathcal{F}^{s_\infty} : L^p(\mathbb{R}^n;\mathbb{R}^m) \to \mathbb{R}$ given by*

$$\mathcal{F}^{s_\infty}(u) = \begin{cases} \displaystyle\int_{\mathbb{R}^n} f(x, D^{s_\infty}u)\,dx & \text{for } u \in H_0^{s_\infty,p}(\Omega;\mathbb{R}^m), \\ \infty & \text{else.} \end{cases}$$

Moreover, the family $(\mathcal{F}_\delta)_\delta$ is equi-coercive.

Theorem 1.1.11 shows that the minimizers of $\mathcal{F}_\delta$ converge in $L^p(\mathbb{R}^n;\mathbb{R}^m)$ as $\delta \to \infty$ up to subsequence to a minimizer of $\mathcal{F}^{s_\infty}$. In particular, with the choice of ρ from Example 1.4.9 a), we find that the models based on the finite-horizon fractional gradient D^s_δ bridge the local and fractional models as the horizon δ moves between 0 and ∞. This complements the localization results in Chapter 3 and 4 for D^s_δ when the fractional parameter s tends to 1.

Applications of nonlocal functionals in the context of image processing is the topic of the final chapter, which is based on the article

[82] E. Davoli, R. Ferreira, C. Kreisbeck and H. Schönberger. Structural changes in nonlocal denoising models arising through bi-level parameter learning. *Applied Mathematics and Optimization*, 88(1):Paper No. 9, 2023. https://doi.org/10.1007/s00245-023-09982-4.

Precisely, we focus on establishing an abstract theory around the bi-level problems in (1.37) where the parameter set Λ need not be compact. As mentioned in Section 1.3.2, we are interested in the extension of the functional $\mathcal{I}$ to the closure $\overline{\Lambda}$ via relaxation, that is

$$\mathcal{I}^{\mathrm{rel}}(\lambda) := \inf\left\{\liminf_{k\to\infty} \mathcal{I}(\lambda_k) : (\lambda_k)_k \subset \Lambda, \lambda_k \to \lambda \text{ in } \overline{\Lambda}\right\} \quad \text{for } \lambda \in \overline{\Lambda},$$

which is the most natural extension from the point of view of minimization; indeed, since we are in the setting (A1) of a first countable topological space $\overline{\Lambda} \subset X$, the functional $\mathcal{I}^{\mathrm{rel}}$ enjoys the properties in Theorem 1.1.7 if $\overline{\Lambda}$ is compact. In order for the relaxation to be practically useful, we determine an explicit characterization of it and show that it actually arises as the upper-level functional of a different bi-level problem.

As preparations, we assume the generic condition that the functionals $\mathcal{R}_\lambda$ are weakly lower semicontinuous on $L^2(\Omega)$ and that $\mathcal{R}_\lambda \not\equiv \infty$ for all $\lambda \in \Lambda$. Moreover, to define the extended problem we assume that the following Mosco-limits exist

$$\overline{\mathcal{R}}_\lambda := \mathrm{Mosc}(L^2)\text{-}\lim_{\lambda' \to \lambda} \mathcal{R}_{\lambda'}, \tag{1.53}$$

for each $\lambda \in \overline{\Lambda}$ with λ' taking values on sequences in Λ; we recall that a L^2-Mosco-limit is equivalent to a Γ-limit with respect to both the strong and weak topology in L^2. We want to highlight that the regularizers $\overline{\mathcal{R}}_\lambda$ for $\lambda \in \overline{\Lambda} \setminus \Lambda$ may have a completely different structure than the original family $\{\mathcal{R}_\lambda\}_\lambda$; for $\lambda \in \Lambda$ on the other hand, we have $\overline{\mathcal{R}}_\lambda = \mathcal{R}_\lambda$ by taking a constant sequence $\lambda' = \lambda$ and recalling (1.4).

We are now in the position to introduce the extended upper-level functional $\overline{\mathcal{I}} : \overline{\Lambda} \longrightarrow [0, \infty)$ defined via the bi-level problem

$$\begin{aligned} &\text{(Upper-level)} \qquad \text{Minimize } \overline{\mathcal{I}}(\lambda) := \inf_{w \in \overline{K}_\lambda} \|w - u^c\|^2_{L^2(\Omega)} \text{ over } \lambda \in \overline{\Lambda}, \\ &\text{(Lower-level)} \qquad \overline{K}_\lambda := \operatorname*{arg\,min}_{u \in L^2(\Omega)} \overline{\mathcal{J}}_\lambda(u), \end{aligned} \tag{1.54}$$

where $\overline{\mathcal{J}}_\lambda := \|\cdot - u^\eta\|^2_{L^2(\Omega)} + \overline{\mathcal{R}}_\lambda$. Note also that $\overline{\mathcal{I}} = \mathcal{I}$ on Λ given the fact that $\overline{\mathcal{J}}_\lambda = \mathcal{J}_\lambda$ for $\lambda \in \Lambda$. Under the additional assumption that

$$\overline{K}_\lambda \text{ is a singleton for every } \lambda \in \overline{\Lambda} \setminus \Lambda, \tag{1.55}$$

we have the following result.

Theorem 1.4.12. *Suppose that* (1.53) *and* (1.55) *are satisfied. Then, the upper-level functional* $\overline{\mathcal{I}}$ *of the extended bi-level problem* (1.54) *agrees with the relaxation of* $\mathcal{I}$ *in* (1.37)*, that is,* $\overline{\mathcal{I}} = \mathcal{I}^{\mathrm{rel}}$.

Remark 1.4.13. a) The result can be generalized to the more realistic setting of data sets with multiple clean and noisy images, see Theorem 7.2.5 for details; the statement and assumptions remain almost identical, though.

b) Both conditions (1.53) and (1.55) are optimal in a sense, since removing either one of them allows for counterexamples to Theorem 1.4.12, cf. Example 7.2.7. △

We illustrate the versatility of Theorem 1.4.12 by considering the extension of four qualitatively different examples of families of regularizers that are relevant for applications: learning the optimal weight, varying the amount of nonlocality, optimizing the integrability exponent, and tuning the fractional parameter. In each of these cases, the parameter domains are one-dimensional non-closed intervals $\Lambda \subset [-\infty, \infty]$, and we determine the relaxation of the bi-level problem by identifying the Mosco-limits from (1.53) in the closure $\overline{\Lambda}$. The obtained Mosco-limits immediately reveal the type of structural changes that occurs at the boundary of the parameter range. Examples include the vanishing of regularizers, the transition from integral to supremal functionals, and the localization of nonlocal regularizers; in fact, this last effect also appears in the family of regularizers depending on finite-horizon fractional discussed in Section 1.3.2. Once the relaxation is characterized, we investigate for what types of data sets the optimal parameter is attained at the boundary or in the interior of Λ; in case it is attained in the interior, then the parameter is also optimal for the original bi-level problem, yielding structure preservation of the regularizers.

Chapter 2

Extending linear growth functionals to functions of bounded fractional variation

This chapter corresponds to the published article

[189] H. Schönberger. Extending linear growth functionals to functions of bounded fractional variation. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 154(1):304–327, 2024. https://doi.org/10.1017/prm.2023.14.

We note that in this chapter, the fractional exponent is denoted by α, the gradients by ∇, and the fractional Sobolev spaces by $S^{\alpha,p}$. The fractional variation and its singular part are indicated by D^α and D^α_s, respectively.

2.1 Introduction

Motivated from both the practical and theoretical point of view, the study of nonlocal aspects in the calculus of variations has received widespread attention in the literature recently. From applications in peridynamics [158, 196], imaging processing [14, 21, 117] and machine learning [13, 134], to the abstract study of lower semicontinuity [34, 139, 140, 174] and localization [8, 29, 35] of various nonlocal functionals. Especially the introduction of the so-called Riesz fractional gradient by Shieh & Spector [193, 194], which for $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\alpha \in (0,1)$ is defined as

$$\nabla^\alpha \varphi(x) = \mu_{n,\alpha} \int_{\mathbb{R}^n} \frac{\varphi(y)-\varphi(x)}{|y-x|^{n+\alpha}} \frac{y-x}{|y-x|}\, dy \quad \text{for } x \in \mathbb{R}^n,$$

has seen a dramatic rise in interest and has opened up the possibility to study new types of fractional problems. We refer to just a few of the recent works [28,31,92,93,140]. The Riesz fractional gradient provides an alternative to the more well-known fractional Laplacian and shares many similarities with the classical gradient. In fact, it is the unique translationally and rotationally invariant α-homogeneous operator [208], which makes it a canonical choice for a fractional gradient.

The definition of the fractional gradient can be extended in a distributional way to define the naturally associated fractional Sobolev spaces

$$S^{\alpha,p}(\mathbb{R}^n;\mathbb{R}^m) = \{u \in L^p(\mathbb{R}^n;\mathbb{R}^m) \,:\, \nabla^\alpha u \in L^p(\mathbb{R}^n;\mathbb{R}^{m\times n})\}, \tag{2.1}$$

with $\alpha \in (0,1)$ and $p \in [1,\infty]$, see [54, 66, 68, 140] for more details. With these new spaces came an inherent class of variational problems to study, that is, integral functionals depending on the Riesz fractional gradient. Precisely, with $\Omega \subset \mathbb{R}^n$ open and bounded, $p \in (1,\infty)$ and $g \in S^{\alpha,p}(\mathbb{R}^n;\mathbb{R}^m)$, one defines the functions subjected to a typical complementary-value condition

$$S^{\alpha,p}_g(\Omega;\mathbb{R}^m) = \{u \in S^{\alpha,p}(\mathbb{R}^n;\mathbb{R}^m) \,:\, u = g \text{ a.e. in } \Omega^c\},$$

and aims to minimize the functional

$$S^{\alpha,p}_g(\Omega;\mathbb{R}^m) \ni u \mapsto \int_{\mathbb{R}^n} f(x, \nabla^\alpha u(x))\,dx; \tag{2.2}$$

here $f : \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a Carathéodory integrand with suitable p-growth and coercivity bounds.

The weak lower semicontinuity and existence of minimizers of these functionals was initially shown in the scalar setting in [193,194] under the condition of convexity in the second argument of f and later extended to the vectorial case under polyconvexity in [28]. More recently, in [140] the weak lower semicontinuity of the functional in (2.2) was fully characterized in terms of the notion α-quasiconvexity, which is a condition on a function $h : \mathbb{R}^{m\times n} \to \mathbb{R}$ that requires that

$$h(A) \le \int_{(0,1)^n} h(A + \nabla^\alpha \varphi(y))\,dy \quad \text{for all } A \in \mathbb{R}^{m\times n} \text{ and } \varphi \in C^\infty_{\mathrm{per}}((0,1)^n;\mathbb{R}^m),$$

see [140, Definition 4.6]. The proof of this result relied on a method to translate fractional gradients into classical gradients and back by using the identities

$$\nabla^\alpha \varphi = \nabla I_{1-\alpha}\varphi \quad \text{and} \quad \nabla\varphi = \nabla^\alpha(-\Delta)^{\frac{1-\alpha}{2}}\varphi \quad \text{for } \varphi \in C^\infty_c(\mathbb{R}^n), \tag{2.3}$$

and actually revealed that the notion of α-quasiconvexity is independent of $\alpha \in (0,1)$ and equivalent to Morrey's well-known quasiconvexity [165]. Therefore, the weak lower semicontinuity of the functionals in (2.2) can be characterized in the same way as the classical integral functionals in the calculus of variations.

Inspired by the rich history on classical linear growth problems, cf. [10,79,113,120,144,181,183], we build upon the above results and exploit the distributional character of the fractional Sobolev spaces to consider the first class of fractional linear growth functionals in the literature. This class constitutes the natural extension of (2.2) to $p = 1$, namely, functionals of the form

$$\mathcal{F}_\alpha(u) = \int_{\mathbb{R}^n} f(x, \nabla^\alpha u(x))\,dx \quad \text{for } u \in S^{\alpha,1}_g(\Omega;\mathbb{R}^m), \tag{2.4}$$

with $f : \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$ a linear growth Carathéodory integrand and $g \in S^{\alpha,1}(\mathbb{R}^n;\mathbb{R}^m)$.

The immediate difficulty in the minimization of the above functional is the non-reflexivity of $S^{\alpha,1}_g(\Omega;\mathbb{R}^m)$, which prevents the direct method from being used with respect to the weak convergence in $S^{\alpha,1}(\mathbb{R}^n;\mathbb{R}^m)$. Therefore, taking up a similar approach as in the classical case, one can suitably extend the functional $\mathcal{F}_\alpha$ to a larger space of bounded fractional variation, in which compactness holds with respect to the weak* convergence.

These spaces of bounded fractional variation and their properties have already been thoroughly studied by Comi & Stefani and coauthors in [54,66,68] and can be understood as

$$BV^\alpha(\mathbb{R}^n;\mathbb{R}^m) = \{u \in L^1(\mathbb{R}^n;\mathbb{R}^m) \,:\, D^\alpha u \in \mathcal{M}(\mathbb{R}^n;\mathbb{R}^{m\times n})\},$$

with $D^\alpha u$ the so-called fractional variation measure of u defined in a distributional sense. We also use the notation

$$D^\alpha u = \nabla^\alpha u\,dx + D^\alpha_s u,$$

where $\nabla^\alpha u \in L^1(\mathbb{R}^n;\mathbb{R}^{m\times n})$ is the absolutely continuous part of $D^\alpha u$ with respect to the Lebesgue measure and $D^\alpha_s u \in \mathcal{M}(\mathbb{R}^n;\mathbb{R}^{m\times n})$ is the singular part. This new class of bounded variation spaces possesses interesting similarities and differences with the classical BV-spaces and has sparked a lot of further investigations. Aspects such as the description of precise representatives [65], Leibniz rules [67], and the failure of a local chain rule [69] have been considered. Very recently, the fractional total variation has been used in the context of image processing providing a nonlocal alternative to the total variation regularization [14].

For the sake of finding an extension of $\mathcal{F}_\alpha$, we introduce the complementary-value space

$$BV^\alpha_g(\Omega;\mathbb{R}^m) = \{u \in BV^\alpha(\mathbb{R}^n;\mathbb{R}^m) \,:\, u = g \text{ a.e. in } \Omega^c\};$$

bounded sequences in $S^{\alpha,1}_g(\Omega;\mathbb{R}^m)$ will converge up to subsequence to an element of $BV^\alpha_g(\Omega;\mathbb{R}^m)$ with respect to the weak* convergence (see Section 2.3). Therefore, with an eye towards minimization, the natural extension of $\mathcal{F}_\alpha$ to $BV^\alpha_g(\Omega;\mathbb{R}^m)$ is the relaxation defined by

$$\mathcal{F}^{\text{rel}}_\alpha(u) = \inf\left\{\liminf_{j\to\infty} \mathcal{F}_\alpha(u_j) \,:\, (u_j)_j \subset S^{\alpha,1}_g(\Omega;\mathbb{R}^m),\ u_j \overset{*}{\rightharpoonup} u \text{ in } BV^\alpha_g(\Omega;\mathbb{R}^m)\right\} \tag{2.5}$$

for $u \in BV^\alpha_g(\Omega;\mathbb{R}^m)$. The useful features of the functional $\mathcal{F}^{\text{rel}}_\alpha$ are that it admits a minimizer under suitable coercivity conditions and that minimizing sequences of $\mathcal{F}_\alpha$ converge up to subsequence to minimizers of $\mathcal{F}^{\text{rel}}_\alpha$.

To benefit from these attributes, it is key to find an explicit representation of the relaxed functional. For this, one must, in particular, account for the concentration effects that fractional gradients of sequences in $S^{\alpha,1}_g(\Omega;\mathbb{R}^m)$ can exhibit and how they relate to the singular part of the limiting fractional variation measure. The well-known concept of the (strong) recession function (cf. [144]), which describes the way an integrand f behaves at infinity, is capable of this and is defined as

$$f^\infty(x,A) = \lim_{\substack{(x',A')\to(x,A)\\ t\to\infty}} \frac{f(x',tA')}{t} \quad \text{for } x \in \mathbb{R}^n \text{ and } A \in \mathbb{R}^{m\times n}, \tag{2.6}$$

whenever it exists. We also recall the upper recession function $f^\#$, which is always well-defined, by replacing the limit in (2.6) with a limit superior. In addition, throughout the paper we use the following growth and coercivity bounds

$$|f(x,A)| \leq M|A| + a(x) \quad \text{for all } x \in \mathbb{R}^n \text{ and } A \in \mathbb{R}^{m\times n}, \tag{G}$$

with $M > 0$ and $a \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ and

$$\mu|A| - c \leq f(x,A) \quad \text{for all } x \in \mathbb{R}^n \text{ and } A \in \mathbb{R}^{m\times n}, \tag{C}$$

with $\mu, c > 0$. Note that the growth bound ensures that f has linear growth and $\mathcal{F}_\alpha$ is well-defined and finite. We now state the following representation result for the relaxation of $\mathcal{F}_\alpha$, which is the main result of the paper.

Theorem 2.1.1. *Let $\alpha \in (0,1)$, $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain and $g \in S^{\alpha,1}(\mathbb{R}^n;\mathbb{R}^m)$. Assume $f : \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a Carathéodory integrand that satisfies* (G) *and* (C)*, and that*

$$f^\infty(x,A) \text{ exists and } (f^{\text{qc}})^\#(x,A) = \limsup_{\substack{A'\to A\\ t\to\infty}} \frac{f^{\text{qc}}(x,tA')}{t} \text{ for all } (x,A) \in \overline{\Omega} \times \mathbb{R}^{m\times n}, \tag{2.7}$$

with f^{qc} the quasiconvex envelope of f with respect to its second argument. Then, the relaxation of $\mathcal{F}_\alpha$ in (2.4) *given by* (2.5) *can be represented as*

$$\mathcal{F}^{\text{rel}}_\alpha(u) = \int_\Omega f^{\text{qc}}(x,\nabla^\alpha u)\,dx + \int_{\overline{\Omega}} (f^{\text{qc}})^\#\left(x, \frac{dD^\alpha_s u}{d|D^\alpha_s u|}\right) d|D^\alpha_s u| + \int_{\Omega^c} f(x,\nabla^\alpha u)\,dx, \tag{2.8}$$

for $u \in BV^\alpha_g(\Omega;\mathbb{R}^m)$.

This theorem provides a fractional analogue to the relaxation result in the classical BV-setting [17, 145] and an extension of the p-growth fractional relaxation in [140, Theorem 1.2]. The reason that the quasiconvex envelope arises in the relaxation is related to the fact that quasiconvexity is the correct characterizing notion for lower semicontinuity similarly as in the p-growth case from [140]. However, the integrand remains unchanged for $x \in \Omega^c$, since fractional gradients of weak* convergent sequences in $BV^\alpha_g(\Omega;\mathbb{R}^m)$ converge strongly in sets with a positive distance from Ω (Lemma 2.3.5). Furthermore, the second integral relating to the singular part of the fractional variation is only integrated over $\overline{\Omega}$, because for $u \in BV^\alpha_g(\Omega;\mathbb{R}^m)$ the measure $D^\alpha_s u$ is supported on $\overline{\Omega}$. This follows since the singular part of the fractional variation actually behaves locally, cf. Remark 2.3.4, implying that $D^\alpha_s u = D^\alpha_s g = 0$ outside of $\overline{\Omega}$. A sufficient condition for (2.7) only in terms of f is given in Remark 2.5.1 c).

The proof of the lower bound of the relaxation result hinges on a characterization of the weak* lower semicontinuity of functionals of the form

$$\overline{\mathcal{F}}_\alpha(u) = \int_{\mathbb{R}^n} f(x, \nabla^\alpha u)\, dx + \int_{\overline{\Omega}} f^\infty\left(x, \frac{dD^\alpha_s u}{d|D^\alpha_s u|}\right) d|D^\alpha_s u| \quad \text{for } u \in BV^\alpha_g(\Omega;\mathbb{R}^m), \tag{2.9}$$

see Theorem 2.4.1. It states that the lower semicontinuity is equivalent to $f(x,\cdot)$ being quasiconvex for a.e. $x \in \Omega$ and is proven by using an analogue of the identities in (2.3) for functions of bounded variation as established in [68]. In addition, we make substantial use of the theory of generalized Young measures developed in [7, 97, 144] for linear growth problems, which allow one to capture the oscillation and concentration effects of sequences of measures. A technical issue arises from the fact that we only assume that f^∞ exists for $x \in \overline{\Omega}$, requiring some care to account for the possible mass that comes from outside Ω and concentrates on the boundary $\partial\Omega$.

The construction of a recovery sequence for the upper bound is carried out in two steps. We first find for $u \in BV^\alpha_g(\Omega;\mathbb{R}^m)$ a sequence $(u_j)_j \subset S^{\alpha,1}_g(\Omega;\mathbb{R}^m)$ that converges to u in a strong enough sense so that the values of the functional along the sequence converge. The natural notion, which has been utilized in the classical case [145], is that of area-strict convergence (Definition 2.3.7). To exploit the properties of area-strict convergence, we prove that $S^{\alpha,1}_g(\Omega;\mathbb{R}^m)$ (and even $g + C^\infty_c(\Omega;\mathbb{R}^m)$) is dense in the larger space $BV^\alpha_g(\Omega;\mathbb{R}^m)$ with respect to this convergence, see Theorem 2.3.9. The second step can then restrict to smooth functions to recover the quasiconvexification of the integrand and relies on adaptations of the argument in [140, Theorem 1.2] and the identities in (2.3).

Finally, we complement the relaxation and lower semicontinuity result with corresponding statements about the existence of minimizers under the coercivity condition (C), see Corollary 2.4.2 and Remark 2.5.1 a). This actually requires an improved version of the fractional Poincaré inequality (Proposition 2.3.6) that only involves the fractional variation over a bounded domain. In particular, the area-integrand $f(A) = \sqrt{1+|A|^2} - 1$ in Example 2.4.3 is an admissible candidate, providing a fractional analogue to the famous Plateau problem [118].

An interesting open problem for further study is the relaxation of $\mathcal{F}_\alpha$ when the integrand admits additional dependence on the values of u. Indeed, in the introduction of [65] it is mentioned for $u \in BV^\alpha(\mathbb{R}^n;\mathbb{R}^m)$ that $D^\alpha u$ can be non-zero on sets of Hausdorff dimension $n-1$, just as the classical variation, while the precise representative of u is only defined for $\mathcal{H}^{n-\alpha+\varepsilon}$-a.e. $x \in \mathbb{R}^n$ for any $\varepsilon > 0$. This discrepancy between $n-1$ and $n-\alpha$ is not present in the classical case and makes it hard to deal with the singular part of the relaxation.

The structure of the text is as follows. In Section 2.2 we present the notation and necessary preliminaries such as generalized Young measure theory and fractional calculus. Section 2.3 revolves around the spaces of bounded fractional variation and contains the proof of the density result with respect to area-strict convergence. The next section is devoted to the characterization of the weak* lower semicontinuity of extended functionals as in (2.9), and Section 2.5 rounds off the paper with the proof of Theorem 2.1.1.

2.2 Preliminaries

2.2.1 Notation

The ball centered at $x \in \mathbb{R}^n$ with radius $r > 0$ is denoted by $B_r(x) = \{y \in \mathbb{R}^n : |x - y| < r\}$. The notation $E \Subset F$ for sets $E, F \subset \mathbb{R}^n$ means that E is compactly contained in F, i.e. $\overline{E} \subset F$ and $\overline{E}$ is compact. We denote by

$$\mathbb{1}_E(x) = \begin{cases} 1 & \text{for } x \in E, \\ 0 & \text{otherwise,} \end{cases} \qquad x \in \mathbb{R}^n,$$

the indicator function of a set $E \subset \mathbb{R}^n$.

By $\mathrm{Lip}_b(\mathbb{R}^n)$ and $\mathrm{Lip}_c(\mathbb{R}^n)$, we refer to all the functions $\psi : \mathbb{R}^n \to \mathbb{R}$ that are Lipschitz continuous and bounded or Lipschitz continuous with compact support on $\mathbb{R}^n$, respectively; we write $\mathrm{Lip}(\psi)$ for the Lipschitz constant of ψ. Furthermore, for $X \subset \mathbb{R}^n$ open or closed we denote by $C_0(X)$ the Banach space obtained by taking the closure of the smooth compactly supported functions $C_c^\infty(X)$ with respect to the supremum norm. In particular, if X is compact then $C_0(X)$ consists of all continuous functions from X to $\mathbb{R}$.

The space $\mathcal{M}(X)$ consists of all finite Radon measures on X and is the dual space of $C_0(X)$. As such, we say that $(\mu_j)_j \subset \mathcal{M}(X)$ converges weak* to $\mu \in \mathcal{M}(X)$ if $\int_X \varphi \, d\mu_j \to \int_X \varphi \, d\mu$ for all $\varphi \in C_0(X)$. More generally, one can define for $f : X \to \mathbb{R}$ Borel measurable and $\mu \in \mathcal{M}(X)$ the duality bracket $\langle f, \mu \rangle = \int_X f \, d\mu$. By $\mathcal{M}^+(X)$ and $\mathcal{M}^1(X)$ we denote the space of positive and probability measures, respectively. We utilize the usual notation for the Radon-Nikodým derivative and for $\mu \in \mathcal{M}(X)$ the Radon-Nikodým derivative with respect to the Lebesgue measure is written as $\frac{d\mu}{dx} \in L^1(X)$, while $\mu^s \in \mathcal{M}(X)$ represents the singular part of μ with respect to the Lebesgue measure. The measure $|\mu| \in \mathcal{M}^+(X)$ constitutes the total variation measure of $\mu \in \mathcal{M}(X)$.

Finally, for $U \subset \mathbb{R}^n$ open we write $BV(U)$ for the space of functions of bounded variation and denote by Du the total variation measure of a function $u \in BV(U)$. We use in this instance ∇u for the absolutely continuous part of Du and $D_s u$ for the singular part of Du with respect to the Lebesgue measure. The variant $BV_{\mathrm{loc}}(\mathbb{R}^n)$ consists of the functions that lie in $BV(U)$ for all open and bounded $U \subset \mathbb{R}^n$. We refer to [11, 210] for more details on functions of bounded variation. All of the mentioned spaces also possess vector-valued counterparts, which are denoted in the second argument like, for example, $BV(U; \mathbb{R}^m)$ and $\mathcal{M}(X; \mathbb{R}^N)$ with $m, N \in \mathbb{N}$.

2.2.2 Generalized Young measures

Generalized Young measures are a tool to study the asymptotic behavior of sequences of functions or even measures and are able to capture both the oscillation and concentration effects. Therefore, they are very well suited for studying linear growth problems in the calculus of variations. In this section we recall the basic definitions and properties that we need in the paper. We refer to [144,182] for more on this topic.

We begin with the definition of the (strong) recession function, which encodes the values of an integrand at infinity. For $U \subset \mathbb{R}^n$ open and bounded and $f : U \times \mathbb{R}^N \to \mathbb{R}$ it is defined as

$$f^\infty(x, A) = \lim_{\substack{(x', A') \to (x, A) \\ t \to \infty}} \frac{f(x', tA')}{t} \quad \text{for } x \in \overline{U} \text{ and } A \in \mathbb{R}^N,$$

provided the limit exists. If the limit exists, then $f^\infty : \overline{U} \times \mathbb{R}^N \to \mathbb{R}$ is automatically jointly continuous and positively homogeneous in the second argument. We now present the definition of a generalized Young measure, see [144] or [17, Definition 2.3].

Definition 2.2.1. *Let $U \subset \mathbb{R}^n$ be open and bounded, then a triple $\nu = (\nu_x, \lambda_\nu, \nu_x^\infty)$ is called a (generalized) Young measure on U with values in $\mathbb{R}^N$, we write $\nu \in Y(U;\mathbb{R}^N)$, if:*

(i) $(\nu_x)_{x\in U} \subset \mathcal{M}^1(\mathbb{R}^N)$ is a parametrized family of probability measures on $\mathbb{R}^N$;

(ii) $\lambda_\nu \in \mathcal{M}_+(\overline{U})$ is a positive measure on $\overline{U}$;

(iii) $(\nu_x^\infty)_{x\in\overline{U}} \subset \mathcal{M}^1(\mathbb{S}^{N-1})$ is a parametrized family of probability measures on $\mathbb{S}^{N-1}$.

Additionally, it is required that $x \mapsto \langle|\cdot|, \nu_x\rangle \in L^1(U)$ and the maps $x \mapsto \langle f(x,\cdot), \nu_x\rangle$ and $x \mapsto \langle f^\infty(x,\cdot), \nu_x^\infty\rangle$ are respectively Lebesgue measurable and λ_ν-measurable for all Carathéodory integrands $f : U \times \mathbb{R}^N \to \mathbb{R}$ for which f^∞ exists.

Intuitively, the Young measure is designed so that $(\nu_x)_{x\in U}$ encodes the oscillations, while λ^ν determines the location and size of the concentrations, and $(\nu_x^\infty)_{x\in\overline{U}}$ the direction of the concentrations. The main result about generalized Young measures is that bounded sequences of measures generate Young measures up to subsequence. Precisely, the following statement is a combination of [144, Theorem 7 and Proposition 2].

Theorem 2.2.2. *Let $U \subset \mathbb{R}^n$ be open and bounded and $(\mu_j)_j \subset \mathcal{M}(\overline{U};\mathbb{R}^N)$ a sequence such that $\sup_j|\mu_j|(\overline{U}) < \infty$. Then, there exists a subsequence (not relabeled) and a Young measure $\nu \in Y(U;\mathbb{R}^N)$ with*

$$\begin{aligned}\lim_{j\to\infty} \int_U f\left(x, \frac{d\mu_j}{dx}\right) dx + \int_{\overline{U}} f^\infty\left(x, \frac{d\mu_j^s}{d|\mu_j^s|}\right) d|\mu_j^s| \\ = \int_U \langle f(x,\cdot), \nu_x\rangle\, dx + \int_{\overline{U}} \langle f^\infty(x,\cdot), \nu_x^\infty\rangle\, d\lambda_\nu,\end{aligned}$$

for all Carathéodory integrands $f : U \times \mathbb{R}^N \to \mathbb{R}$ for which f^∞ exists.

In the setting of the above theorem, we say that $(\mu_j)_j$ generates the Young measure ν. We can also associate to a $\mu \in \mathcal{M}(\overline{U};\mathbb{R}^N)$ the elementary Young measure $\delta[\mu] \in Y(U;\mathbb{R}^N)$ with

$$(\delta[\mu])_x = \delta_{\frac{d\mu}{dx}(x)}, \quad \lambda_{\delta[\mu]} = |\mu^s| \quad \text{and} \quad (\delta[\mu])_x^\infty = \delta_{\frac{d\mu^s}{d|\mu^s|}(x)}, \tag{2.10}$$

with δ_A the dirac measure at a point $A \in \mathbb{R}^N$. One can then interpret the convergence in Theorem 2.2.2 as weak* convergence of $\delta[\mu_j]$ to ν in $Y(U;\mathbb{R}^N)$, where the duality arises from testing Young measures with integrands f. When we have an integrand without a well-defined strong recession function, we can still define the upper recession function

$$f^\#(x,A) = \limsup_{\substack{(x',A')\to(x,A)\\ t\to\infty}} \frac{f(x',tA')}{t} \quad \text{for } x \in \overline{U} \text{ and } A \in \mathbb{R}^N. \tag{2.11}$$

Then, if $(\mu_j)_j \in \mathcal{M}(\overline{U};\mathbb{R}^N)$ generates the Young measure $\nu \in Y(U;\mathbb{R}^N)$ and $f : U \times \mathbb{R}^N \to \mathbb{R}$ is jointly upper semicontinuous, it holds that

$$\begin{aligned}\limsup_{j\to\infty} \int_U f\left(x, \frac{d\mu_j}{dx}\right) dx + \int_{\overline{U}} f^\#\left(x, \frac{d\mu_j^s}{d|\mu_j^s|}\right) d|\mu_j^s| \\ \leq \int_U \langle f(x,\cdot), \nu_x\rangle\, dx + \int_{\overline{U}} \langle f^\#(x,\cdot), \nu_x^\infty\rangle\, d\lambda_\nu,\end{aligned} \tag{2.12}$$

see [17, Corollary 2.10].

2.2.3 Fractional calculus

Here we introduce the fractional operators and their properties, which we use throughout the paper. Firstly, for an integrable function $u \in L^1(\mathbb{R}^n)$, the Riesz potential $I_\alpha u$ of order $\alpha \in (0,n)$ is defined as

$$I_\alpha u(x) = \frac{1}{\gamma_{n,\alpha}} \int_{\mathbb{R}^n} \frac{u(y)}{|x-y|^{n-\alpha}}\, dy \quad \text{for } x \in \mathbb{R}^n,$$

where $\gamma_{n,\alpha} = \pi^{n/2} 2^\alpha \frac{\Gamma(\alpha/2)}{\Gamma((n-\alpha)/2)}$. It is well-known that the above integral is finite for a.e. $x \in \mathbb{R}^n$ and $I_\alpha u \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, cf. [163, 199]. Next up, we have the three different fractional differential operators: The Riesz fractional gradient, the fractional divergence and the fractional Laplacian. We introduce these notions for the class of bounded Lipschitz functions. Precisely, for $\alpha \in (0,1)$ and $\varphi \in \mathrm{Lip}_b(\mathbb{R}^n)$ the Riesz fractional gradient $\nabla^\alpha \varphi : \mathbb{R}^n \to \mathbb{R}^n$ is given by

$$\nabla^\alpha \varphi(x) = \mu_{n,\alpha} \int_{\mathbb{R}^n} \frac{\varphi(y)-\varphi(x)}{|y-x|^{n+\alpha}} \frac{y-x}{|y-x|}\, dy \quad \text{for } x \in \mathbb{R}^n, \tag{2.13}$$

with $\mu_{n,\alpha} = 2^\alpha \pi^{-n/2} \frac{\Gamma((n+\alpha+1)/2)}{\Gamma((1-\alpha)/2)}$, and the fractional Laplacian $(-\Delta)^{\alpha/2}\varphi : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$(-\Delta)^{\alpha/2}\varphi(x) = \nu_{n,\alpha} \int_{\mathbb{R}^n} \frac{\varphi(y)-\varphi(x)}{|y-x|^{n+\alpha}}\, dy \quad \text{for } x \in \mathbb{R}^n,$$

with $\nu_{n,\alpha} = 2^\alpha \pi^{-n/2} \frac{\Gamma((n+\alpha)/2)}{\Gamma(-\alpha/2)}$. Both these operators are well-defined and bounded functions for $\varphi \in \mathrm{Lip}_b(\mathbb{R}^n)$, see [66, Section 2.2]. Finally, for a vector-valued function $\varphi \in \mathrm{Lip}_b(\mathbb{R}^n;\mathbb{R}^n)$, the fractional divergence $\mathrm{div}^\alpha \varphi : \mathbb{R}^n \to \mathbb{R}$ is the natural analogue of the fractional gradient

$$\mathrm{div}^\alpha \varphi(x) = \mu_{n,\alpha} \int_{\mathbb{R}^n} \frac{\varphi(y)-\varphi(x)}{|y-x|^{n+\alpha}} \cdot \frac{y-x}{|y-x|}\, dy \quad \text{for } x \in \mathbb{R}^n, \tag{2.14}$$

which is also a well-defined bounded function. We note that it is proven in [208] that these three fractional differential operators are the unique operators that satisfy translational and rotational invariance, α-homogeneity and a weak requirement of continuity. The fractional gradient and divergence are dual, in the sense that for all $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\psi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$ the integration by parts

$$\int_{\mathbb{R}^n} \varphi\, \mathrm{div}^\alpha \psi\, dx = -\int_{\mathbb{R}^n} \nabla^\alpha \varphi \cdot \psi\, dx \tag{2.15}$$

holds. For more on these differential operators, such as composition rules and extension to different orders than $\alpha \in (0,1)$, we refer to [208].

2.3 Spaces of bounded fractional variation

The spaces of bounded fractional variation were first introduced by Comi & Stefani in the recent series of papers [54,66,68]. We recall the definition of these spaces, which is based on the fractional divergence in (2.14).

Definition 2.3.1. *Let $\alpha \in (0,1)$. A function $u \in L^1(\mathbb{R}^n)$ belongs to $BV^\alpha(\mathbb{R}^n)$ if*

$$\sup\left\{ \int_{\mathbb{R}^n} u\, \mathrm{div}^\alpha \varphi\, dx \ :\ \varphi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n), \|\varphi\|_{L^\infty(\mathbb{R}^n;\mathbb{R}^n)} \leq 1 \right\} < \infty. \tag{2.16}$$

It follows from the structure theorem [66, Theorem 3.2] that $u \in BV^\alpha(\mathbb{R}^n)$ if and only if $u \in L^1(\mathbb{R}^n)$ and there exists a (necessarily unique) finite vector-valued Radon measure $D^\alpha u \in \mathcal{M}(\mathbb{R}^n;\mathbb{R}^n)$ such that

$$\int_{\mathbb{R}^n} u \operatorname{div}^\alpha \varphi \, dx = -\int_{\mathbb{R}^n} \varphi \cdot dD^\alpha u \qquad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n).$$

The measure $D^\alpha u$ is called the fractional variation measure of u and it constitutes a natural extension of the Riesz fractional gradient (2.13) to the space $BV^\alpha(\mathbb{R}^n)$ based on the integration by parts formula (2.15). The space $BV^\alpha(\mathbb{R}^n)$ endowed with the norm

$$\|u\|_{BV^\alpha(\mathbb{R}^n)} = \|u\|_{L^1(\mathbb{R}^n)} + |D^\alpha u|(\mathbb{R}^n)$$

is a Banach space [66, Corollary 3.4], where $|D^\alpha u|(\mathbb{R}^n)$ denotes the total variation of $D^\alpha u$ on $\mathbb{R}^n$ and equals the left-hand side of (2.16). One can also decompose

$$D^\alpha u = \nabla^\alpha u \, dx + D^\alpha_s u,$$

where $\nabla^\alpha u \in L^1(\mathbb{R}^n;\mathbb{R}^n)$ is the absolutely continuous part of $D^\alpha u$ with respect to the Lebesgue measure and $D^\alpha_s u \in \mathcal{M}(\mathbb{R}^n;\mathbb{R}^n)$ is the singular part. We write $BV^\alpha(\mathbb{R}^n;\mathbb{R}^m)$ for the vector-valued analogue with matrix-valued fractional variation. We also introduce the fractional Sobolev space with exponent $p = 1$

$$S^{\alpha,1}(\mathbb{R}^n) = \{u \in BV^\alpha(\mathbb{R}^n) \,:\, D^\alpha_s u = 0\},$$

which consists of those BV^α-functions with an absolutely continuous fractional variation. In fact, this corresponds to the space $S^{\alpha,p}(\mathbb{R}^n)$ defined in (2.1) when $p = 1$, that is, the functions $u \in L^p(\mathbb{R}^n)$ with weak fractional gradient $\nabla^\alpha u \in L^p(\mathbb{R}^n;\mathbb{R}^n)$; see e.g. [54,66,68,140] for more on these fractional Sobolev spaces.

As in [140], the main tool we use to prove the lower semicontinuity and relaxation result is a method to transform the fractional gradient into the classical gradient and back. It relies on the Riesz potential and fractional Laplacian and is proven in the BV-framework in [66, Lemma 3.28].

Proposition 2.3.2. *Let $\alpha \in (0,1)$, then the following holds:*

(*i*) *For $u \in BV^\alpha(\mathbb{R}^n)$ one has that $v = I_{1-\alpha} u \in BV_{\mathrm{loc}}(\mathbb{R}^n)$ with $Dv = D^\alpha u$ in $\mathcal{M}(\mathbb{R}^n;\mathbb{R}^n)$.*

(*ii*) *For $v \in BV(\mathbb{R}^n)$ one has that $u = (-\Delta)^{\frac{1-\alpha}{2}} v \in BV^\alpha(\mathbb{R}^n)$ with $D^\alpha u = Dv$ in $\mathcal{M}(\mathbb{R}^n;\mathbb{R}^n)$ and*

$$\|u\|_{BV^\alpha(\mathbb{R}^n)} \leq c_{n,\alpha} \|v\|_{BV(\mathbb{R}^n)}.$$

Another ingredient we need is the Leibniz rule for the fractional variation, in order to employ localization techniques. We define for $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\psi \in \mathrm{Lip}_c(\mathbb{R}^n)$ the operator

$$\nabla^\alpha_{\mathrm{NL}}(\varphi,\psi)(x) = \mu_{n,\alpha} \int_{\mathbb{R}^n} \frac{(y-x)(\varphi(y)-\varphi(x))(\psi(y)-\psi(x))}{|y-x|^{n+\alpha+1}} \, dy, \quad \text{for } x \in \mathbb{R}^n,$$

which can be continuously extended to $\varphi \in L^1(\mathbb{R}^n)$. The following Leibniz rule for BV^α-functions is from [65, Lemma 5.6], see also [67] for more general Leibniz rules.

Lemma 2.3.3. *Let $\alpha \in (0,1)$, $\psi \in \mathrm{Lip}_c(\mathbb{R}^n)$ and $u \in BV^\alpha(\mathbb{R}^n)$. Then, $\psi u \in BV^\alpha(\mathbb{R}^n)$ with*

$$D^\alpha(\psi u) = \psi D^\alpha u + (u \nabla^\alpha \psi + \nabla^\alpha_{\mathrm{NL}}(u,\psi)) \, dx$$

and there is a constant $C = C(n,\alpha) > 0$ such that

$$\|u\nabla^\alpha \psi + \nabla^\alpha_{\mathrm{NL}}(u,\psi)\|_{L^1(\mathbb{R}^n;\mathbb{R}^n)} \leq C \|\psi\|^{1-\alpha}_{L^\infty(\mathbb{R}^n)} \mathrm{Lip}(\psi)^\alpha \|u\|_{L^1(\mathbb{R}^n)}. \tag{2.17}$$

Remark 2.3.4. Even though the fractional variation is a nonlocal object, the above Leibniz rule implies that the singular part of the fractional variation behaves locally. Indeed, if we have $u, v \in BV^\alpha(\mathbb{R}^n)$ with $u = v$ in an open set $U \subset \mathbb{R}^n$, we find for all $\chi \in C_c^\infty(U)$ (extended to $\mathbb{R}^n$ as zero) that

$$\chi D_s^\alpha u = D_s^\alpha(\chi u) = D_s^\alpha(\chi v) = \chi D_s^\alpha v. \qquad \triangle$$

For our minimization problems, we restrict to functions satisfying a complementary-value condition, which is a nonlocal counterpart of the common Dirichlet boundary conditions. For $\Omega \subset \mathbb{R}^n$ open and bounded we define

$$BV_0^\alpha(\Omega) = \{u \in BV^\alpha(\mathbb{R}^n) \,:\, u = 0 \text{ a.e. in } \Omega^c\},$$

and for $g \in S^{\alpha,1}(\mathbb{R}^n)$

$$BV_g^\alpha(\Omega) = g + BV_0^\alpha(\Omega).$$

Here, we take $g \in S^{\alpha,1}(\mathbb{R}^n)$ since our initial motivation comes from studying linear growth functionals on the fractional Sobolev space. With this in mind, we also introduce the spaces $S_0^{\alpha,1}(\Omega)$ and $S_g^{\alpha,1}(\Omega)$ in a similar way as above. For $u \in BV_g^\alpha(\Omega)$, it follows that the singular part $D_s^\alpha u$ has support inside $\overline{\Omega}$, because of the local behavior of the singular part of the fractional variation (cf. Remark 2.3.4).

A key reason to consider the fractional BV-spaces as extension of the fractional Sobolev spaces, is the property that bounded sequences have convergent subsequences in $BV_g^\alpha(\Omega)$ in an appropriate sense. We say that $(u_j)_j \subset BV_g^\alpha(\Omega)$ converges weak* to $u \in BV_g^\alpha(\Omega)$ if

$$u_j \to u \text{ in } L^1(\mathbb{R}^n) \quad \text{and} \quad D^\alpha u_j \overset{*}{\rightharpoonup} D^\alpha u \text{ in } \mathcal{M}(\mathbb{R}^n;\mathbb{R}^n) \text{ as } j \to \infty.$$

A direct application of the Banach-Alaoglu theorem and the compactness result [66, Theorem 3.16] shows that bounded sequences in $BV_g^\alpha(\Omega)$ admit weak* convergent subsequences. Moreover, we have the following result stating that weak* convergence improves to strong L^1-convergence outside Ω. We omit the proof as it is almost identical to that of [140, Lemma 2.12].

Lemma 2.3.5. *Let $\alpha \in (0,1)$, Ω be open and bounded and $g \in S^{\alpha,1}(\mathbb{R}^n)$. If $(u_j)_j \subset BV_g^\alpha(\Omega)$ converges weak* to u in $BV_g^\alpha(\Omega)$, then for every open $\Omega' \Supset \Omega$ we find*

$$\nabla^\alpha u_j \to \nabla^\alpha u \textit{ in } L^1((\Omega')^c;\mathbb{R}^n) \textit{ as } j \to \infty.$$

We also need an improved version of the Poincaré inequality for fractional BV-functions in [68], which only requires a bound on the fractional variation on some open and bounded set as opposed to the whole space $\mathbb{R}^n$. This allows us to consider interesting integrands with slightly weaker coercivity properties, such as the area-integrand in Example 2.4.3.

Proposition 2.3.6. *Let $\alpha \in (0,1)$ and Ω be open and bounded, then there exists an open and bounded set $\Omega' \Supset \Omega$ and a constant $C = C(\Omega, n, \alpha) > 0$ such that*

$$\|u\|_{BV^\alpha(\mathbb{R}^n)} \le C|D^\alpha u|(\Omega'),$$

for all $u \in BV_0^\alpha(\Omega)$.

Proof. Define for $r > 0$ the function

$$\chi : \mathbb{R}^n \to \mathbb{R}, \quad \chi(x) = \max\{1 - r\, d(x,\Omega), 0\},$$

with $d(x,\Omega)$ the distance from x to Ω. Then, we have $\chi \in \mathrm{Lip}_c(\mathbb{R}^n)$, $\mathrm{Lip}(\chi) \leq r$, $\chi \equiv 1$ on Ω and

$$\operatorname{supp}(\chi) = \Omega_r := \{x \in \mathbb{R}^n \,:\, d(x,\Omega) \leq 1/r\}.$$

We deduce that $u = \chi u$ and conclude from the Leibniz rule (Lemma 2.3.3) that

$$D^\alpha u = D^\alpha(\chi u) = \chi D^\alpha u + (u\nabla^\alpha \chi + \nabla^\alpha_{\mathrm{NL}}(u,\chi))\,dx.$$

Therefore, we find by (2.17) that

$$\|\nabla^\alpha u\|_{L^1(\Omega_r^c;\mathbb{R}^n)} = \|u\nabla^\alpha\chi + \nabla^\alpha_{\mathrm{NL}}(u,\chi)\|_{L^1(\Omega_r^c;\mathbb{R}^n)} \leq Cr^\alpha \|u\|_{L^1(\mathbb{R}^n)}. \tag{2.18}$$

Now using Hölder's inequality on the scale of Lorentz spaces, see [122, Chapter 1.4] for an introduction on Lorentz spaces, in combination with the weak Gagliardo-Nirenberg-Sobolev inequality from [68, Theorem 3.8] yields

$$\|u\|_{L^1(\Omega)} \leq \|\mathbb{1}_\Omega\|_{L^{\frac{n}{\alpha},1}(\mathbb{R}^n)} \|u\|_{L^{\frac{n}{n-\alpha},\infty}(\mathbb{R}^n)} \leq \frac{n|\Omega|^{\alpha/n}}{\alpha} c_{n,\alpha}|D^\alpha u|(\mathbb{R}^n) = c_{n,\alpha,\Omega}|D^\alpha u|(\mathbb{R}^n), \tag{2.19}$$

for all $u \in BV^\alpha_0(\Omega)$. If we choose $r > 0$ such that $Cr^\alpha \leq (2c_{n,\alpha,\Omega})^{-1}$ we obtain from (2.18) and (2.19) that

$$\begin{aligned}\|u\|_{L^1(\Omega)} &\leq c_{n,\alpha,\Omega}\left(|D^\alpha u|(\Omega_r) + |D^\alpha u|(\Omega_r^c)\right)\\ &\leq c_{n,\alpha,\Omega}\left(|D^\alpha u|(\Omega_r) + \frac{1}{2c_{n,\alpha,\Omega}}\|u\|_{L^1(\Omega)}\right),\end{aligned}$$

which, after rewriting, becomes

$$\|u\|_{L^1(\Omega)} \leq 2c_{n,\alpha,\Omega}|D^\alpha u|(\Omega_r). \tag{2.20}$$

Therefore, we obtain

$$\begin{aligned}\|u\|_{BV^\alpha(\mathbb{R}^n)} &= \|u\|_{L^1(\Omega)} + |D^\alpha u|(\Omega_r) + |D^\alpha u|(\Omega_r^c)\\ &\leq \left(1 + \frac{1}{2c_{n,\alpha,\Omega}}\right)\|u\|_{L^1(\Omega)} + |D^\alpha u|(\Omega_r)\\ &\leq (2c_{n,\alpha,\Omega}+2)|D^\alpha u|(\Omega_r),\end{aligned}$$

which proves the result with any open set $\Omega' \Supset \Omega_r$. □

Next, to extend the linear growth functionals from $S^{\alpha,1}_g(\Omega;\mathbb{R}^m)$ to $BV^\alpha_g(\Omega;\mathbb{R}^m)$ we need to be able to approximate functions in $BV^\alpha_g(\Omega;\mathbb{R}^m)$ with functions in $S^{\alpha,1}_g(\Omega;\mathbb{R}^m)$ in a strong enough sense to also have convegence of the functional values. However, since $S^{\alpha,1}_g(\Omega;\mathbb{R}^m)$ is closed with respect to the BV^α-norm, we have to consider a convergence notion that is also weaker than the one induced by the norm. The relevant notion here is a type of area-strict convergence, which is in between norm convergence and weak* convergence. Like in [145], we define the area-functional for $\mu \in \mathcal{M}(\mathbb{R}^n;\mathbb{R}^N)$ and $U \subset \mathbb{R}^n$ Borel measurable as

$$\langle\mu\rangle(U) := \int_U \sqrt{1 + \left|\frac{d\mu}{dx}\right|^2}\,dx + |\mu^s|(U),$$

with μ^s the singular part of μ with respect to the Lebesgue measure. We also write $\langle A\rangle := \sqrt{1+|A|^2}$ for $A \in \mathbb{R}^{m\times n}$.

Definition 2.3.7 (area-strict convergence). *We say that a sequence $(u_j)_j \subset BV^\alpha_g(\Omega;\mathbb{R}^m)$ converges area-strictly to $u \in BV^\alpha_g(\Omega;\mathbb{R}^m)$ if $u_j \to u$ in $L^1(\mathbb{R}^n;\mathbb{R}^m)$,*

$$\langle D^\alpha u_j\rangle(\overline{\Omega}) \to \langle D^\alpha u\rangle(\overline{\Omega}) \quad \textit{and} \quad \nabla^\alpha u_j \to \nabla^\alpha u \textit{ in } L^1(\Omega^c;\mathbb{R}^{m\times n}) \textit{ as } j \to \infty.$$

Remark 2.3.8. The key property of area-strict convergence is that when restricted to $\overline{\Omega}$, the sequence $(D^\alpha u_j)_j \subset \mathcal{M}(\overline{\Omega};\mathbb{R}^{m\times n})$ generates the elementary Young measure $\delta[D^\alpha u]$ (cf. (2.10) and [182, Proposition 12.4]). The convergence $\nabla^\alpha u_j \to \nabla^\alpha u$ in $L^1(\Omega^c;\mathbb{R}^{m\times n})$ also excludes any concentration effects happening outside Ω, which are in general not ruled out by Lemma 2.3.5. △

We now prove a density result with respect to the area-strict convergence, which plays a key role in the construction of a recovery sequence when extending the linear growth functionals. The proof exploits the fractional Leibniz rule and invariance properties of the fractional variation to incorporate the partition of unity and mollification techniques from the classical case (as in e.g. [182, Lemma 11.1]). Note that we implicitly assume that functions in $C_c^\infty(\Omega;\mathbb{R}^m)$ are extended to $\mathbb{R}^n$ as zero.

Theorem 2.3.9. *Let $\alpha \in (0,1)$, $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain and $g \in S^{\alpha,1}(\mathbb{R}^n;\mathbb{R}^m)$. For every $u \in BV^\alpha_g(\Omega;\mathbb{R}^m)$ there exists a sequence $(u_j)_j \subset g + C_c^\infty(\Omega;\mathbb{R}^m)$ such that*

$$u_j \to u \textit{ area-strictly in } BV^\alpha_g(\Omega;\mathbb{R}^m).$$

Proof. Step 1: Shrinking the support. We show that for every $\varepsilon > 0$, we can find a $v \in BV^\alpha_g(\Omega;\mathbb{R}^m)$ such that $\operatorname{supp}(v-g) \Subset \Omega$,

$$\|u-v\|_{L^1(\mathbb{R}^n;\mathbb{R}^m)} + \|\nabla^\alpha u - \nabla^\alpha v\|_{L^1(\Omega^c;\mathbb{R}^{m\times n})} \le \varepsilon \quad \text{and} \quad \langle D^\alpha v\rangle(\overline{\Omega}) \le \langle D^\alpha u\rangle(\overline{\Omega}) + \varepsilon. \tag{2.21}$$

To this aim, we take a representative of u that is identical to g in Ω^c and set $u_0 := u - g$. Then, since Ω is a Lipschitz domain, we find a partition of unity $\chi_0, \chi_1, \cdots, \chi_N \subset C_c^\infty(\mathbb{R}^n)$ and vectors $\zeta_1, \cdots, \zeta_N \subset \mathbb{R}^n$ such that

$$\sum_{i=0}^N \chi_i = 1 \text{ on } \Omega, \quad \chi_0 \in C_c^\infty(\Omega), \quad \text{and} \quad \operatorname{supp}(\tau_{\lambda\zeta_i}(\chi_i u_0)) \Subset \Omega, \tag{2.22}$$

for all $\lambda > 0$ small enough, where $\tau_\zeta(w)(x) := w(x+\zeta)$ denotes translation by $\zeta \in \mathbb{R}^n$. In view of Lemma 2.3.3 we can define the function

$$v = g + \chi_0 u_0 + \sum_{i=1}^N \tau_{\lambda\zeta_i}(\chi_i u_0) \in BV^\alpha_g(\Omega;\mathbb{R}^m),$$

which satisfies $\operatorname{supp}(v-g) \Subset \Omega$ due to (2.22). Using the first identity from (2.22), we have that

$$\|u-v\|_{L^1(\mathbb{R}^n;\mathbb{R}^m)} \le \sum_{i=1}^N \|\chi_i u_0 - \tau_{\lambda\zeta_i}(\chi_i u_0)\|_{L^1(\mathbb{R}^n;\mathbb{R}^m)} \le \varepsilon/2,$$

for λ small enough given the continuity of translation on $L^1(\mathbb{R}^n;\mathbb{R}^m)$. Moreover, we have by the translation invariance of ∇^α that

$$\nabla^\alpha v = \nabla^\alpha u + \sum_{i=1}^N \tau_{\lambda\zeta_i}(\nabla^\alpha(\chi_i u_0)) - \nabla^\alpha(\chi_i u_0)$$

so that the continuity of translation on $L^1(\mathbb{R}^n;\mathbb{R}^{m\times n})$ again yields

$$\|\nabla^\alpha v - \nabla^\alpha u\|_{L^1(\mathbb{R}^n;\mathbb{R}^{m\times n})} = \sum_{i=1}^N \|\tau_{\lambda\zeta_i}(\nabla^\alpha(\chi_i u_0)) - \nabla^\alpha(\chi_i u_0)\|_{L^1(\mathbb{R}^n;\mathbb{R}^{m\times n})} \leq \varepsilon/2$$

for λ small enough. This shows the first part of (2.21) and at the same time that

$$\int_\Omega \langle \nabla^\alpha v\rangle\, dx \leq \int_\Omega \langle \nabla^\alpha u\rangle\, dx + \|\nabla^\alpha v - \nabla^\alpha u\|_{L^1(\Omega;\mathbb{R}^{m\times n})} \leq \int_\Omega \langle \nabla^\alpha u\rangle\, dx + \varepsilon/2, \tag{2.23}$$

where we have exploited the 1-Lipschitz continuity of $A \mapsto \langle A\rangle$. Finally, for the singular part we note that

$$D^\alpha_s v = \chi_0 D^\alpha_s u_0 + \sum_{i=1}^N \tau_{\lambda\zeta_i}(\chi_i D^\alpha_s u_0) = \chi_0 D^\alpha_s u + \sum_{i=1}^N \tau_{\lambda\zeta_i}(\chi_i D^\alpha_s u)$$

in virtue of Lemma 2.3.3. Hence, it follows with (2.22) that

$$\begin{aligned} |D^\alpha_s v|(\overline{\Omega}) &\leq \int_{\overline{\Omega}} \chi_0\, d|D^\alpha_s u| + \sum_{i=1}^N \int_{\tau_{-\lambda\zeta_i}(\overline{\Omega})} \chi_i\, d|D^\alpha_s u| \\ &\leq \sum_{i=0}^N \int_{\mathbb{R}^n} \chi_i\, d|D^\alpha_s u| = |D^\alpha_s u|(\overline{\Omega}), \end{aligned}$$

which proves the second part of (2.21) in combination with (2.23).

Step 2: Mollification. Let v be as in Step 1, then we show that there is a $w \in g + C_c^\infty(\Omega;\mathbb{R}^m)$ such that

$$\|v - w\|_{L^1(\mathbb{R}^n;\mathbb{R}^m)} + \|\nabla^\alpha v - \nabla^\alpha w\|_{L^1(\Omega^c;\mathbb{R}^{m\times n})} \leq \varepsilon \quad \text{and} \quad \langle D^\alpha w\rangle(\overline{\Omega}) \leq \langle D^\alpha v\rangle(\overline{\Omega}) + \varepsilon. \tag{2.24}$$

Let $\eta_\delta \in C_c^\infty(B_\delta(0))$ for $\delta > 0$ be a standard mollifier and choose δ small enough such that

$$\eta_\delta * (v - g) \in C_c^\infty(\Omega;\mathbb{R}^m),$$

which is possible since $\operatorname{supp}(v-g) \Subset \Omega$. Setting $w = g + \eta_\delta * (v-g)$, standard properties of mollification show that

$$\|v - w\|_{L^1(\mathbb{R}^n;\mathbb{R}^m)} \leq \varepsilon/2,$$

for δ small enough. Furthermore, by [66, Lemma 3.5] we have

$$\nabla^\alpha w = \nabla^\alpha g + \eta_\delta * \nabla^\alpha(v-g) + \eta_\delta * D^\alpha_s v.$$

In particular, since $\eta_\delta * D^\alpha_s v$ has support inside Ω we have

$$\|\nabla^\alpha v - \nabla^\alpha w\|_{L^1(\Omega^c;\mathbb{R}^{m\times n})} = \|\nabla^\alpha(v-g) - \eta_\delta * \nabla^\alpha(v-g)\|_{L^1(\Omega^c;\mathbb{R}^{m\times n})} \leq \varepsilon/2,$$

for small δ, thus proving the first part of (2.24). Furthermore,

$$\begin{aligned} \langle D^\alpha w\rangle(\overline{\Omega}) &= \int_\Omega \langle \nabla^\alpha g + \eta_\delta * \nabla^\alpha(v-g) + \eta_\delta * D^\alpha_s v\rangle\, dx \\ &\leq \int_\Omega \langle \nabla^\alpha g + \eta_\delta * \nabla^\alpha(v-g)\rangle\, dx + \int_\Omega |\eta_\delta * D^\alpha_s v|\, dx \\ &\leq \int_\Omega \langle \nabla^\alpha g + \eta_\delta * \nabla^\alpha(v-g)\rangle\, dx + |D^\alpha_s v|(\overline{\Omega}) \\ &\leq \int_\Omega \langle \nabla^\alpha v\rangle\, dx + \varepsilon + |D^\alpha_s v|(\overline{\Omega}) = \langle D^\alpha v\rangle(\overline{\Omega}) + \varepsilon, \end{aligned}$$

where in the last line we utilize Lebesgue's dominated convergence for small enough δ, recalling the fact that $\eta_\delta * \nabla^\alpha(v-g) \to \nabla^\alpha(v-g)$ in $L^1(\mathbb{R}^n;\mathbb{R}^{m\times n})$. This yields (2.24).

Step 3: Conclusion. By combining Step 1 and 2 we may find a sequence $(u_j)_j \subset g + C_c^\infty(\Omega;\mathbb{R}^m)$ such that

$$u_j \to u \text{ in } L^1(\mathbb{R}^n;\mathbb{R}^m), \quad \nabla^\alpha u_j \to \nabla^\alpha u \text{ in } L^1(\Omega^c;\mathbb{R}^{m\times n}) \text{ as } j\to\infty$$

and

$$\limsup_{j\to\infty} \langle D^\alpha u_j\rangle(\overline{\Omega}) \leq \langle D^\alpha u\rangle(\overline{\Omega}).$$

In view of this bound we have that $u_j \overset{*}{\rightharpoonup} u$ in $BV_g^\alpha(\Omega;\mathbb{R}^m)$. Therefore, we may use the weak* lower semicontinuity of the area-functional on $\mathcal{M}(\Omega';\mathbb{R}^{m\times n})$ for some $\Omega' \Supset \Omega$ open and bounded, which follows from the convexity of $A \mapsto \langle A\rangle$, and Lebesgue's dominated convergence theorem to conclude that

$$\begin{aligned}\liminf_{j\to\infty} \langle D^\alpha u_j\rangle(\overline{\Omega}) &= \liminf_{j\to\infty} \langle D^\alpha u_j\rangle(\Omega') - \lim_{j\to\infty}\int_{\Omega'\setminus\Omega} \langle \nabla^\alpha u_j\rangle\,dx \\ &\geq \langle D^\alpha u\rangle(\Omega') - \int_{\Omega'\setminus\Omega}\langle\nabla^\alpha u\rangle\,dx = \langle D^\alpha u\rangle(\overline{\Omega}),\end{aligned}$$

which finishes the proof. □

2.4 Lower semicontinuity

In this section we characterize the weak* lower semicontinuity of functionals as in (2.9), which is interesting in its own right and is used in the proof of the main relaxation result in Section 2.5. Recall that a continuous function $h : \mathbb{R}^{m\times n} \to \mathbb{R}$ is called quasiconvex if

$$h(A) \leq \int_{(0,1)^n} h(A+\nabla\varphi(y))\,dy \quad \text{for all } A\in\mathbb{R}^{m\times n} \text{ and } \varphi \in W_0^{1,\infty}((0,1)^n;\mathbb{R}^m),$$

see [75, 165]. We prove the following statement, whose proof relies on the connection between the classical and fractional variation and the theory of generalized Young measures. We note that even though $f^\infty(x,A)$ is only assumed to exist for $x\in\overline{\Omega}$, we do allow the sequences $x'\to x$ from (2.6) to approach from outside $\overline{\Omega}$.

Theorem 2.4.1. *Let $\alpha\in(0,1)$, $\Omega\subset\mathbb{R}^n$ be open and bounded with $|\partial\Omega| = 0$, $g\in S^{\alpha,1}(\mathbb{R}^n;\mathbb{R}^m)$ and $f:\mathbb{R}^n\times\mathbb{R}^{m\times n}\to\mathbb{R}$ a Carathéodory integrand that satisfies* (G). *If*

$$f^\infty(x,A) \text{ exists for all } (x,A)\in\overline{\Omega}\times\mathbb{R}^{m\times n},$$

then the functional

$$\overline{\mathcal{F}}_\alpha(u) = \int_{\mathbb{R}^n} f(x,\nabla^\alpha u)\,dx + \int_{\overline{\Omega}} f^\infty\left(x, \frac{dD_s^\alpha u}{|dD_s^\alpha u|}\right) d|D_s^\alpha u| \quad \text{for } u\in BV_g^\alpha(\Omega;\mathbb{R}^m),$$

is weak lower semicontinuous if and only if $f(x,\cdot)$ is quasiconvex for a.e. $x\in\Omega$.*

Proof. Step 1: Necessity. The weak* lower semicontinuity of $\overline{\mathcal{F}}_\alpha$ implies, in particular, that

$$\mathcal{F}_\alpha(u) = \int_{\mathbb{R}^n} f(x,\nabla^\alpha u(x))\,dx \quad \text{for } u\in S_g^{\alpha,1}(\Omega;\mathbb{R}^m),$$

is weakly lower semicontinuous on $S^{\alpha,1}_g(\Omega;\mathbb{R}^m)$. A simple adaptation of [140, Theorem 4.5] to the case $p = 1$ yields that $f(x,\cdot)$ is quasiconvex for a.e. $x \in \Omega$.

Step 2: Sufficiency. Let $u_j \overset{*}{\rightharpoonup} u$ in $BV^\alpha_g(\Omega;\mathbb{R}^m)$ and fix $\Omega' \Supset \Omega$ open and bounded. By Proposition 2.3.2 (i) we can find a sequence $(v_j)_j \subset BV(\Omega';\mathbb{R}^m)$ and $v \in BV(\Omega';\mathbb{R}^m)$ such that

$$Dv_j = D^\alpha u_j \text{ on } \Omega' \text{ for } j \in \mathbb{N} \quad \text{and} \quad Dv = D^\alpha u \text{ on } \Omega'; \tag{2.25}$$

we can also ensure that $v_j \overset{*}{\rightharpoonup} v$ in $BV(\Omega';\mathbb{R}^m)$ by continuity properties of the Riesz potential, see [163, Theorem 2.1 (i)]. Up to a non-relabeled subsequence, $(Dv_j)_j \subset \mathcal{M}(\overline{\Omega'};\mathbb{R}^{m\times n})$ generates a BV-Young measure $\nu \in Y(\Omega';\mathbb{R}^{m\times n})$ on Ω'. Before we proceed, we can redefine f, similarly to [182, Proof of Theorem 12.25], such that its recession function is defined in a larger region. Indeed, by definition of the strong recession function $f^\infty : \overline{\Omega} \times \mathbb{R}^{m\times n} \to \mathbb{R}$, it is automatically jointly continuous, so that we can continuously extend it to $\overline{\Omega'} \times \mathbb{R}^{m\times n}$. If we set $f' : \overline{\Omega'} \times \mathbb{R}^{m\times n} \to \mathbb{R}$ equal to f on $\overline{\Omega} \times \mathbb{R}^{m\times n}$ and f^∞ on $(\overline{\Omega'} \setminus \overline{\Omega}) \times \mathbb{R}^{m\times n}$, then f' is a Carathéodory integrand with a well-defined strong recession function on $\overline{\Omega'} \times \mathbb{R}^{m\times n}$. Now, applying Theorem 2.2.2 to f' and Ω' yields

$$\begin{aligned}
&\liminf_{j\to\infty} \int_{\Omega'} f(x,\nabla^\alpha u_j)\,dx + \int_{\overline{\Omega}} f^\infty\left(x, \frac{dD^\alpha_s u_j}{|dD^\alpha_s u_j|}\right) d|D^\alpha_s u_j| \\
&\quad\geq \liminf_{j\to\infty} \int_{\Omega'} f'(x,\nabla^\alpha u_j)\,dx + \int_{\overline{\Omega'}} (f')^\infty\left(x, \frac{dD^\alpha_s u_j}{|dD^\alpha_s u_j|}\right) d|D^\alpha_s u_j| \\
&\qquad - \sup_{j\in\mathbb{N}} \int_{\Omega'\setminus\Omega} |(f-f')(x,\nabla^\alpha u_j)|\,dx \\
&\quad\geq \int_{\Omega'} \langle f'(x,\cdot),\nu_x\rangle\,dx + \int_{\overline{\Omega'}} \langle (f')^\infty(x,\cdot),\nu_x^\infty\rangle\,d\lambda^\nu - \sup_{j\in\mathbb{N}} \int_{\Omega'\setminus\Omega} |(f-f')(x,\nabla^\alpha u_j)|\,dx.
\end{aligned} \tag{2.26}$$

Due to the strong convergence $\nabla v_j = \nabla^\alpha u_j \to \nabla^\alpha u = \nabla v$ in sets away from Ω (Lemma 2.3.5), we also find that the support of the concentration measure λ_ν is contained inside $\overline{\Omega}$ and $\nu_x = \delta_{\nabla^\alpha u(x)}$ for a.e. $x \in \Omega' \setminus \Omega$; that is, we find

$$\int_{\Omega'\setminus\Omega} \langle f'(x,\cdot),\nu_x\rangle\,dx = \int_{\Omega'\setminus\Omega} f'(x,\nabla^\alpha u)\,dx. \tag{2.27}$$

Furthermore, since ν is a BV-Young measure generated by $(Dv_j)_j$, we may argue as in [144, Theorem 10] and [182, Theorem 12.25] using the generalized Jensen's inequalities from [144, Theorem 9] in combination with the quasiconvexity, continuity and linear growth of $f(x,\cdot)$ for a.e. $x \in \Omega$ and of $f^\infty(x,\cdot)$ for all $x \in \overline{\Omega}$ (by continuity of f^∞) to conclude

$$\begin{aligned}
\int_\Omega \langle f(x,\cdot),\nu_x\rangle\,dx + \int_{\overline{\Omega}} \langle f^\infty(x,\cdot),\nu_x^\infty\rangle\,d\lambda_\nu &\geq \int_\Omega f(x,\nabla v)\,dx + \int_{\overline{\Omega}} f^\infty\left(x,\frac{dD_s v}{|dD_s v|}\right) d|D_s v| \\
&= \int_\Omega f(x,\nabla^\alpha u)\,dx + \int_{\overline{\Omega}} f^\infty\left(x,\frac{dD^\alpha_s u}{|dD^\alpha_s u|}\right) d|D^\alpha_s u|,
\end{aligned} \tag{2.28}$$

with the last equality exploiting (2.25). Additionally, since $\nabla^\alpha u_j \to \nabla^\alpha u$ strongly in $L^1((\Omega')^c;\mathbb{R}^{m\times n})$ by Lemma 2.3.5, the growth bound (G) and Lebesgue's dominated convergence theorem yields

$$\lim_{j\to\infty} \int_{(\Omega')^c} f(x,\nabla^\alpha u_j)\,dx = \int_{(\Omega')^c} f(x,\nabla^\alpha u)\,dx.$$

Combining this with (2.26), (2.27) and (2.28) results in

$$\begin{aligned}\liminf_{j\to\infty} \overline{\mathcal{F}}_\alpha(u_j) \geq \int_{\Omega'} f'(x,\nabla^\alpha u)\,dx + \int_{(\Omega')^c} f(x,\nabla^\alpha u)\,dx + \int_{\overline{\Omega}} f^\infty\left(x, \frac{dD^\alpha_s u}{|dD^\alpha_s u|}\right) d|D^\alpha_s u| \\ - \sup_{j\in\mathbb{N}} \int_{\Omega'\setminus\Omega} |(f-f')(x,\nabla^\alpha u_j)|\,dx.\end{aligned} \tag{2.29}$$

Finally, since

$$(f-f')^\infty(x,A) = 0 \quad \text{for all } x \in \overline{\Omega} \text{ and } A \in \mathbb{R}^{m\times n},$$

we may choose Ω' potentially smaller and find a $R > 0$ such that

$$|(f-f')(x,A)| \leq \varepsilon|A| \quad \text{for all } x \in \Omega' \text{ and } A \in \mathbb{R}^{m\times n} \text{ with } |A| \geq R,$$

for any given $\varepsilon > 0$. With $C := \sup_j \|\nabla^\alpha u_j\|_{L^1(\mathbb{R}^n;\mathbb{R}^{m\times n})} < \infty$, we obtain

$$\sup_{j\in\mathbb{N}} \int_{\Omega'\setminus\Omega} |(f-f')(x,\nabla^\alpha u_j)|\,dx \leq (MR + \|a\|_{L^\infty(\mathbb{R}^n)})|\Omega'\setminus\Omega| + \varepsilon C,$$

so that we can deduce the result by first letting $\Omega' \downarrow \Omega$, given that $f = f'$ on $\Omega\times\mathbb{R}^{m\times n}$, and secondly letting $\varepsilon \downarrow 0$ in (2.29). □

In order to get the existence of minimizers, we also impose the coercivity bound (C) and utilize the improved Poincaré inequality from Proposition 2.3.6.

Corollary 2.4.2. *Let $\alpha \in (0,1)$, $\Omega \subset \mathbb{R}^n$ be open and bounded with $|\partial\Omega| = 0$, $g \in S^{\alpha,1}(\mathbb{R}^n;\mathbb{R}^m)$ and $f : \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$ a Carathéodory integrand that satisfies* (G) *and* (C)*. If*

$$f^\infty(x,A) \text{ exists for all } (x,A) \in \overline{\Omega} \times \mathbb{R}^{m\times n},$$

and $f(x,\cdot)$ is quasiconvex for a.e. $x \in \Omega$, then

$$\overline{\mathcal{F}}_\alpha(u) = \int_{\mathbb{R}^n} f(x,\nabla^\alpha u)\,dx + \int_{\overline{\Omega}} f^\infty\left(x, \frac{dD^\alpha_s u}{|dD^\alpha_s u|}\right) d|D^\alpha_s u| \quad \text{for } u \in BV^\alpha_g(\Omega;\mathbb{R}^m),$$

admits a minimizer on $BV^\alpha_g(\Omega;\mathbb{R}^m)$.

Proof. We fix $\Omega' \Supset \Omega$ large enough as in Proposition 2.3.6 and such that $M|D^\alpha v|((\Omega')^c) \leq \frac{\mu}{2}|D^\alpha v|(\Omega')$ for all $v \in BV^\alpha_0(\Omega;\mathbb{R}^m)$, which is possible by (2.18) and (2.20). Now using the coercivity condition of f on Ω', the growth bound on $(\Omega')^c$ and Proposition 2.3.6, we find for all $u \in BV^\alpha_g(\Omega;\mathbb{R}^m)$ that

$$\begin{aligned}\overline{\mathcal{F}}_\alpha(u) &\geq \mu|D^\alpha u|(\Omega') - M|D^\alpha u|((\Omega')^c) - C' \\ &\geq \mu|D^\alpha(u-g)|(\Omega') - M|D^\alpha(u-g)|((\Omega')^c) - C'' \\ &\geq \frac{\mu}{2C}\|u-g\|_{BV^\alpha(\mathbb{R}^n;\mathbb{R}^m)} - C''.\end{aligned}$$

Hence, a standard argument using the direct method and the weak* lower semicontinuity from Theorem 2.4.1 finishes the proof. □

Example 2.4.3. An example integrand that satisfies all the hypotheses of Corollary 2.4.2 is

$$f : \mathbb{R}^n \times \mathbb{R}^{m\times n} \to [0,\infty), \quad f(x,A) = \sqrt{1+|A|^2} - 1,$$

since f is convex, $f^\infty(x,A) = |A|$ and

$$|A| - 1 \le f(x,A) \le |A| \quad \text{for all } x \in \mathbb{R}^n \text{ and } A \in \mathbb{R}^{m\times n}.$$

Hence, the following type of fractional area-functional

$$\overline{\mathcal{F}}_\alpha(u) = \int_{\mathbb{R}^n} \sqrt{1+|\nabla^\alpha u|^2} - 1\,dx + |D^\alpha_s u|(\overline{\Omega}),$$

is weak* lower semicontinuous on $BV^\alpha_g(\Omega;\mathbb{R}^m)$ and admits a minimizer.

2.5 Relaxation

We are now in the position to give the proof of the main result. For a Carathéodory integrand $f : \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$ that satisfies the bounds (G) and (C), it follows from [75, Proposition 9.5] that

$$f^{\mathrm{qc}}(x,A) = \inf\left\{\int_{(0,1)^n} f(x, A+\nabla\varphi(y))\,dy \;:\; \varphi \in W^{1,\infty}_0((0,1)^n;\mathbb{R}^m)\right\}$$

for $(x,A) \in \mathbb{R}^n \times \mathbb{R}^{m\times n}$, is a Carathéodory integrand and from [75, Theorem 6.9] that the function $f^{\mathrm{qc}}(x,\cdot)$ is the largest quasiconvex function below $f(x,\cdot)$. Note also that f^{qc} still satisfies (G) and (C) for $x \in \overline{\Omega}$ since $f^{\mathrm{qc}} \le f$ and the lower bound in (C) is quasiconvex in the second argument.

Proof of Theorem 2.1.1. Denote the functional on the right-hand side of (2.8) by $\overline{\mathcal{F}}_\alpha$. We split up the proof into the lower and upper bound.

Step 1: Lower bound. Let $(u_j)_j \subset S^{\alpha,1}_g(\Omega;\mathbb{R}^m)$ with $u_j \overset{*}{\rightharpoonup} u$ in $BV^\alpha_g(\Omega;\mathbb{R}^m)$ as $j \to \infty$, then we can completely follow the proof of Theorem 2.4.1 without using the generalized Jensen's inequalities to conclude (up to a non-relabeled subsequence) that

$$\liminf_{j\to\infty} \mathcal{F}_\alpha(u_j) \ge \int_\Omega \langle f(x,\cdot), \nu_x\rangle\,dx + \int_{\overline{\Omega}} \langle f^\infty(x,\cdot), \nu^\infty_x\rangle\,d\lambda_\nu + \int_{\Omega^c} f(x, \nabla^\alpha u)\,dx,$$

with ν the generalized Young measure generated by the sequence $(\nabla^\alpha u_j)_j$ (on some domain containing Ω). Using the bounds $f \ge f^{\mathrm{qc}}$ and $f^\infty \ge (f^{\mathrm{qc}})^\#$, we can now proceed as in Theorem 2.4.1 by using the Jensen's inequalities for the quasiconvexification f^{qc}, to obtain the lower bound. Here, we make crucial use of the second part of (2.7), since the Jensen's inequalities for upper recession functions in [144, Theorem 9] can only be directly applied in the x-independent case.

Step 2: Upper bound. We first show that we can restrict to the case $u \in g + C^\infty_c(\Omega;\mathbb{R}^m)$ for the upper bound. To this aim, we take $u \in BV^\alpha_g(\Omega;\mathbb{R}^m)$ and a sequence $(u_j)_j \subset g + C^\infty_c(\Omega;\mathbb{R}^m)$ which converges area-strictly to u, possible by Theorem 2.3.9. Hence, Lebesgue's dominated convergence theorem and the growth bound (G) yields

$$\lim_{j\to\infty} \int_{\Omega^c} f(x, \nabla^\alpha u_j)\,dx = \int_{\Omega^c} f(x, \nabla^\alpha u)\,dx. \tag{2.30}$$

Next, if we denote by $g : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ the (jointly) upper semicontinuous envelope of f^{qc} restricted to $\Omega \times \mathbb{R}^{m\times n}$, then it is not hard to verify that $g^\# = (f^{\mathrm{qc}})^\#$ via the definition of the upper recession function in (2.11). We take for $R > 0$ a truncation function $T_R \in C^\infty_c(\mathbb{R}^{m\times n})$ with $0 \le T_R(A) \le 1$ and $T_R \equiv 1$ on $B_R(0)$ and bound

$$f^{\mathrm{qc}}(x,A) \le T_R(A) f^{\mathrm{qc}}(x,A) + T^c_R(A) g(x,A),$$

with $T_R^c(A) := 1 - T_R(A)$. The first integrand on the right-hand side has zero recession function, whereas the second integrand is jointly upper semicontinuous with upper recession function $g^\# = (f^{\mathrm{qc}})^\#$. Applying Theorem 2.2.2 and (2.12) then gives, in combination with the fact that $(D^\alpha u_j)_j \subset \mathcal{M}(\overline{\Omega}; \mathbb{R}^{m\times n})$ generates the elementary Young measure $\delta[D^\alpha u]$ (cf. Remark 2.3.8),

$$\begin{aligned}\liminf_{j\to\infty} \int_\Omega f^{\mathrm{qc}}(x, \nabla^\alpha u_j)\,dx \\ &\leq \lim_{j\to\infty} \int_\Omega T_R(\nabla^\alpha u_j) f^{\mathrm{qc}}(x, \nabla^\alpha u_j)\,dx + \limsup_{j\to\infty} \int_\Omega T_R^c(\nabla^\alpha u_j) g(x, \nabla^\alpha u_j)\,dx \\ &\leq \int_\Omega T_R(\nabla^\alpha u) f^{\mathrm{qc}}(x, \nabla^\alpha u) + T_R^c(\nabla^\alpha u) g(x, \nabla^\alpha u)\,dx + \int_{\overline{\Omega}} (f^{\mathrm{qc}})^\# \left(x, \frac{dD_s^\alpha u}{d|D_s^\alpha u|}\right) d|D_s^\alpha u|.\end{aligned}$$

Letting $R \to \infty$, using the dominated convergence theorem and adding the limit in (2.30) results in

$$\liminf_{j\to\infty} \overline{\mathcal{F}}_\alpha(u_j) \leq \overline{\mathcal{F}}_\alpha(u).$$

Therefore, if we find for each $j \in \mathbb{N}$ a recovery sequence for u_j, then we can conclude the result using a diagonal argument; here, the coercivity of f is important to be able to extract convergent diagonal sequences. We can restrict to the case $u \in g + C_c^\infty(\Omega; \mathbb{R}^n)$ from now on.

The remaining argument is an adaptation of [140, Theorem 1.2] to the linear growth setting. To prove the upper bound when $u \in g + C_c^\infty(\Omega; \mathbb{R}^m)$, we take a Lipschitz domain $O \Subset \Omega$ and apply Proposition 2.3.2 (i) to find a $v \in W^{1,1}(O; \mathbb{R}^m)$ such that

$$\nabla v = \nabla^\alpha u \ \text{ on } O. \tag{2.31}$$

Then, we apply a classical relaxation theorem [75, Theorem 9.8] to find a sequence $(v_k)_k \subset W^{1,1}(O; \mathbb{R}^m)$ with the same trace values as v on the boundary $\partial\Omega$ such that $v_k \to v$ in $L^1(O; \mathbb{R}^m)$ and

$$\lim_{k\to\infty} \int_O f(x, \nabla v_k)\,dx = \int_O f^{\mathrm{qc}}(x, \nabla v)\,dx. \tag{2.32}$$

In view of the coercivity of f we may also suppose that $v_k \overset{*}{\rightharpoonup} v$ in $BV(O; \mathbb{R}^m)$. Now define the auxiliary sequence $(\tilde{v}_k)_k \subset W^{1,1}(\mathbb{R}^n; \mathbb{R}^m)$ via $\tilde{v}_k := v_k - v$ on O and $\tilde{v}_k = 0$ in O^c. By Proposition 2.3.2 (ii) and [140, Eq. (3.3)], we find that the sequence $(\tilde{u}_k)_k \subset S^{\alpha,1}(\mathbb{R}^n; \mathbb{R}^m)$ defined by $\tilde{u}_k = (-\Delta)^{\frac{1-\alpha}{2}} \tilde{v}_k$ satisfies

$$\tilde{u}_k \to 0 \ \text{ in } L^1(\mathbb{R}^n; \mathbb{R}^m) \text{ as } k \to \infty \tag{2.33}$$

and its fractional gradients are given by

$$\nabla^\alpha \tilde{u}_k = \nabla(v_k - v) \ \text{ in } O \quad \text{and} \quad \nabla^\alpha \tilde{u}_k = 0 \ \text{ in } O^c. \tag{2.34}$$

Take a cut-off function $\chi \in C_c^\infty(\Omega)$ such that $0 \leq \chi \leq 1$ and $\chi|_O \equiv 1$. Then, we define the sequence $(w_k)_k \subset S_g^{\alpha,1}(\Omega; \mathbb{R}^m)$ by

$$w_k = u + \chi \tilde{u}_k \overset{*}{\rightharpoonup} u \ \text{ in } BV_g^\alpha(\Omega; \mathbb{R}^m) \text{ as } k \to \infty,$$

where the convergence follows from (2.33) and the Leibniz rule (Lemma 2.3.3). Moreover, we have by (2.17) the convergence of the residuals

$$R_k := \nabla^\alpha w_k - \nabla^\alpha u - \chi \nabla^\alpha \tilde{u}_k \to 0 \ \text{ in } L^1(\mathbb{R}^n; \mathbb{R}^{m\times n}). \tag{2.35}$$

In O, we find in view of (2.31) and (2.34) that $\nabla^\alpha w_k = \nabla v_k + R_k$. Due to the strong convergence of R_k to zero it follows by testing with the Lipschitz basis from [144, Lemma 3] that the sequences

$(\nabla v_k)_k$ and $(\nabla^\alpha w_k)_k$ when restricted to O generate (up to a non-relabeled subsequence) the same generalized Young measure $\nu \in Y(O;\mathbb{R}^{m\times n})$. As a result, we use Theorem 2.2.2 twice to conclude

$$\begin{aligned}\liminf_{k\to\infty} \int_O f(x,\nabla^\alpha w_k)\,dx &\leq \int_O \langle f(x,\cdot),\nu_x\rangle\,dx + \int_{\overline{O}} \langle f^\infty(x,\cdot),\nu_x^\infty\rangle\,d\lambda_\nu \\ &= \lim_{k\to\infty}\int_O f(x,\nabla v_k)\,dx = \int_O f^{\mathrm{qc}}(x,\nabla v)\,dx = \int_O f^{\mathrm{qc}}(x,\nabla^\alpha u)\,dx,\end{aligned} \tag{2.36}$$

where we use (2.32) and (2.31) in the final two equalities. Additionally, in O^c we have that $\nabla^\alpha w_k = \nabla^\alpha u + R_k$ thanks to (2.34). Hence, we find using (2.35) and (G) that

$$\limsup_{k\to\infty}\int_{\Omega\setminus O} f(x,\nabla^\alpha w_k)\,dx = \limsup_{k\to\infty}\int_{\Omega\setminus O} f(x,\nabla^\alpha u + R_k)\,dx \leq \|M|\nabla^\alpha u| + a\|_{L^1(\Omega\setminus O)}. \tag{2.37}$$

Finally, using Lebesgue's dominated convergence theorem and (2.35) we derive

$$\lim_{k\to\infty}\int_{\Omega^c} f(x,\nabla^\alpha w_k)\,dx = \lim_{k\to\infty}\int_{\Omega^c} f(x,\nabla^\alpha u + R_k)\,dx = \int_{\Omega^c} f(x,\nabla^\alpha u)\,dx. \tag{2.38}$$

Summing (2.36), (2.37) and (2.38) together, we obtain

$$\liminf_{k\to\infty}\int_{\mathbb{R}^n} f(x,\nabla^\alpha w_k)\,dx \leq \int_O f^{\mathrm{qc}}(x,\nabla^\alpha u)\,dx + \int_{\Omega^c} f(x,\nabla^\alpha u)\,dx + \|M|\nabla^\alpha u| + a\|_{L^1(\Omega\setminus O)},$$

which yields the result if we let $O \uparrow \Omega$ and extract a diagonal sequence. □

Remark 2.5.1. a) Because of the coercivity condition of f, the functional $\mathcal{F}_\alpha^{\mathrm{rel}}$ is in particular weak* lower semicontinuous on $BV_g^\alpha(\Omega;\mathbb{R}^m)$. Interestingly, this fact does not immediately follow from the lower semicontinuity result in Theorem 2.4.1, since the strong recession function of $\mathbb{1}_\Omega f^{\mathrm{qc}} + \mathbb{1}_{\Omega^c} f$ need not exist. An application of the direct method as in Corollary 2.4.2 provides the existence of minimizers of $\mathcal{F}_\alpha^{\mathrm{rel}}$.

b) A simple argument using the theory of Young measures shows that the functional

$$\overline{\mathcal{F}}_\alpha(u) = \int_{\mathbb{R}^n} f(x,\nabla^\alpha u)\,dx + \int_{\overline{\Omega}} f^\infty\left(x,\frac{dD_s^\alpha u}{|dD_s^\alpha u|}\right)\,d|D_s^\alpha u| \quad \text{for } u \in BV_g^\alpha(\Omega;\mathbb{R}^m),$$

is the area-strictly continuous extension of $\mathcal{F}_\alpha$ to $BV_g^\alpha(\Omega;\mathbb{R}^m)$. This immediately implies that this functional is also the relaxation of $\mathcal{F}_\alpha$ if $f(x,\cdot)$ is quasiconvex for a.e. $x \in \Omega$, given the lower semicontinuity result from Theorem 2.4.1 and the density with respect to area-strict convergence.

c) The requirement (2.7) on f^{qc} is needed for the application of the Jensen's inequalities in the lower bound and allows the relaxation result to be phrased for general Carathéodory integrands. However, one can dispose of this assumption if we assume a continuity condition similarly as in [17, Theorem 1.7], that is,

$$|f(x,A) - f(y,A)| \leq \omega(|x-y|)(1+|A|) \quad \text{for all } x,y \in \overline{\Omega} \text{ and } A \in \mathbb{R}^{m\times n},$$

where $\omega : [0,\infty) \to [0,\infty)$ is a continuous and increasing function with $\omega(0) = 0$. Indeed, one can utilize (G) and (C) as in [182, Theorem 7.6] to deduce that f^{qc} inherits the same continuity condition (up to a different modulus of continuity), after which (2.7) readily follows. △

Chapter 3

A variational theory for integral functionals involving finite-horizon fractional gradients

This chapter agrees with

[72] J. Cueto, C. Kreisbeck and H. Schönberger. A variational theory for integral functionals involving finite-horizon fractional gradients. *Fractional Calculus and Applied Analysis*, 26(5):2001–2056, 2023. https://doi.org/10.1007/s13540-023-00196-7.

3.1 Introduction

Nonlocality has long been a recurring theme in the calculus of variations, appearing in various facets and applications. When modeling phenomena in nature and technology, nonlocal operators, whose values result from integrating over a neighborhood, have become a popular alternative to differential operators. A main advantage of this derivative-free approach is that it allows functions to be less regular and, therefore, makes it possible to capture discontinuity effects, and also long-range interactions are naturally included. In the context of mechanics, this is exploited in peridynamic modeling [158, 196] or to cover fracture and cavitation of deformed elastic materials [28, 30]. From the analytical viewpoint, dealing with nonlocality brings along new mathematical challenges, since it is intrinsically opposed to the standard techniques for classical variational problems. And yet, local and nonlocal problems can be closely intertwined: while localization causes nonlocal features to vanish [35, 158, 161], they can, on the other hand, arise from local ones e.g., through limit processes such as homogenization and disrete-to-continuum passages [39, 49].

In a recent series of works, different authors have studied problems involving integral functionals that depend instead of usual gradients on fractional-order ones through the Riesz fractional gradients [28, 140, 193, 194]. Even though the latter had appeared in the literature before [137], Shieh & Spector brought it back into the spotlight in [193, 194] and discussed properties of the associated fractional Sobolev spaces, which are equivalent to the Bessel potential spaces, see also [28, 54, 66, 140]. In contrast to the standard fractional Sobolev spaces defined via Gagliardo semi-norms, these spaces have a distributional character, and are, therefore, particularly well-suited for variational problems. Another asset is that the Riesz fractional gradient enjoys a unique combination of desirable homogeneity and invariance properties as shown by Šilhavý in [208], which makes it the natural choice of a fractional derivative among operators with infinite interaction range. Motivated by mechanical models of hyperelastic materials, which call for operators on bounded domains with finite interaction, Bellido, Cueto & Mora-Corral [31] recently proposed to consider nonlocal operators that result from the Riesz fractional gradient by truncation with a suitable cut-off function. This is the same setting we are adopting in the following.

Overall, this paper deals with variational integrals in the truncated framework of [31], for which we contribute new insights into the existence theory of minimizers as well as their asymptotic analysis. More precisely, the set-up is as follows: Let $\Omega \subset \mathbb{R}^n$ be a bounded open set, $s \in (0,1)$ the fractional-order parameter, $\delta > 0$ the horizon, which stipulates the maximal length scale of the interaction distance between points, and $\Omega_\delta = \Omega + B(0,\delta)$ the nonlocal closure of Ω.

We consider functionals of the form

$$\mathcal{F}(u) = \int_\Omega f\big(x, u(x), D^s_\delta u(x)\big)\, dx, \tag{3.1}$$

where the integrand function $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is Carathéodory with standard p-growth and p-coercivity for some $1 < p < \infty$ and $D^s_\delta u$ is the truncated Riesz fractional gradient (see (3.3) below) for functions u in a suitable linear subspace of $L^p(\Omega_\delta;\mathbb{R}^m)$. This function space, which is called $H^{s,p,\delta}(\Omega;\mathbb{R}^m)$ and introduced in Definition 3.2.7, is defined in analogy to the classical Sobolev spaces by requiring that the nonlocal gradient is p-integrable. In addition, we assume volumetric-type boundary conditions by prescribing complementary values in a tubular neighborhood or collar of radius 2δ around Ω; in the basic case of zero complementary values, we write $H^{s,p,\delta}_0(\Omega;\mathbb{R}^m)$ for the set of functions admissible for (3.1).

It remains to specify the nonlocal gradient $D^s_\delta u$. With $\mathcal{G}_\rho$ a general nonlocal gradient with kernel ρ, that is,

$$\mathcal{G}_\rho u(x) = \int_{\mathbb{R}^n} \frac{u(x)-u(y)}{|x-y|}\frac{x-y}{|x-y|}\rho(x-y)\, dy, \tag{3.2}$$

whenever the integral exists for a function $u : \mathbb{R}^n \to \mathbb{R}$, we first recall that the Riesz fractional gradient is defined as the nonlocal gradient with the Riesz potential kernel I_{1-s}, i.e.,

$$D^s u \propto \mathcal{G}_{I_{1-s}} u \quad \text{with} \quad I_{1-s} \propto \frac{1}{|\cdot|^{n+s-1}}.$$

To introduce the truncated version, let us consider a certain smooth, radial cut-off function $w_\delta : \mathbb{R}^n \to [0,\infty)$ supported in a ball of radius δ around the origin. Then,

$$D^s_\delta u = \mathcal{G}_{\rho^s_\delta} u \quad \text{with} \quad \rho^s_\delta \propto w_\delta I_{1-s}. \tag{3.3}$$

Throughout the paper, we refer to D^s_δ simply as nonlocal gradient to keep the terminology short. For more details on these definitions of nonlocal and fractional gradients, we refer the reader to Section 3.2.2. Alternative choices for the kernel function in (3.2) can be found in the literature, for example, kernels defined on half-balls [129, 148], and variable horizon kernels [103, 198, 201].

Our methodology for proving the results about the functionals (3.1) builds substantially on their relation with classical functionals with a dependence on the usual gradient, namely

$$v \mapsto \int_\Omega f(x, v(x), \nabla v(x))\, dx, \tag{3.4}$$

and also the relation with the fractional variational integrals

$$u \mapsto \int_{\mathbb{R}^n} f(x, u(x), D^s u(x))\, dx \tag{3.5}$$

provides useful insights. To set a foundation for a comparison of $\mathcal{F}$ with (3.4) and (3.5), we discuss the connection between the three differential operators

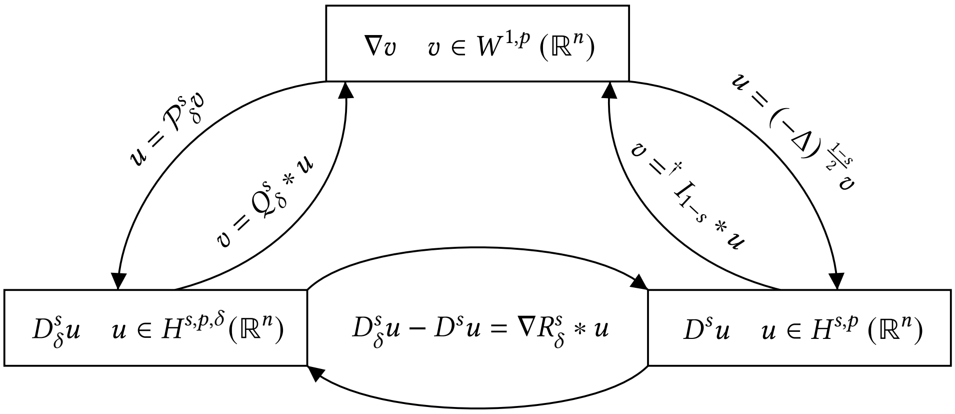

Figure 3.1: Illustration of the relations between classical, fractional, and nonlocal gradients, which enable the transfer of results between the corresponding settings. † When $I_{1-s} * u$ is well-defined.

$$\text{classical gradient } \nabla, \text{ fractional gradient } D^s, \text{ nonlocal gradient } D^s_\delta,$$

and the associated Sobolev-type function spaces

$$W^{1,p}(\mathbb{R}^n),\ H^{s,p}(\mathbb{R}^n),\ H^{s,p,\delta}(\mathbb{R}^n),$$

respectively; for an illustrative overview, see Figure 3.1.

Fractional vs. classical: For smooth compactly supported functions $\varphi \in C_c^\infty(\mathbb{R}^n)$, it is by now well-known that

$$D^s\varphi = \nabla(I_{1-s} * \varphi) \quad \text{and} \quad \nabla\varphi = D^s(-\Delta)^{\frac{1-s}{2}}\varphi, \tag{3.6}$$

where I_{1-s} is the Riesz potential and $(-\Delta)^{\frac{1-s}{2}}$ is the fractional Laplacian of order $1-s$, see e.g., [193, 208]. In [140, Proposition 3.1], two of the authors extended these identities to the setting of Sobolev and fractional Sobolev functions, showing that for any $u \in H^{s,p}(\mathbb{R}^n)$, there exists a $v \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$ such that $\nabla v = D^s u$, and for every $u \in H^{s,p}(\mathbb{R}^n)$, one can find a $v \in W^{1,p}(\mathbb{R}^n)$ with $D^s u = \nabla v$. The latter follows immediately from the observation that $(-\Delta)^{\frac{1-s}{2}} : W^{1,p}(\mathbb{R}^n) \to H^{s,p}(\mathbb{R}^n)$ is a bounded linear operator. This way, one can translate from the fractional gradient to the classical one and vice-versa, up to a gap related to an issue of local integrability. For a similar statement in the space of fractional BV-functions, we refer to [66, Lemma 3.28].

Nonlocal vs. classical: Providing analogous translation formulas between the nonlocal and classical setting is one of the major steps in the analysis of this paper. The fact that $D^s_\delta u$ is defined over a bounded domain brings about some technical complications compared with $D^s u$; for instance, as opposed to $D^s u$, the operator $D^s_\delta u$ is no longer homogeneous and it does not enjoy a semigroup property, which the fractional one inherits from its relationship with the Riesz potential. The foundations for finding a suitable replacement for the generalization of (3.6), were laid by Bellido, Cueto & Mora-Corral [31] (see also [30]). They identified an integrable finite-horizon counterpart of the Riesz potential kernel, called Q^s_δ, which provides one of the directions of the translation mechanism for smooth functions. For the other direction, we heuristically invert the convolution with Q^s_δ in Fourier space, i.e., we consider the operator

$$\mathcal{P}^s_\delta\varphi = \left(\frac{\widehat{\varphi}}{\widehat{Q^s_\delta}}\right)^\vee$$

for any Schwartz function φ. This operator can be considered as an analogue of the fractional Laplacian of order $(1-s)/2$ in the nonlocal framework. Another way of interpreting $\mathcal{P}_\delta^s\varphi$ is as the convolution of the gradient of φ with the kernel from the nonlocal fundamental theorem of calculus in [31, Theorem 4.5], see Remark 3.2.14 d).

Here, we prove that the convolution with Q_δ^s and $\mathcal{P}_\delta^s$ can both be extended to the Sobolev spaces in such a way that they are each other's inverses. This gives a perfect isomorphism between $H^{s,p,\delta}(\mathbb{R}^n)$ and $W^{1,p}(\mathbb{R}^n)$ with the property that for any $u \in H^{s,p,\delta}(\mathbb{R}^n)$ and $v \in W^{1,p}(\mathbb{R}^n)$

$$D_\delta^s u = \nabla(Q_\delta^s * u) \qquad \text{and} \qquad \nabla v = D_\delta^s \mathcal{P}_\delta^s v, \tag{3.7}$$

see Theorem 3.2.13 and the discussion thereafter. It is noteworthy that in the fractional case there is no such isomorphism, since the Riesz potential is only locally integrable as opposed to Q_δ^s.

Fractional vs. nonlocal: A comparison between the kernels Q_δ^s and I_{1-s}, where R_δ^s denotes their difference, gives us a basic and direct way for switching between the fractional and nonlocal setting. Indeed, we show in Section 3.2.5, that

$$D_\delta^s u = D^s u + \nabla R_\delta^s * u, \tag{3.8}$$

for all $u \in H^{s,p,\delta}(\mathbb{R}^n) = H^{s,p}(\mathbb{R}^n)$, where $\nabla R_\delta^s \in L^1(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n)$.

Having the translation mechanism of (3.7) and (3.8) at hand paves the way for shifting results between the three variational settings. Note, however, that not all results can be directly carried over, since boundary conditions are not preserved in the translation procedure and problems involving both the function and its nonlocal gradient require additional techniques. Here, we list and discuss the main contributions of this paper to the existence and asymptotic analysis of the functionals $\mathcal{F}$ in (3.1):

(1) Characterization of weak lower semicontinuity of $\mathcal{F}$. One of the crucial steps to conclude the existence of minimizers of integral functionals, like $\mathcal{F}$ or those in (3.4) or (3.5), via the direct method, is to establish weak lower semicontinuity. A well-known fundamental result from the vectorial calculus of variations with roots in the 1950s states that, for the functionals (3.4), quasiconvexity (in the sense of Morrey) regarding the third variable of f is necessary and sufficient for weak lower semicontinuity in $W^{1,p}(\Omega;\mathbb{R}^m)$, see [3, 155, 162, 165]. In the fractional setting (3.5), the efforts are more recent. After convexity [193] and polyconvexity [28] had been identified as sufficient conditions for weak lower semicontinuity in $H_0^{s,p}(\Omega;\mathbb{R}^m)$, the problem of characterization was solved in [140, Theorem 1.1]. Interestingly, the correct condition on f is the same as in the local case, namely quasiconvexity.

We complement the picture in Theorem 3.4.1, by proving that, altogether, quasiconvexity is the intrinsic convexity notion in all three situations. In fact,

$$\begin{aligned}\mathcal{F} \text{ is weakly lower semicontinuous in } H_0^{s,p,\delta}(\Omega;\mathbb{R}^m) \text{ if and only if}\\ f(x,z,\cdot) \text{ is quasiconvex for a.e. } x \in \Omega_{-\delta} \text{ and all } z \in \mathbb{R}^m;\end{aligned} \tag{3.9}$$

note that, due to a boundary layer effect, which yields even strong L^p-convergence of weakly convergent sequences in $H_0^{s,p,\delta}(\Omega;\mathbb{R}^m)$, quasiconvexity is not required in the collar. Moreover, we introduce a nonlocal notion of quasiconvexity defined through testing with nonlocal gradients that turns out to be equivalent with quasiconvexity, cf. Remark 3.4.3.

The proof of (3.9) exploits the parallels between the nonlocal and fractional gradient in their relation to the standard one (cf. (3.7) and (3.6)) by using similar arguments and techniques as in [140]. An alternative proof strategy that reduces (3.9) directly to the statement of [140, Theorem 1.1] via (3.8) is also possible, as we demonstrate under simplified assumptions.

(2) Variational convergence, homogenization and relaxation. Considering sequences of nonlocal functionals $\{\mathcal{F}_{f_j}\}_{j\in\mathbb{N}}$ as in (3.1) with specific integrand functions f_j, we study their asymptotic behavior as $j \to \infty$. The intention of finding a versatile method that makes Γ-convergence (see [49, 80]) accessible to a number of cases and applications motivates the statement of Theorem 3.5.1. If we denote the counterparts of $\mathcal{F}_{f_j}$ with dependence on classical gradients defined on $W^{1,p}(\Omega_{-\delta};\mathbb{R}^m)$ by $\mathcal{I}_{f_j}$, it says that the convergence of $\{\mathcal{I}_{f_j}\}_{j\in\mathbb{N}}$ to a Γ-limit $\mathcal{I}_{f_\infty}$ as $j \to \infty$ along with the pointwise convergence of the integrals over the collar, $L^p(\Omega_\delta;\mathbb{R}^{m\times n}) \ni V \mapsto \int_{\Omega\setminus\Omega_{-\delta}} f_j(x,V)\,dx$ yields

$$\Gamma\text{-}\lim_{j\to\infty} \mathcal{F}_{f_j} = \mathcal{F}_{f_\infty};$$

note that all Γ-limits are taken with respect to the strong L^p-topology.

To demonstrate how this observation can help to carry various Γ-convergence results in the literature from the local to the nonlocal setting, we choose homogenization theory as a specific case. Indeed, Corollary 3.5.2 shows that the fundamental Γ-limit of [48, 168], where the homogenized functional is again of integral form with integrand determined by a multi-cell formula, gives rise to a new homogenization limit for problems involving nonlocal gradients. As an immediate consequence of this homogenization, one can obtain relaxation of nonlocal functions $\mathcal{F}$, that is, a representation for their lower semicontinuous envelopes. In the case of a homogeneous integrand f, the latter arises from the quasiconvexification of f on $\Omega_{-\delta}$, while f remains unchanged in $\Omega \setminus \Omega_{-\delta}$, see Corollary 3.5.3.

(3) Asymptotics for varying fractional order and localization. It is a natural question to investigate the dependence of our nonlocal variational problems, in particular, their minimizers and minima, on the fractional order $s \in (0,1)$; for an analogous study for functionals of the type (3.5), see [29]. To this end, we take functionals as in (3.1), with f independent of the second variable and quasiconvex in the third one, and highlight the dependence of s with a subscript index $\mathcal{F}_s$. The functional $\mathcal{F}_1$ can be defined in the same way with $D^1_\delta u := \nabla u$ the classical gradient and $\mathcal{F}_0$, after extension of the definition in (3.3) to $s = 0$, lives on $L^p(\Omega_\delta;\mathbb{R}^m)$.

The main result in this context is Theorem 3.5.1, which says the following:

$$\text{The sequence } \{\mathcal{F}_s\}_s \ \Gamma\text{-converges to } \mathcal{F}_{s'} \text{ as } s \to s' \in [0,1];$$

since sequential compactness of bounded-energy sequences holds strongly in $L^p(\Omega;\mathbb{R}^m)$ when $s' \in (0,1]$ and weakly in $L^p(\Omega;\mathbb{R}^m)$ if $s' = 0$, it is natural to state the Γ-convergence results regarding the strong and weak topology, respectively, see Lemma 3.3.9. We point out that the limit $s \to 1$ provides a localization statement, and as such, establishes another interesting connection between classical local and nonlocal theories.

The proof of the above-mentioned compactness for bounded-energy sequences in nonlocal spaces of different order involves, besides the continuous dependence of the nonlocal gradient $D^s_\delta u$ on s (see Lemma 3.3.2), also a new technical tool that is worth mentioning in its own right. This is the nonlocal Poincaré inequality with a constant independent of the fractional order presented in Theorem 3.3.8; we refer to recent progress on nonlocal Poincaré-type inequalities, for example, in problems involving radial kernels [31, 104] or asymmetric and inhomogeneous kernels [114, 129]. The difficulty in establishing a parameter-independent bound is the fact that the kernel in the nonlocal fundamental theorem of calculus from [31, Theorem 4.5] is implicitly defined via a Fourier transform, which makes it hard to isolate the dependence on s in the proof of the Poincaré inequality from [31, Theorem 6.2]. Instead, we utilize a fine analysis of the decay of the Fourier transform of Q^s_δ, an application of the Mihlin-Hörmander multiplier theorem and an extension of the nonlocal fundamental theorem to the case $s = 0$ (see Proposition 3.2.9) to prove the Poincaré inequality with an s-independent constant.

Note that besides the above localization result for $s \to 1$, a different type of localization could be obtained in the limit of vanishing horizon, i.e., for $\delta \to 0$. Indeed, such a Γ-convergence statement for integral functionals depending on a class of closely related nonlocal gradients is already proven in [161], yielding a classical local model in the limit. However, it remains an interesting open problem for the future to prove the required equi-compactness in order to deduce the convergence of minimizers. For readers interested in localization results in the context of other nonlocal variational problems there is a broad literature available, we refer e.g. to [8, 35, 47], which discuss double-integrals that depend on difference quotients, convolution-type integrals, and functionals arising from models in peridynamics, respectively.

This manuscript is organized as follows. We begin in Section 3.2 with notations and a detailed introduction to our set-up and nonlocal calculus. Moreover, we collect and establish the relevant technical tools, especially, the connections between classical, nonlocal and fractional gradients along with the corresponding translation keys. Section 3.3 deals then with the asymptotics of the nonlocal gradient, and we derive as a main application a Poincaré inequality with a constant uniform in s, which opens the way for compactness results for sequences in nonlocal spaces of different order. The variational results for the nonlocal integral functionals are proven from Section 3.4 onwards, based on the comparison with the classical and fractional setting. First, we prove the characterization of weak lower semicontinuity in terms of quasiconvexity of the integrand and state an existence statement for minimizers of $\mathcal{F}$ (see Corollary 3.4.4) based on it. In Section 3.5, we then provide a general Γ-convergence result, from which homogenization and relaxation can be deduced as corollaries. Finally, we prove the convergence of minimizers of the functionals $\{\mathcal{F}_s\}_s$ for the limit $s \to s' \in [0, 1]$ in Section 3.6, showing, in particular, the localization to a classical local limit as $s \to 1$.

3.2 Preliminaries and technical tools

The aim of this section is to introduce the notation and several important definitions and tools regarding the nonlocal gradient and Sobolev spaces.

3.2.1 Notation

General notation

Unless mentioned otherwise, $s \in (0, 1)$ and $Y = (0, 1)^n \subset \mathbb{R}^n$. We use $\mathbb{R}_\infty$ to denote $\mathbb{R} \cup \{\infty\}$. We write $|x| = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$ for the Euclidean norm of a vector $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ and similarly, $|A|$ for the Frobenius norm of a matrix $A \in \mathbb{R}^{m\times n}$. The ball centered at $x \in \mathbb{R}^n$ and with radius $\rho > 0$ is denoted by $B(x, \rho) = \{y \in \mathbb{R}^n : |x - y| < \rho\}$ and the distance between $x \in \mathbb{R}^n$ and a set $E \subset \mathbb{R}^n$ is written as $d(x, E)$. For an open set $\Omega \subset \mathbb{R}^n$ and $\delta > 0$, we write Ω_δ for its nonlocal closure, that is,

$$\Omega_\delta = \Omega + B(0, \delta) = \{x \in \mathbb{R}^n \,:\, d(x, \Omega) < \delta\}.$$

The complement of a set $E \subset \mathbb{R}^n$ is indicated by $E^c := \mathbb{R}^n \setminus E$ and its closure by $\overline{E}$. The notation $E \Subset F$ for sets $E, F \subset \mathbb{R}^n$ means that E is compactly contained in F, i.e., $\overline{E} \subset F$ and $\overline{E}$ is compact. Let

$$\mathbb{1}_E(x) = \begin{cases} 1 & \text{for } x \in E, \\ 0 & \text{otherwise,} \end{cases} \qquad x \in \mathbb{R}^n,$$

be the indicator function of a set $E \subset \mathbb{R}^n$.

Let $U \subset \mathbb{R}^n$ be an open set. The notation $C_c^\infty(U)$ symbolizes the smooth functions $\varphi : U \to \mathbb{R}$ with compact support in $U \subset \mathbb{R}^n$. Our convention is that functions in $C_c^\infty(U)$ are identified with their trivial extension to $\mathbb{R}^n$ by zero. Further, by $C^\infty(\mathbb{R}^n)$, $C_0(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$ we denote the space of smooth functions, continuous functions vanishing at infinity and Schwartz functions on $\mathbb{R}^n$, respectively. We utilize multi-index notation, in particular, we write ∂^α for the partial derivative with respect to a multi-index $\alpha \in \mathbb{N}_0^n$.

By $\mathrm{Lip}_b(\mathbb{R}^n)$, we refer to all the functions $\psi : \mathbb{R}^n \to \mathbb{R}$ that are Lipschitz continuous and bounded on $\mathbb{R}^n$ and we write $\mathrm{Lip}(\psi)$ for the Lipschitz constant of ψ.

The Lebesgue measure of $U \subset \mathbb{R}^n$ is written $|U|$ and the convolution of two functions $u, v : \mathbb{R}^n \to \mathbb{R}$ is denoted by $u * v$. If one of the functions is vector-valued, the convolution should be understood componentwise. We use the common notation for Lebesgue- and Sobolev-spaces, that is, $L^p(U)$ for $p \in [1, \infty]$ is the space of p-real-valued integrable functions on U with the norm

$$\|u\|_{L^p(U)} = \begin{cases} \left(\int_U |u(x)|\, dx\right)^{1/p} & \text{if } p \in [1,\infty), \\ \operatorname{ess\,sup}_{x\in U} |u(x)| & \text{if } p = \infty, \end{cases} \qquad u \in L^p(U).$$

Moreover, $W^{1,p}(U)$ for $p \in [1, \infty]$ consists of all L^p-functions on U with p-integrable weak derivatives, endowed with the norm

$$\|u\|_{W^{1,p}(U)} = \|u\|_{L^p(U)} + \|\nabla u\|_{L^p(U;\mathbb{R}^n)};$$

here ∇u stands for the weak gradient of u.

The functions that lie locally in L^p and $W^{1,p}$ are denoted by $L^p_{\mathrm{loc}}(\mathbb{R}^n)$ and $W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$. Besides, $W^{1,p}_0(U)$ stands for those functions in $W^{1,p}(U)$ with zero boundary value in the sense of the trace and $W^{1,\infty}_\#(Y)$ indicates the Y-periodic functions in $W^{1,\infty}(\mathbb{R}^n)$.

In general, the spaces defined above can be extended componentwise to vector-valued functions. The target space is explicitly mentioned in the notation, like, for example, $L^p(U;\mathbb{R}^m)$. Whenever convenient, we identify a function on a subset of $\mathbb{R}^n$ with its trivial extension by zero. Finally, we use C to denote a generic constant, which may change from one estimate to the next without further mention. If we wish to indicate the dependence of C on certain quantities, we add them in brackets.

Riesz potential and Fourier transform

We recall the definition of Riesz potential. Given $0 < s < n$, the Riesz potential kernel $I_s : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$ is

$$I_s(x) = \gamma_{n,s}^{-1} \frac{1}{|x|^{n-s}}, \tag{3.10}$$

where

$$\gamma_{n,s} = \frac{\pi^{\frac{n}{2}}\, 2^s\, \Gamma(\frac{s}{2})}{\Gamma(\frac{n-s}{2})}$$

with Γ denoting the Gamma function. For notational convenience, we also define $I_1 : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ as

$$I_1(x) = -\frac{1}{\pi} \log(|x|) \tag{3.11}$$

when $n = 1$. The Riesz potential of a locally integrable function f is given via convolution as

$$I_s * f(x) = \frac{1}{\gamma_{n,s}} \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-s}} dy,$$

whenever the integral exists for a.e. $x \in \mathbb{R}^n$.

Since we will also deal with the use of the Fourier transform, we clarify here the notation we are going to use. For $f \in L^1(\mathbb{R}^n)$, we define the Fourier transform of f as

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i x\cdot\xi}\, dx \quad \xi \in \mathbb{R}^n.$$

Notice that this definition can also be used in the Schwartz space $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, where it defines an isomorphism. By continuity and duality extensions, it also defines isomorphism on the spaces $L^2(\mathbb{R}^n;\mathbb{C})$ and in the space of tempered distributions $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Moreover, the inverse Fourier transform is denoted by $f^\vee$ and corresponds with $x \mapsto \widehat{f}(-x)$. Notable references in Fourier analysis are [106, 122].

3.2.2 Nonlocal calculus and function spaces

In this section, we present the definition of the nonlocal gradient used throughout this paper, introduce the naturally associated function spaces, and collect several auxiliary results. A delicate issue is the choice of suitable boundary values, which is addressed below in Section 3.2.3.

In what follows, let $\delta > 0$ and $w_\delta : \mathbb{R}^n \to [0,\infty)$ be a non-negative cut-off function satisfying these hypotheses:

(H1) w_δ is radial, i.e., there is a $\overline{w}_\delta : \mathbb{R} \to [0,\infty)$ such that $w_\delta(x) = \overline{w}_\delta(|x|)$ for $x \in \mathbb{R}^n$;

(H2) w_δ is smooth and compactly supported in $B(0,\delta)$, i.e., $w_\delta \in C_c^\infty(B(0,\delta))$;

(H3) there is a constant $b_0 \in (0,1)$ such that $w_\delta = 1$ on $B(0, b_0\delta)$;

(H4) w_δ is radially decreasing, that is, $w_\delta(x) \geq w_\delta(y)$ if $|x| \leq |y|$.

In accordance with [31, Definition 3.1], we define the nonlocal gradient and divergence for smooth functions as follows: For $s \in [0,1)$, the nonlocal gradient of $\varphi \in C^\infty(\mathbb{R}^n)$ is given by

$$D_\delta^s \varphi(x) = c_{n,s} \int_{\mathbb{R}^n} \frac{\varphi(x)-\varphi(y)}{|x-y|} \frac{x-y}{|x-y|} \frac{w_\delta(x-y)}{|x-y|^{n+s-1}}\, dy \quad \text{for } x \in \mathbb{R}^n, \tag{3.12}$$

and the nonlocal divergence of $\psi \in C^\infty(\mathbb{R}^n;\mathbb{R}^n)$ is

$$\operatorname{div}_\delta^s \psi(x) = c_{n,s} \int_{\mathbb{R}^n} \frac{\psi(x)-\psi(y)}{|x-y|} \cdot \frac{x-y}{|x-y|} \frac{w_\delta(x-y)}{|x-y|^{n+s-1}}\, dy \quad \text{for } x \in \mathbb{R}^n, \tag{3.13}$$

with the scaling constant

$$c_{n,s} := \frac{\Gamma\left(\frac{n+s+1}{2}\right)}{\pi^{n/2} 2^{-s} \Gamma\left(\frac{1-s}{2}\right)}.$$

Note that the integral in (3.12) is absolutely convergent given that φ is in particular locally Lipschitz continuous and $w_\delta(\cdot)/|\cdot|^{n+s-1} \in L^1(\mathbb{R}^n)$ with compact support. Moreover, the above definitions show that $\operatorname{supp}(D_\delta^s\varphi) \subset \operatorname{supp}(\varphi) + \overline{B(0,\delta)}$ and Proposition 3.2.2 below establishes $D_\delta^s\varphi \in C^\infty(\mathbb{R}^n;\mathbb{R}^n)$. Analogous observations hold for the nonlocal divergence.

Remark 3.2.1. a) Due to the radial symmetry of w_δ from (H1), an equivalent way of expressing $D_\delta^s\varphi$ for $\varphi \in C^\infty(\mathbb{R}^n)$ is as

$$D_\delta^s \varphi(x) = \lim_{r\downarrow 0} \int_{B(x,r)^c} \varphi(y) d_\delta^s(x-y)\, dy \quad \text{for } x \in \mathbb{R}^n, \tag{3.14}$$

with

$$d_\delta^s(x) = -c_{n,s} \frac{x w_\delta(x)}{|x|^{n+s+1}} \quad \text{for } x \in \mathbb{R}^n \setminus \{0\}. \tag{3.15}$$

When $x \notin \operatorname{supp}(\varphi)$, this allows us to write $D_\delta^s \varphi(x) = (d_\delta^s * \varphi)(x)$.

b) It is straightforward to check for the nonlocal gradient that it is translation and rotation invariant, i.e.,

$$D_\delta^s\big(\varphi(\cdot + b)\big) = D_\delta^s \varphi(\cdot + b) \quad \text{and} \quad D_\delta^s\big(\varphi(R\,\cdot)\big) = R^{-1} D_\delta^s \varphi(R\,\cdot)$$

for all $\varphi \in C^\infty(\mathbb{R}^n)$, $b \in \mathbb{R}^n$ and $R \in \mathrm{O}(n)$. The rotation invariance relies on the radiality of w_δ. If, in addition, $w_\delta(\cdot/\lambda) = w_{\lambda\delta}(\cdot)$ for all $\lambda > 0$, then D_δ^s is also positively s-homogeneous in the sense that

$$D_\delta^s\big(\varphi(\lambda\,\cdot)\big) = \lambda^s D_{\delta/\lambda}^s \varphi(\lambda\,\cdot)$$

for all $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\lambda > 0$.

To put this observation in context, we remark that Šilhavý in [208] identified the Riesz fractional gradient as the unique fractional derivative operator that is suitably continuous, rotation and translation invariant and s-homogeneous. Hence, one can view D_δ^s as a nonlocal derivative operator with finite interaction range that enjoys the same desirable properties. △

As recently shown in [31], the nonlocal gradient can be written as the convolution of a certain integrable kernel with the classical gradient. To formulate this result, which is in analogy to the representation of the Riesz fractional gradient as the Riesz potential of the usual gradient, we first introduce for $s \in [0, 1)$ the kernel

$$Q_\delta^s : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}, \quad Q_\delta^s(x) = c_{n,s} \int_{|x|}^{\delta} \frac{\overline{w}_\delta(t)}{t^{n+s}}\, dt. \tag{3.16}$$

Proposition 3.2.2. *Let $s \in [0, 1)$. It holds for every $\varphi \in C^\infty(\mathbb{R}^n)$ that*

$$D_\delta^s \varphi = Q_\delta^s * \nabla\varphi \ \in C^\infty(\mathbb{R}^n).$$

In particular, when $\varphi \in \mathcal{S}(\mathbb{R}^n)$ then $D_\delta^s \varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{R}^n)$.

Proof. The statement for $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $s \in (0, 1)$ is exactly [31, Proposition 4.3], and the case $s = 0$ is proven analogously. Since any $\varphi \in C^\infty(\mathbb{R}^n)$ locally coincides with a smooth function with compact support, the same holds for such functions. Finally, since $Q_\delta^s \in L^1(\mathbb{R}^n)$, the statement for Schwartz functions follows. □

Remark 3.2.3 (Properties of Q_δ^s). For easier referencing, we list here a few relevant properties of Q_δ^s for $s \in [0, 1)$ that will be used later in the paper. The details for $s \in (0, 1)$ can be found in [31, Lemma 4.2, Propositions 5.2 and 5.5], and the same arguments extend also to the case $s = 0$.

a) The kernel Q_δ^s lies in $L^1(\mathbb{R}^n)$ with $\operatorname{supp}(Q_\delta^s) \subset B(0, \delta)$ and is radially decreasing.

b) Since Q_δ^s has compact support, its Fourier transform is analytic and thus smooth. Moreover, $\widehat{Q_\delta^s}$ is bounded, radial, and strictly positive. △

The nonlocal gradient and divergence as defined in (3.12) and (3.13) act as dual operators in the sense of integration by parts. While several versions of nonlocal integration by parts for related fractional or nonlocal operators have been studied in the literature [66, 161, 208], we employ here the following formula, stated for smooth functions.

Lemma 3.2.4 (Nonlocal integration by parts formula). *Let $s \in [0,1)$ and suppose that $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\psi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$. Then,*

$$\int_{\mathbb{R}^n} D_\delta^s \varphi \cdot \psi \, dx = -\int_{\mathbb{R}^n} \varphi \operatorname{div}_\delta^s \psi \, dx.$$

Proof. According to Proposition 3.2.2, it holds that $D_\delta^s\varphi = Q_\delta^s * \nabla\varphi = \nabla(Q_\delta^s * \varphi) \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$ and similarly, $\operatorname{div}_\delta^s \psi = Q_\delta^s * \operatorname{div}\psi \in C_c^\infty(\mathbb{R}^n)$. Hence, we may calculate

$$\begin{aligned}\int_{\mathbb{R}^n} D_\delta^s\varphi \cdot \psi \, dx &= \int_{\mathbb{R}^n} \nabla(Q_\delta^s * \varphi)\cdot\psi \, dx \\ &= -\int_{\mathbb{R}^n} (Q_\delta^s * \varphi) \operatorname{div}\psi \, dx = -\int_{\mathbb{R}^n} \varphi\,(Q_\delta^s * \operatorname{div}\psi)\, dx = -\int_{\mathbb{R}^n} \varphi \operatorname{div}_\delta^s \psi \, dx;\end{aligned}$$

the second identity is due to classical integration by parts, while the third one follows via Fubini's theorem. □

In light of this integration by parts formula, the definition in (3.12) can be extended to a broader class of functions using a distributional approach. We will work with functions defined on an open set $\Omega \subset \mathbb{R}^n$. As nonlocal boundary of this set, we choose a volumetric type as is common in nonlocal models, considering a tubular neighborhood or collar of radius $\delta > 0$ around Ω. Precisely, $\Omega_\delta = \Omega + B(0,\delta)$ is the nonlocal closure of Ω and $\Omega_\delta \setminus \Omega$ plays the role of nonlocal boundary.

Definition 3.2.5 (Weak nonlocal gradient). *Let $s \in [0,1)$, $\delta > 0$, $\Omega \subset \mathbb{R}^n$ open and $u \in L^1_{\mathrm{loc}}(\Omega_\delta)$. We say that $v \in L^1_{\mathrm{loc}}(\Omega;\mathbb{R}^n)$ is the weak nonlocal gradient of u, written as $v = D_\delta^s u$, if*

$$\int_\Omega v \cdot \psi \, dx = -\int_{\Omega_\delta} u \operatorname{div}_\delta^s \psi \, dx \qquad \text{for all } \psi \in C_c^\infty(\Omega;\mathbb{R}^n).$$

Remark 3.2.6. In the case $s = 0$, it holds for each $\varphi \in C^\infty(\mathbb{R}^n)$ by (3.14) that

$$D_\delta^0\varphi(x) = \lim_{r\downarrow 0}\int_{B(x,r)^c} \varphi(y) d_\delta^0(x-y)\, dy \quad \text{for } x \in \mathbb{R}^n,$$

with d_δ^0 as in (3.15). The theory of singular integrals (see e.g., [122, Theorem 5.4.1]) implies that D_δ^0 can be uniquely extended to a continuous linear operator from $L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n;\mathbb{R}^n)$ when $p \in (1,\infty)$; indeed, one can easily verify that d_δ^0 satisfies the size and cancellation conditions [122, Eq. (5.4.1) and (5.4.3)], while the Hörmander condition [122, Eq. (5.4.2)] follows from the stronger property

$$|\nabla d_\delta^0| \le \frac{C}{|\cdot|^{n+1}},$$

which holds due to $\nabla w_\delta = 0$ in $B(0, b_0\delta)$.

We therefore find for each $u \in L^p(\Omega_\delta)$ (after extension to $\mathbb{R}^n$ by zero) that $D_\delta^0 u \in L^p(\Omega;\mathbb{R}^n)$ and

$$\|D_\delta^0 u\|_{L^p(\Omega;\mathbb{R}^n)} \le C\|u\|_{L^p(\Omega_\delta)}$$

with $C > 0$ a constant independent of u. Note that via a density argument, $D_\delta^0 u$ coincides with the weak nonlocal gradient from Definition 3.2.5. △

In analogy with the definition of the standard and fractional Sobolev spaces, it is now quite natural to consider the space of L^p-functions whose weak nonlocal gradient is also an L^p-function. Our default choice for the integrability parameters throughout the manuscript is $p \in (1,\infty)$. Especially when dealing with compactness results, weak convergence, and the coercivity of functionals, the reflexivity of the nonlocal Sobolev spaces is essential. For the sake of generality, however, we introduce the following spaces for the full range of integrability exponents $p \in [1, \infty]$ and extend results to the cases $p = 1$ and $p = \infty$ whenever this is possible at moderate technical expense.

Definition 3.2.7 (Nonlocal Sobolev spaces). *Let $s \in [0, 1)$, $p \in [1, \infty]$ and $\Omega \subset \mathbb{R}^n$ be open. We define the nonlocal Sobolev space $H^{s,p,\delta}(\Omega)$ as*

$$H^{s,p,\delta}(\Omega) := \{u \in L^p(\Omega_\delta) \ : \ D^s_\delta u \in L^p(\Omega; \mathbb{R}^n)\},$$

equipped with the norm

$$\|u\|_{H^{s,p,\delta}(\Omega)} = \left(\|u\|^p_{L^p(\Omega_\delta)} + \|D^s_\delta u\|^p_{L^p(\Omega;\mathbb{R}^n)}\right)^{\frac{1}{p}}.$$

The corresponding spaces of vector-valued functions $H^{s,p,\delta}(\Omega; \mathbb{R}^m)$ are defined componentwise.

In parallel with the classical Sobolev spaces, $H^{s,p,\delta}(\Omega)$ is a Banach space and, when $p \in (1, \infty)$, also reflexive. Moreover, a sequence $\{u_j\}_{j\in\mathbb{N}} \subset H^{s,p,\delta}(\Omega)$ converges weakly to u in $H^{s,p,\delta}(\Omega)$ for $p \in (1, \infty)$ if and only if $u_j \rightharpoonup u$ in $L^p(\Omega_\delta)$ and $D^s_\delta u_j \rightharpoonup D^s_\delta u$ in $L^p(\Omega; \mathbb{R}^n)$ as $j \to \infty$. In view of Remark 3.2.6, it holds that $H^{0,p,\delta}(\Omega) = L^p(\Omega_\delta)$ for $p \in (1, \infty)$ with an equivalent norm. Additionally, we set

$$H^{1,p,\delta}(\mathbb{R}^n) := W^{1,p}(\mathbb{R}^n) \quad \text{with } D^1_\delta u := \nabla u \text{ for } u \in H^{1,p,\delta}(\mathbb{R}^n), \tag{3.17}$$

which provides a consistent notation for the range of fractional orders $s \in [0, 1]$.

When we consider the whole space, i.e., $\Omega = \mathbb{R}^n$, and $s \in (0, 1)$, then by Lemma 3.2.16 the nonlocal Sobolev spaces of Definition 3.2.7 correspond to the fractional Sobolev spaces $H^{s,p}(\mathbb{R}^n)$ consisting of L^p-functions with weak fractional gradient in L^p, which are known to be equivalent to the Bessel potential spaces for $p \in (1, \infty)$ [54, 66, 140, 193]; in formulas,

$$H^{s,p,\delta}(\mathbb{R}^n) = H^{s,p}(\mathbb{R}^n).$$

We point out that for $s \in (0, 1)$ and $p \in [1, \infty)$, the Definition 3.2.7 is different from how nonlocal Sobolev spaces are introduced in [31, Definition 3.3], where the authors use the closure of $C^\infty_c(\mathbb{R}^n)$ functions under the norm in (3.2.7). However, both definitions are equivalent for Lipschitz domains as the following density result shows. It corresponds to a nonlocal version of the Meyers-Serrin theorem for classical Sobolev spaces, and the proof, which is based on approximate extension, can be found in Appendix 3.B.

Theorem 3.2.8. *Let $s \in [0, 1)$, $p \in [1, \infty)$ and let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain or $\Omega = \mathbb{R}^n$. Then, for every $u \in H^{s,p,\delta}(\Omega)$, there exists a sequence $\{\varphi_j\}_{j\in\mathbb{N}} \subset C^\infty_c(\mathbb{R}^n)$ that converges (when restricted to Ω_δ) to u in $H^{s,p,\delta}(\Omega)$.*

An important ingredient for the analysis of the nonlocal gradient are suitable versions of the fundamental theorem of calculus (FTC). For the case $s \in (0, 1)$, this has been proven in [31, Proposition 4.4]. Now we generalize it to $s = 0$ as well, which is needed to obtain a nonlocal Poincaré inequality independent of the fractional parameter. The proof takes inspiration from the arguments in [31, Appendix].

Proposition 3.2.9 (Nonlocal FTC for $s = 0$). *There is a function $W_\delta \in L^\infty(\mathbb{R}^n; \mathbb{R}^n)$ such that every $\varphi \in \mathcal{S}(\mathbb{R}^n)$ can be expressed as*

$$\varphi = -RD^0_\delta \varphi + W_\delta * D^0_\delta \varphi,$$

where $\widehat{R\psi}(\xi) = \frac{i\xi\cdot\widehat{\psi}(\xi)}{|\xi|}$ denotes the Riesz transform of $\psi \in \mathcal{S}(\mathbb{R}^n; \mathbb{R}^n)$ and $W_\delta * D^0_\delta \varphi$ is the sum of the componentwise convolutions of W_δ and $D^0_\delta \varphi$.

Proof. Consider the tempered distribution $Z_\delta \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C}^n)$, given by

$$\langle Z_\delta, \eta\rangle = \lim_{r\downarrow 0}\int_{B(0,r)^c}\left(\frac{i\xi}{|\xi|} - \frac{i\xi}{2\pi|\xi|^2\widehat{Q}_\delta^0(\xi)}\right)\eta(\xi)\,d\xi \quad \text{for } \eta\in\mathcal{S}(\mathbb{R}^n).$$

We may decompose Z_δ into the sum of another tempered distribution $Y_\delta \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C}^n)$ given by

$$\langle Y_\delta, \eta\rangle = \lim_{r\downarrow 0}\int_{B(0,r)^c}\mathbb{1}_{B(0,1)}(\xi)\frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\delta^0(0)}\eta(\xi)\,d\xi \quad \text{for } \eta\in\mathcal{S}(\mathbb{R}^n),$$

and the locally integrable function $X_\delta \in L^1_{\text{loc}}(\mathbb{R}^n;\mathbb{C}^n)$,

$$\xi \mapsto \left(\frac{i\xi}{2\pi|\xi|^2}\left(\frac{1}{\widehat{Q}_\delta^0(0)} - \frac{1}{\widehat{Q}_\delta^0(\xi)}\right) + \frac{i\xi}{|\xi|}\right)\mathbb{1}_{B(0,1)}(\xi) + \frac{i\xi}{|\xi|}\left(1 - \frac{1}{2\pi|\xi|\widehat{Q}_\delta^0(\xi)}\right)\mathbb{1}_{B(0,1)^c}(\xi).$$

The inverse Fourier transform of Y_δ corresponds to a bounded function; for the case $n \geq 2$, this is because Y_δ agrees with an integrable function, whereas for the case $n = 1$ this follows from [31, Lemma A.1 *b*)]. Moreover, we can show that X_δ is actually integrable. For the first term this follows from the fact that $\widehat{Q}_\delta^0$ is smooth and strictly positive, cf. Remark 3.2.3. For the second term, we use (3.76) to write for $|\xi| \geq 1$

$$1 - \frac{1}{2\pi|\xi|\widehat{Q}_\delta^0(\xi)} = 1 - \frac{1}{1 + 2\pi|\xi|\widehat{R}_\delta^0(\xi)} = \frac{2\pi|\xi|\widehat{R}_\delta^0(\xi)}{1 + 2\pi|\xi|\widehat{R}_\delta^0(\xi)},$$

which is integrable by Lemma 3.A.1. We conclude that X_δ also has a bounded inverse Fourier transform.

All in all, we conclude that there is a $W_\delta \in L^\infty(\mathbb{R}^n;\mathbb{R}^n)$ such that

$$\widehat{W_\delta} = Z_\delta.$$

Note that W_δ takes values in $\mathbb{R}^n$ as $\langle Z_\delta, \eta(-\,\cdot)\rangle = \overline{\langle Z_\delta, \eta\rangle}$ for $\eta \in \mathcal{S}(\mathbb{R}^n)$. Finally, using Proposition 3.2.2 for the Fourier transform of $D_\delta^0\varphi$, we have for $\varphi, \eta \in \mathcal{S}(\mathbb{R}^n)$ and $\psi = -RD_\delta^0\varphi + W_\delta * D_\delta^0\varphi \in \mathcal{S}'(\mathbb{R}^n)$ that

$$\begin{aligned}\langle\widehat{\psi}, \eta\rangle &= \int_{\mathbb{R}^n}\frac{-i\xi}{|\xi|}\cdot\widehat{D_\delta^0\varphi(\xi)}\eta(\xi)\,d\xi + \langle Z_\delta\widehat{D_\delta^0\varphi}, \eta\rangle\\ &= \lim_{r\downarrow 0}\int_{B(0,r)^c}\frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\delta^0(\xi)}\cdot\widehat{D_\delta^0\varphi}(\xi)\eta(\xi)\,d\xi\\ &= \lim_{r\downarrow 0}\int_{B(0,r)^c}\frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\delta^0(\xi)}\cdot\widehat{Q}_\delta^0(\xi)2\pi i\xi\widehat{\varphi}(\xi)\eta(\xi)\,d\xi = \int_{\mathbb{R}^n}\widehat{\varphi}(\xi)\eta(\xi)\,d\xi,\end{aligned}$$

which proves $\psi = \varphi \in \mathcal{S}(\mathbb{R}^n)$ after taking the inverse Fourier transform. □

3.2.3 Complementary-value spaces

Our study of variational problems involving the nonlocal gradient is carried out on affine subspaces of $H^{s,p,\delta}(\Omega)$ satisfying a complementary-value condition. For $\Omega \subset \mathbb{R}^n$ open and bounded, let

$$\Omega_{-\delta} = \{x \in \Omega \,:\, d(x,\partial\Omega) > \delta\}.$$

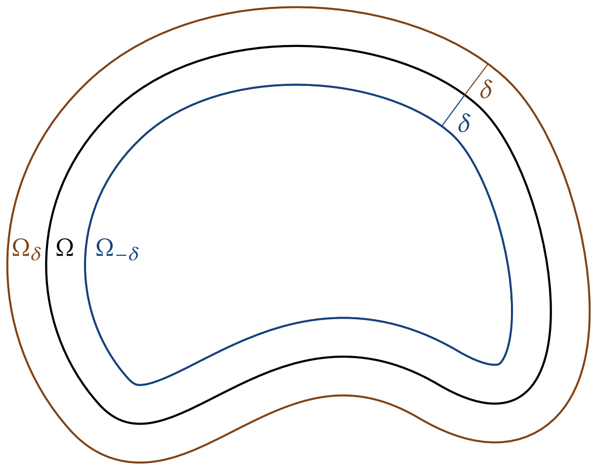

Figure 3.2: Illustration of the set Ω with its nonlocal closure Ω_δ, its nonlocal boundary $\Omega_\delta \setminus \Omega$, and the collar $\Omega_\delta \setminus \Omega_{-\delta}$ of thickness 2δ, where complementary values are prescribed.

Whenever we work with the complementary-value spaces throughout the paper, we assume implicitly that $\delta > 0$ is small enough so that $\Omega_{-\delta}$ is non-empty, that is, $0 < \delta < \max_{x\in\Omega} d(x, \partial\Omega)$; for an illustration of Ω and its inner and outer collar, see Figure 3.2. We define for $s \in [0, 1)$ and $p \in [1, \infty)$,

$$H_0^{s,p,\delta}(\Omega) = \overline{C_c^\infty(\Omega_{-\delta})}^{H^{s,p,\delta}(\Omega)} \tag{3.18}$$

and for $g \in H^{s,p,\delta}(\Omega)$ the complementary-value space

$$H_g^{s,p,\delta}(\Omega) = g + H_0^{s,p,\delta}(\Omega).$$

In a similar vain, we set

$$H_0^{1,p,\delta}(\Omega) := \overline{C_c^\infty(\Omega_{-\delta})}^{W^{1,p}(\Omega_\delta)}, \tag{3.19}$$

which will be used to study the asymptotics $s \to 1$.

In order to avoid confusion, we clarify that the notation used in this document for $H_g^{s,p,\delta}(\Omega)$ slightly differs from the one used in [30, 31], where the same spaces were denoted by $H_g^{s,p,\delta}(\Omega_{-\delta})$.

When $\Omega_{-\delta}$ is a Lipschitz domain, these affine subspaces comprise exactly those functions in $H^{s,p,\delta}(\Omega)$ that have prescribed values in $\Omega_\delta \setminus \Omega_{-\delta}$. Indeed, for the case $s = 1$, it is well-known that

$$H_0^{1,p,\delta}(\Omega) = \{u \in W^{1,p}(\Omega_\delta) \,:\, u = 0 \text{ a.e. in } \Omega_\delta \setminus \Omega_{-\delta}\},$$

whereas the case $s \in [0, 1)$ is treated in the next statement, which we prove in Appendix 3.B.

Proposition 3.2.10. *Let $s \in [0, 1)$, $p \in [1, \infty)$ and $\Omega \subset \mathbb{R}^n$ be open and bounded such that $\Omega_{-\delta}$ is a Lipschitz domain. Then,*

$$H_g^{s,p,\delta}(\Omega) = \{u \in H^{s,p,\delta}(\Omega) \,:\, u = g \textit{ a.e. in } \Omega_\delta \setminus \Omega_{-\delta}\}.$$

It may be surprising at first glance that we prescribe values in a collar of width 2δ around the boundary of Ω, yet, this choice leads to a natural treatment of the nonlocal variational problems in this paper, as can be seen for instance from the Poincaré inequality in [31, Theorem 6.2] and the Euler-Lagrange equations in [31, Theorem 8.2].

For a discussion of relevant properties and useful results on these function spaces, like Poincaré inequalities and compact embeddings, we refer to [30, 31]. Apart from those, there is the following Leibniz rule from [30, Lemma 3.2 and 3.3], which we will use among other things to enforce complementary-values via cut-off procedures.

Lemma 3.2.11 (Nonlocal Leibniz rule). *Let $s \in [0,1)$, $\delta > 0$, $p \in [1,\infty]$, and $\Omega \subset \mathbb{R}^n$ open. If $u \in H^{s,p,\delta}(\Omega)$ and $\chi \in C_c^\infty(\mathbb{R}^n)$, then $\chi u \in H^{s,p,\delta}(\Omega)$ with*

$$D_\delta^s(\chi u) = \chi D_\delta^s u + K_\chi(u),$$

where $K_\chi : L^p(\Omega_\delta) \to L^p(\Omega;\mathbb{R}^n)$ is the bounded linear operator given by

$$K_\chi(u)(x) = c_{n,s} \int_{B(x,\delta)} u(y) \frac{\chi(x)-\chi(y)}{|x-y|^{n+s}} \frac{x-y}{|x-y|} w_\delta(x-y)\, dy \quad \text{for } x \in \Omega,$$

and there is a $C > 0$ such that

$$\|K_\chi(u)\|_{L^p(\Omega;\mathbb{R}^n)} \leq C \mathrm{Lip}(\chi) \|u\|_{L^p(\Omega_\delta)}.$$

Proof. The statement for $u \in C_c^\infty(\mathbb{R}^n)$ with $s \in (0,1)$ and the bound for K_χ follow immediately from [30, Lemma 3.2 and 3.3] (the arguments remain valid for unbounded sets). We can extend it to $u \in H^{s,p,\delta}(\Omega)$ via a distributional argument as in Lemma 3.B.1. The case $s = 0$ can be proven analogously. □

In a similar spirit, one obtains with a slight abuse of notation that

$$\operatorname{div}_\delta^s(\chi u) = \chi \operatorname{div}_\delta^s u + K_\chi(u^\intercal) \tag{3.20}$$

for $u \in H^{s,p,\delta}(\Omega;\mathbb{R}^n)$ and $\chi \in C_c^\infty(\mathbb{R}^n)$; here $\zeta^\intercal$ indicates the transpose of a vector $\zeta \in \mathbb{R}^n$.

As a consequence of the Leibniz rule above, we can prove that in complementary-value spaces, weak convergence of nonlocal gradients improves to strong convergence in the strip where the values are prescribed. The following result shows natural parallels with [140, Lemma 2.12] in the context of Riesz fractional gradients.

Lemma 3.2.12 (Strong convergence in the collar). *Let $s \in (0,1)$, $p \in (1,\infty)$, $\Omega \subset \mathbb{R}^n$ open and bounded, $O \subset \Omega$ open with $\Omega_{-\delta} \Subset O$, and $g \in H^{s,p,\delta}(\Omega)$. If $\{u_j\}_{j\in\mathbb{N}} \subset H_g^{s,p,\delta}(\Omega)$ converges weakly to u in $H^{s,p,\delta}(\Omega)$, then*

$$D_\delta^s u_j \to D_\delta^s u \quad \text{in } L^p(\Omega \setminus O;\mathbb{R}^n).$$

Proof. Due to linearity, it suffices to prove the statement for the special case $u = 0$ and $g = 0$. Let us consider therefore a sequence $\{u_j\}_{j\in\mathbb{N}} \subset H_0^{s,p,\delta}(\Omega)$ with $u_j \rightharpoonup 0$ in $H^{s,p,\delta}(\Omega)$. With $\chi \in C_c^\infty(O)$ a cut-off function with $\chi \equiv 1$ on $\Omega_{-\delta}$, we obtain that $u_j = \chi u_j$ for $j \in \mathbb{N}$ and $u_j \to 0$ in $L^p(\Omega_\delta)$ as a consequence of [31, Theorem 7.3]. Hence, by Lemma 3.2.11,

$$\begin{aligned}\|D_\delta^s u_j\|_{L^p(\Omega\setminus O;\mathbb{R}^n)} &= \|D_\delta^s(\chi u_j)\|_{L^p(\Omega\setminus O;\mathbb{R}^n)} \\ &\leq \|D_\delta^s(\chi u_j) - \chi D_\delta^s u_j\|_{L^p(\Omega;\mathbb{R}^n)} = \|K_\chi(u_j)\|_{L^p(\Omega;\mathbb{R}^n)} \to 0 \quad \text{as } j \to \infty,\end{aligned}$$

exploiting the continuity of $K_\chi : L^p(\Omega_\delta) \to L^p(\Omega;\mathbb{R}^n)$. □

3.2.4 Connection between nonlocal and classical Sobolev spaces

One of the key tools for our analysis is the following proposition, which allows us to switch between nonlocal and classical gradients and is the technical basis for an effective translation mechanism. It is the counterpart of [140, Proposition 3.1], where fractional gradients and their relation with classical ones are analyzed.

We first introduce the operator

$$\mathcal{P}_\delta^s : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n), \quad \varphi \mapsto \left(\frac{\widehat{\varphi}}{\widehat{Q_\delta^s}} \right)^\vee, \tag{3.21}$$

which is well-defined since $1/\widehat{Q_\delta^s}$ is a smooth function with polynomially bounded derivatives (cf. Remark 3.2.3 and [31, Eq. (29)]). Moreover, as a consequence of the Fourier representation,

$$\mathcal{P}_\delta^s(Q_\delta^s * \varphi) = Q_\delta^s * (\mathcal{P}_\delta^s \varphi) = \varphi \quad \text{for all } \varphi \in \mathcal{S}(\mathbb{R}^n), \tag{3.22}$$

which implies, in particular, that $D_\delta^s(\mathcal{P}_\delta^s\varphi) = \nabla(Q_\delta^s * (\mathcal{P}_\delta^s\varphi)) = \nabla\varphi$. We now extend these properties to the Sobolev spaces.

Theorem 3.2.13 (Translating between nonlocal and classical gradients). *Let $s \in (0,1)$, $p \in [1,\infty]$ and $\Omega \subset \mathbb{R}^n$ be open. The following two statements hold:*

- (i) *The operator $\mathcal{Q}_\delta^s : H^{s,p,\delta}(\Omega) \to W^{1,p}(\Omega)$, $u \mapsto Q_\delta^s * u$ is bounded and if $u \in H^{s,p,\delta}(\Omega)$, then $v = \mathcal{Q}_\delta^s u$ satisfies $\nabla v = D_\delta^s u$ on Ω.*
- (ii) *The operator $\mathcal{P}_\delta^s$ in (3.21) can be extended to a bounded linear operator from $W^{1,p}(\mathbb{R}^n)$ to $H^{s,p,\delta}(\mathbb{R}^n)$ such that $\mathcal{P}_\delta^s = (\mathcal{Q}_\delta^s)^{-1}$, i.e.,*

$$\mathcal{P}_\delta^s\mathcal{Q}_\delta^s u = u \ \text{ for } u \in H^{s,p,\delta}(\mathbb{R}^n) \quad \text{and} \quad \mathcal{Q}_\delta^s\mathcal{P}_\delta^s v = v \ \text{ for } v \in W^{1,p}(\mathbb{R}^n);$$

in particular, if $v \in W^{1,p}(\mathbb{R}^n)$, then $u = \mathcal{P}_\delta^s v$ satisfies $D_\delta^s u = \nabla v$ on $\mathbb{R}^n$.

Proof. Part (i): Let $u \in H^{s,p,\delta}(\Omega)$, then $v = \mathcal{Q}_\delta^s u \in L^p(\Omega)$ since $Q_\delta^s \in L^1(\mathbb{R}^n)$. For every $\varphi \in C_c^\infty(\Omega;\mathbb{R}^n)$, we find that

$$\begin{aligned}\int_\Omega v \operatorname{div}\varphi\,dx &= \int_{\Omega_\delta} u\,(Q_\delta^s * \operatorname{div}\varphi)\,dx \\ &= \int_{\Omega_\delta} u \operatorname{div}_\delta^s \varphi\,dx = -\int_\Omega D_\delta^s u \cdot \varphi\,dx,\end{aligned}$$

where the first identity uses Fubini's theorem, the second one follows from Proposition 3.2.2, and the third one is simply the definition of the weak nonlocal gradient. This proves $v \in W^{1,p}(\Omega)$ with $\nabla v = D_\delta^s u$ on Ω. The boundedness of $\mathcal{Q}_\delta^s$ follows from Young's convolution inequality.

Part (ii): Since $\mathcal{P}_\delta^s$ is the inverse of the mapping $\mathcal{Q}_\delta^s$ on $\mathcal{S}(\mathbb{R}^n)$ (cf. (3.22)), it is sufficient to prove that $\mathcal{Q}_\delta^s$ is boundedly invertible. Indeed, we can then find the suitable extension by setting $\mathcal{P}_\delta^s := (\mathcal{Q}_\delta^s)^{-1}$. Since $\mathcal{Q}_\delta^s$ is bounded by part (i), we only need to prove bijectivity to deduce the statement via Banach's isomorphism theorem.

Step 1: Injectivity. Suppose that $\mathcal{Q}_\delta^s u = Q_\delta^s * u = 0$ for $u \in H^{s,p,\delta}(\mathbb{R}^n)$. Then, $D_\delta^s u = \nabla(Q_\delta^s * u) = 0$, and in particular,

$$\int_{\mathbb{R}^n} u \operatorname{div}_\delta^s \varphi\,dx = 0 \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n);$$

by density, this also holds for all $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{R}^n)$. By taking any $\psi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$ and setting $\varphi = \mathcal{P}_\delta^s \psi \in \mathcal{S}(\mathbb{R}^n;\mathbb{R}^n)$, we obtain

$$0 = \int_{\mathbb{R}^n} u \operatorname{div}_\delta^s \varphi \, dx = \int_{\mathbb{R}^n} u \operatorname{div}_\delta^s \mathcal{P}_\delta^s \psi \, dx = \int_{\mathbb{R}^n} u \operatorname{div} \psi \, dx.$$

Hence, u is constant. Together with $Q_\delta^s * u = 0$, this shows that $u = 0$ and proves the injectivity of $\mathcal{Q}_\delta^s$.

Step 2: Surjectivity. Take $v \in W^{1,p}(\mathbb{R}^n)$ and $\chi \in C_c^\infty(\mathbb{R}^n)$ an even function with $\chi \equiv 1$ on $B(0,1)$. Define the functions $\varphi_1, \varphi_2 \in \mathcal{S}(\mathbb{R}^n)$ by

$$\varphi_1 = -\chi^\vee \quad \text{and} \quad \varphi_2 = \left(\frac{\chi}{\widehat{Q_\delta^s}} + (1-\chi)\left(\frac{1}{\widehat{Q_\delta^s}} - |2\pi \cdot|^{1-s} \right) \right)^\vee;$$

here, $\varphi_2 \in \mathcal{S}(\mathbb{R}^n)$ since

$$\frac{1}{\widehat{Q_\delta^s}} - |2\pi \cdot|^{1-s} = \frac{-|2\pi \cdot|^{1-s}\widehat{R_\delta^s}}{\widehat{Q_\delta^s}} \quad \text{in } B(0,1)^c$$

in view of (3.76), and $\widehat{R_\delta^s}$ agrees with a Schwartz function on $B(0,1)^c$ by (3.77). Note also that φ_1 and φ_2 are real-valued and even, since the same holds for their Fourier transforms. Because $(-\Delta)^{\frac{1-s}{2}}$ may be extended to a bounded linear operator from $W^{1,p}(\mathbb{R}^n)$ to $H^{s,p}(\mathbb{R}^n)$ (cf. [140, Proposition 3.1 (ii)]) and $H^{s,p}(\mathbb{R}^n) = H^{s,p,\delta}(\mathbb{R}^n)$ by Lemma 3.24 and Remark 3.2.17, we can define

$$w := (-\Delta)^{\frac{1-s}{2}} v + \varphi_1 * (-\Delta)^{\frac{1-s}{2}} v + \varphi_2 * v \in H^{s,p,\delta}(\mathbb{R}^n).$$

Using Fubini's theorem and the duality for the fractional Laplacian (see e.g., [140, Eq. (3.6)]), we find for $\varphi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$

$$\begin{aligned}\int_{\mathbb{R}^n} w \operatorname{div}_\delta^s \varphi \, dx &= \int_{\mathbb{R}^n} v\big((-\Delta)^{\frac{1-s}{2}} \operatorname{div}_\delta^s \varphi + (-\Delta)^{\frac{1-s}{2}}(\varphi_1 * \operatorname{div}_\delta^s \varphi) + \varphi_2 * \operatorname{div}_\delta^s \varphi\big)\, dx \\ &= \int_{\mathbb{R}^n} v \operatorname{div} \varphi \, dx,\end{aligned}$$

where the last inequality follows from

$$\begin{aligned}&\left((-\Delta)^{\frac{1-s}{2}} \operatorname{div}_\delta^s \varphi + (-\Delta)^{\frac{1-s}{2}}(\varphi_1 * \operatorname{div}_\delta^s \varphi) + \varphi_2 * \operatorname{div}_\delta^s \varphi\right)^\wedge(\xi) \\ &\quad = \big(|2\pi\xi|^{1-s} + |2\pi\xi|^{1-s}\widehat{\varphi_1}(\xi) + \widehat{\varphi_2}(\xi)\big)\widehat{Q_\delta^s}(\xi) 2\pi i \xi \cdot \widehat{\varphi}(\xi) = 2\pi i\xi \cdot \widehat{\varphi}(\xi) = \widehat{\operatorname{div}\varphi}(\xi).\end{aligned}$$

We conclude that $D_\delta^s w = \nabla v$, which means that $\mathcal{Q}_\delta^s w - v \equiv c$ for some $c \in \mathbb{R}$; if $p < \infty$, then $c = 0$ since both $\mathcal{Q}_\delta^s w$ and v lie in $L^p(\mathbb{R}^n)$. Therefore, we obtain

$$\mathcal{Q}_\delta^s \left(w - \frac{c}{\|Q_\delta^s\|_{L^1(\mathbb{R}^n)}} \right) = v,$$

which shows the surjectivity of $\mathcal{Q}_\delta^s$ and finishes the proof. □

Remark 3.2.14. a) Note that the proof of the fractional version of (i) in [140, Proposition 3.1 (i)] has to deal with a technical difficulty that the Riesz kernel I_{1-s} is not integrable as opposed to Q_δ^s. Therefore, the convolution of I_{1-s} with an L^p-function for large p is not always well-defined, whereas Q_δ^s can be convolved with any L^p-function. In particular, there is also no perfect identification between $H^{s,p}(\mathbb{R}^n)$ and $W^{1,p}(\mathbb{R}^n)$ that turns fractional gradients into classical gradients as for the nonlocal case in part (ii) above.

b) Regarding part (ii), when $p < \infty$ then the extension of $\mathcal{P}_\delta^s$ can also be seen as the unique extension via density. Moreover, if Ω is a Lipschitz domain, then any $v \in W^{1,p}(\Omega)$ can be extended to a function in $W^{1,p}(\mathbb{R}^n)$, after which we can apply the result to find a $u \in H^{s,p,\delta}(\Omega)$ with $D_\delta^s u = \nabla v$ on Ω.

c) The proof of the surjectivity in part (ii) shows that $\mathcal{P}_\delta^s v$ corresponds, up to a constant, to

$$(-\Delta)^{\frac{1-s}{2}} v + \varphi_1 * (-\Delta)^{\frac{1-s}{2}} v + \varphi_2 * v,$$

for $v \in W^{1,p}(\mathbb{R}^n)$; when $p < \infty$, then the correspondence is even an identity, given that there are no non-zero constants in $L^p(\mathbb{R}^n)$.

As a particular consequence of this observation, along with the fact that the convolution with a periodic function remains periodic, we observe that both $\mathcal{Q}_\delta^s$ and $\mathcal{P}_\delta^s$ preserve periodicity. Precisely, if Y denotes the unit cube $(0,1)^n$, and $W_\#^{1,\infty}(Y)$ and $H_\#^{s,\infty,\delta}(Y)$ comprise all Y-periodic functions in $W^{1,\infty}(\mathbb{R}^n)$ and $H^{s,\infty,\delta}(\mathbb{R}^n)$, respectively, then there is a bijection between the gradients of $W_\#^{1,\infty}(Y)$-functions and the nonlocal gradients of $H_\#^{s,\infty,\delta}(Y)$-functions.

d) For $\varphi \in \mathcal{S}(\mathbb{R}^n)$, it holds that

$$\mathcal{P}_\delta^s \varphi(x) = \int_{\mathbb{R}^n} V_\delta^s(x-y) \cdot \nabla \varphi(y)\, dy \quad \text{for } x \in \mathbb{R}^n, \tag{3.23}$$

where $V_\delta^s \in C^\infty(\mathbb{R}^n \backslash \{0\})$ is the kernel from the nonlocal version of the fundamental theorem of calculus [31, Theorem 4.5]. Indeed, this follows directly from the formula for the Fourier transform of V_δ^s in [31, Theorem 5.9]. The representation in (3.23) extends naturally to functions in $W^{1,p}(\mathbb{R}^n)$ with compact support, given that V_δ^s is locally integrable.

e) The translation procedure of Theorem 3.2.13 allows us to give an alternative proof for the nonlocal Poincaré inequality in [31, Theorem 6.2]. Since $\mathcal{Q}_\delta^s$ maps $H_0^{s,p,\delta}(\Omega)$ into $W_0^{1,p}(\Omega)$, we infer from the classical Poincaré inequality that

$$\|u\|_{L^p(\Omega)} = \|\mathcal{P}_\delta^s \mathcal{Q}_\delta^s u\|_{L^p(\Omega)} \le C\|\mathcal{Q}_\delta^s u\|_{W_0^{1,p}(\Omega)} \le C\|\nabla \mathcal{Q}_\delta^s u\|_{L^p(\Omega;\mathbb{R}^n)} = C\|D_\delta^s u\|_{L^p(\Omega;\mathbb{R}^n)},$$

for any $u \in H_0^{s,p,\delta}(\Omega)$ with a constant $C > 0$ depending on s, δ, p, and Ω. △

We conclude this section with a compactness result that will be used below in the proof of Theorem 3.4.1.

Lemma 3.2.15. *Let $s \in (0,1)$, $p \in (1,\infty)$ and $\Omega \subset \mathbb{R}^n$ be open and bounded. If $\{v_j\}_{j\in\mathbb{N}} \subset W^{1,p}(\mathbb{R}^n)$ is a bounded sequence, then $\{\mathcal{P}_\delta^s v_j\}_{j\in\mathbb{N}}$ (when restricted to Ω_δ) is relatively compact in $L^p(\Omega_\delta)$.*

Proof. Let $\chi \in C_c^\infty(\mathbb{R}^n)$ with $\chi \equiv 1$ on Ω_δ and set $R > 0$ such that $\operatorname{supp}(\chi) \subset B(0, R-\delta)$. Since $\{\mathcal{P}_\delta^s v_j\}_j$ is bounded in $H^{s,p,\delta}(\mathbb{R}^n)$ by Theorem 3.2.13 (ii), the sequence $\{\chi(\mathcal{P}_\delta^s v_j)\}_{j\in\mathbb{N}}$ is bounded in $H_0^{s,p,\delta}(B(0,R))$ by Lemma 3.2.11. The relative compactness of $\{\chi(\mathcal{P}_\delta^s v_j)\}_{j\in\mathbb{N}}$ in $L^p(B(0,R))$ now follows from [31, Theorem 7.3] and since $\chi \equiv 1$ on Ω_δ, the statement follows. □

3.2.5 Connection between nonlocal and fractional gradients

After the comparison of the nonlocal gradients with classical weak gradients, let us now discuss their connection with the Riesz fractional gradient. We start by recalling that the nonlocal gradient is a truncated version of the latter.

In the following, let $p \in [1, \infty)$ and $s \in (0, 1)$. The upcoming lemma presents the equivalence between the nonlocal and fractional Sobolev spaces $H^{s,p}(\mathbb{R}^n)$ and $H^{s,p,\delta}(\mathbb{R}^n)$ and also a version with prescribed complementary values. To recall the definition of the fractional and nonlocal complementary-value spaces, we have for $\Omega \subset \mathbb{R}^n$ open and bounded that $H_0^{s,p}(\Omega)$ comprises all functions $u \in H^{s,p}(\mathbb{R}^n)$ such that $u = 0$ a.e. in Ω^c and $H_0^{s,p,\delta}(\Omega)$ is given as in (3.18). We mention that one of the inclusions was already provided by [31, Proposition 3.5]. For the sake of the reader, we show here a complete proof.

Lemma 3.2.16. *It holds that*

$$H^{s,p}(\mathbb{R}^n) = H^{s,p,\delta}(\mathbb{R}^n) \tag{3.24}$$

with equivalent norms, and

$$D_\delta^s u = D^s u + \nabla R_\delta^s * u \tag{3.25}$$

for all $u \in H^{s,p}(\mathbb{R}^n) = H^{s,p,\delta}(\mathbb{R}^n)$ with $\nabla R_\delta^s \in C^\infty(\mathbb{R}^n;\mathbb{R}^n) \cap L^1(\mathbb{R}^n;\mathbb{R}^n)$ as in (3.74). *Moreover, for $\Omega \subset \mathbb{R}^n$ open and bounded with $\Omega_{-\delta}$ Lipschitz, it holds that*

$$H_0^{s,p}(\Omega_{-\delta}) = H_0^{s,p,\delta}(\Omega),$$

with equivalent norms, and (3.25) *holds for $u \in H_0^{s,p}(\Omega_{-\delta}) = H_0^{s,p,\delta}(\Omega)$ on Ω.*

Proof. Let $\varphi \in C_c^\infty(\mathbb{R}^n)$. Since by (3.75)

$$D_\delta^s \varphi - D^s \varphi = \nabla R_\delta^s * \varphi,$$

we obtain the estimates

$$\|D^s\varphi\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \leq \|D_\delta^s\varphi\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} + \|\nabla R_\delta^s\|_{L^1(\mathbb{R}^n;\mathbb{R}^n)}\|\varphi\|_{L^p(\mathbb{R}^n)} \leq C\|\varphi\|_{H^{s,p,\delta}(\mathbb{R}^n)}$$

and

$$\|D_\delta^s\varphi\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \leq \|D^s\varphi\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} + \|\nabla R_\delta^s\|_{L^1(\mathbb{R}^n;\mathbb{R}^n)}\|\varphi\|_{L^p(\mathbb{R}^n)} \leq C\|\varphi\|_{H^{s,p}(\mathbb{R}^n)}$$

with a constant $C > 0$. In light of the density of $C_c^\infty(\mathbb{R}^n)$ in $H^{s,p}(\mathbb{R}^n)$ and $H^{s,p,\delta}(\mathbb{R}^n)$ (see [140, Theorem 2.7] and Theorem 3.2.8), the identity (3.24) and (3.25) follow via approximation. For the case of a bounded domain, we note that

$$H_0^{s,p,\delta}(\Omega) = \{u \in H^{s,p,\delta}(\Omega) \,:\, u = 0 \text{ a.e. in } \Omega_\delta \setminus \Omega_{-\delta}\}$$

since $\Omega_{-\delta}$ is Lipschitz (cf. Proposition 3.2.10). Observe also that for any $u \in H_0^{s,p,\delta}(\Omega)$ its extension $\bar{u}$ to $\mathbb{R}^n$ by zero lies in $H^{s,p,\delta}(\mathbb{R}^n)$ with

$$\|\bar{u}\|_{L^p(\mathbb{R}^n)} + \|D_\delta^s\bar{u}\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} = \|u\|_{L^p(\Omega)} + \|D_\delta^s u\|_{L^p(\Omega;\mathbb{R}^n)},$$

since $D_\delta^s\bar{u}$ is simply the extension of $D_\delta^s u$ by zero. Hence, we may identify

$$H_0^{s,p,\delta}(\Omega) = \{u \in H^{s,p,\delta}(\mathbb{R}^n) \,:\, u = 0 \text{ a.e. in } (\Omega_{-\delta})^c\},$$

after which the equality with $H^{s,p}(\Omega_{-\delta})$ becomes obvious given (3.24). □

Remark 3.2.17. We mention that it also holds that

$$H^{s,\infty}(\mathbb{R}^n) = H^{s,\infty,\delta}(\mathbb{R}^n),$$

with equivalent norms and $D^s_\delta u = D^s u + \nabla R^s_\delta * u$ for $u \in H^{s,\infty}(\mathbb{R}^n)$. This can be proven via a distributional approach instead of utilizing density as above. △

As already indicated in the introduction, Lemma 3.2.5 opens up a new proof strategy for some of the results in this paper. Instead of exploiting well-known result for problems involving classical gradients, one can resort to established findings in the fractional setting. We illustrate this approach below by presenting an alternative proof for the characterization of lower semicontinuity in Section 3.4, which follows as a corollary of [140, Theorems 4.1 and 4.5]. An analogous reasoning could also be used, for instance, to deduce the relaxation below in Corollary 3.5.3 from [140, Theorem 1.2]. Note that the transfer of results between the nonlocal and fractional set-up also works in the reverse direction, giving rise to analogues of the general Γ-convergence statement in Theorem 3.5.1 and homogenization result of Corollary 3.5.2.

3.3 Asymptotics of the nonlocal gradient and applications

Our next goal is to study the localization of the nonlocal gradient as $s \to 1$, and more generally, to understand how the nonlocal gradient depends on the fractional parameter s. In particular, the findings in this section serve as necessary preparations for proving the Γ-convergence of nonlocal integral functionals in Section 3.6.

We start by investigating the s-dependence of the convolution kernel Q^s_δ from (3.16) and its Fourier transform.

Lemma 3.3.1. *Let $\varepsilon > 0$ and $R > 0$.*

(*i*) *The map $[0,1) \to L^1(\mathbb{R}^n)$, $s \mapsto Q^s_\delta$ is continuous with*

$$\lim_{s\to 1} \|Q^s_\delta\|_{L^1(\mathbb{R}^n)} = 1 \quad \text{and} \quad \lim_{s\to 1} \|Q^s_\delta\|_{L^1(B(0,\varepsilon)^c)} = 0. \tag{3.26}$$

(*ii*) *The map $[0,1) \to C(\overline{B(0,R)})$, $s \mapsto \widehat{Q}^s_\delta$ is continuous with $\widehat{Q}^s_\delta \to 1$ uniformly on $\overline{B(0,R)}$ as $s \to 1$.*

Proof. As for (*i*), we calculate first that for $s \in [0,1)$,

$$\begin{aligned}
\|Q^s_\delta\|_{L^1(\mathbb{R}^n)} &= c_{n,s} \int_{B(0,\delta)} \int_{|x|}^{\delta} \frac{\overline{w}_\delta(r)}{r^{n+s}}\, dr\, dx = c_{n,s} |\partial B(0,1)| \int_0^\delta \int_\rho^\delta \frac{\overline{w}_\delta(r)}{r^{n+s}} \rho^{n-1}\, dr\, d\rho \\
&= c_{n,s} |\partial B(0,1)| \int_0^\delta \frac{\overline{w}_\delta(r)}{r^{n+s}} \int_0^r \rho^{n-1}\, d\rho\, dr = c_{n,s} \frac{|\partial B(0,1)|}{n} \int_0^\delta \frac{\overline{w}_\delta(r)}{r^s}\, dr \\
&= c_{n,s}\, \omega_n \int_0^\delta \frac{\overline{w}_\delta(r)}{r^s}\, dr,
\end{aligned} \tag{3.27}$$

where ω_n denotes the volume of the unit ball in $\mathbb{R}^n$. Since $s \mapsto c_{n,s}$ is continuous on $[0,1)$, cf. [29, Lemma 2.4], it follows via Lebesgue's dominated convergence that $\|Q^s_\delta\|_{L^1(\mathbb{R}^n)}$ depends continuously on s. Now, if $\{s_j\}_{j\in\mathbb{N}} \subset [0,1)$ is a sequence converging to $s \in [0,1)$, we can apply once again Lebesgue's dominated convergence theorem to find that $Q^{s_j}_\delta \to Q^s_\delta$ pointwise a.e. as $j \to \infty$. Together with $\lim_{j\to\infty} \|Q^{s_j}_\delta\|_{L^1(\mathbb{R}^n)} = \|Q^s_\delta\|_{L^1(\mathbb{R}^n)}$ as shown above, this implies $Q^{s_j}_\delta \to Q^s_\delta$ in $L^1(\mathbb{R}^n)$ for $j \to \infty$.

To see the first convergence in (3.26), we observe that $\mathbb{1}_{B(0,b_0\delta)} \leq w_\delta \leq \mathbb{1}_{B(0,\delta)}$ by (H3) and (H4), which gives

$$\frac{(b_0\delta)^{1-s}}{1-s} \leq \int_0^\delta \frac{\overline{w}_\delta(r)}{r^s}\,dr \leq \frac{\delta^{1-s}}{1-s}, \tag{3.28}$$

and exploit $c_{n,s}/(1-s) \to 1/\omega_n$ as $s \to 1$ according to [29, Lemma 2.4]. The localization of Q^s_δ for $s \to 1$ follows from a calculation similar to (3.27) and (3.28), integrating instead over $B(\varepsilon,\delta)$ and using that $c_{n,s}/(1-s)$ stays bounded as $s \to 1$.

The first part of (ii) can be deduced from the continuity of the map in (i) in combination with the fact that the Fourier transform is a bounded linear operator from $L^1(\mathbb{R}^n)$ to $C_0(\mathbb{R}^n;\mathbb{C})$.

Due to (3.26), the kernel Q^s_δ behaves like a mollifier, satisfying

$$\lim_{s\to 1}\|Q^s_\delta * \varphi - \varphi\|_{L^\infty(\mathbb{R}^n)} = 0 \tag{3.29}$$

for all $\varphi \in \mathrm{Lip}_b(\mathbb{R}^n)$, where this convergence is uniform on bounded sets of $\mathrm{Lip}_b(\mathbb{R}^n)$. Indeed,

$$\begin{aligned}\|Q^s_\delta * \varphi - \varphi\|_{L^\infty(\mathbb{R}^n)} &\leq \|(\mathbb{1}_{B(0,\varepsilon)}Q^s_\delta) * \varphi - \varphi\|_{L^\infty(\mathbb{R}^n)} + \|Q^s_\delta\|_{L^1(B(0,\varepsilon)^c)}\|\varphi\|_{L^\infty(\mathbb{R}^n)}\\ &\leq \varepsilon\mathrm{Lip}(\varphi)\|Q^s_\delta\|_{L^1(B(0,\varepsilon))} + \big(|1-\|Q^s_\delta\|_{L^1(B(0,\varepsilon))}| + \|Q^s_\delta\|_{L^1(B(0,\varepsilon)^c)}\big)\,\|\varphi\|_{L^\infty(\mathbb{R}^n)},\end{aligned}$$

for any $\varepsilon > 0$. Considering now $\varphi_\xi(x) = e^{-2\pi i\xi\cdot x}$ for $\xi \in \overline{B(0,R)}$, we have

$$\|\varphi_\xi\|_{L^\infty(\mathbb{R}^n;\mathbb{C})} + \mathrm{Lip}(\varphi_\xi) \leq 1 + 2\pi|\xi| \leq 1 + 2\pi R,$$

so that by (3.29),

$$\lim_{s\to 1}\widehat{Q^s_\delta}(\xi) = \lim_{s\to 1}(Q^s_\delta * \varphi_\xi)(0) = \varphi_\xi(0) = 1,$$

uniformly for $\xi \in \overline{B(0,R)}$. □

The next lemma addresses the continuous dependence of the nonlocal gradient and divergence on the fractional parameter in the case of smooth test functions with compact support. Recall the notation $D^1_\delta u := \nabla u$.

Lemma 3.3.2. *Let $s \in [0,1]$ and $\{s_j\}_{j\in\mathbb{N}} \subset [0,1]$ a sequence converging to s. Then, it holds for every $\varphi \in C^\infty_c(\mathbb{R}^n)$ and $\psi \in C^\infty_c(\mathbb{R}^n;\mathbb{R}^n)$ that*

$$D^{s_j}_\delta\varphi \to D^s_\delta\varphi \qquad \text{and} \qquad \mathrm{div}^{s_j}_\delta\,\psi \to \mathrm{div}^s_\delta\,\psi$$

uniformly on $\mathbb{R}^n$ as $j \to \infty$.

Proof. It suffices to focus on proving the convergence of the nonlocal gradient; the argument for the divergence is an immediate consequence. If $s < 1$, we conclude from Proposition 3.2.2, Young's convolution inequality, and Lemma 3.3.1 (i) that

$$\|D^{s_j}_\delta\varphi - D^s_\delta\varphi\|_{L^\infty(\mathbb{R}^n;\mathbb{R}^n)} \leq \|Q^{s_j}_\delta - Q^s_\delta\|_{L^1(\mathbb{R}^n)}\|\nabla\varphi\|_{L^\infty(\mathbb{R}^n;\mathbb{R}^n)} \to 0 \ \text{ as } j \to \infty.$$

The case $s = 1$ follows immediately from (3.29), since $\nabla\varphi \in \mathrm{Lip}_b(\mathbb{R}^n)$ allows us to conclude that

$$\lim_{j\to\infty}\|D^{s_j}_\delta\varphi - \nabla\varphi\|_{L^\infty(\mathbb{R}^n;\mathbb{R}^n)} = \lim_{j\to\infty}\|Q^{s_j}_\delta * \nabla\varphi - \nabla\varphi\|_{L^\infty(\mathbb{R}^n;\mathbb{R}^n)} = 0.$$

□

Our approach to extending the previous results for smooth functions in a suitable way to nonlocal Sobolev spaces, relies on the following estimate (see Corollary 3.3.4 below),

$$\|D^s_\delta u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \leq C\|D^t_\delta u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \quad \text{for all } u \in C_c^\infty(\mathbb{R}^n) \text{ and } 0 \leq s \leq t \leq 1, \tag{3.30}$$

with a constant $C > 0$ depending only on n, p, δ. If $t = 1$, (3.30) simply follows from Young's convolution inequality

$$\|D^s_\delta u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \leq \|Q^s_\delta\|_{L^1(\mathbb{R}^n)}\|\nabla u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \leq C\|\nabla u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)}, \tag{3.31}$$

where we have exploited that $\|Q^s_\delta\|_{L^1(\mathbb{R}^n)}$ is bounded by C uniformly in s as a consequence of Lemma 3.3.1. If $t = s$, one can obviously take the constant to be 1. For the other cases, we build on Fourier multiplier theory (see e.g., [122, Chapter 5]) and show via the Mihlin-Hörmander theorem that the maps

$$m^s_t : \mathbb{R}^n \to \mathbb{R}, \quad \xi \mapsto \frac{\widehat{Q}^s_\delta(\xi)}{\widehat{Q}^t_\delta(\xi)} \tag{3.32}$$

are L^p-multipliers with uniformly bounded norms. This requires control on the decay behavior of m^s_t and its derivatives. The idea for deriving suitable bounds for large frequencies is to compare Q^s_δ with the well-known Riesz potential kernel I_{1-s} (cf. (3.10)) and exploit the decay of the difference of their Fourier transforms uniformly in s (see Lemma 3.A.1).

Lemma 3.3.3. *The map $m^s_t : \mathbb{R}^n \to \mathbb{R}$ from* (3.32) *with $0 \leq s \leq t < 1$ is an L^p-multiplier for every $p \in (1, \infty)$ with multiplier norm independent of the parameters s, t.*

Proof. According to the Mihlin-Hörmander multiplier theorem, see e.g., [122, Theorem 6.2.7], the statement follows immediately once these estimates have been established: There exists a constant $C > 0$ depending only on n and δ such that for every $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq n/2 + 1$ and every $0 \leq s \leq t < 1$,

$$|\xi|^{|\alpha|}\left|\partial^\alpha m^s_t(\xi)\right| \leq C \qquad \text{for all } \xi \in \mathbb{R}^n. \tag{3.33}$$

The proof is split in two parts, where we distinguish bounds for large and small frequencies. Note that in the following all the constants $C, c > 0$ are independent of s, t.

Step 1: Bounds away from zero. In this step, we show that there is some $R \geq 1$ such that (3.33) holds for all $|\xi| \geq R$. Since

$$\widehat{Q}^s_\delta(\xi) = |2\pi\xi|^{-(1-s)} + \widehat{R}^s_\delta(\xi) \qquad \text{for } |\xi| \geq 1$$

for any $s \in [0, 1)$ by (3.76), we can express m^s_t on $B(0, 1)^c$ as

$$m^s_t(\xi) = \frac{\widehat{Q}^s_\delta(\xi)}{\widehat{Q}^t_\delta(\xi)} = |2\pi\xi|^{-(t-s)} + r^s_t(\xi) \tag{3.34}$$

with

$$r^s_t(\xi) := \frac{-|2\pi\xi|^{-(t-s)}\widehat{R}^t_\delta(\xi) + \widehat{R}^s_\delta(\xi)}{|2\pi\xi|^{-(1-t)} + \widehat{R}^t_\delta(\xi)}.$$

Given $t \geq s$, it is clear that

$$\partial^\alpha\left(|2\pi\xi|^{-(t-s)}\right) \leq C|\xi|^{-|\alpha|} \quad \text{for } |\xi| \geq 1. \tag{3.35}$$

Along with (3.77), one can estimate the denominator of r_t^s and find some $R \geq 1$ such that for all $\xi \in \mathbb{R}^n$ with $|\xi| \geq R$,

$$|2\pi\xi|^{-(1-t)} + \widehat{R}_\delta^t(\xi) \geq |2\pi\xi|^{-1} - c|\xi|^{-2} \geq C|\xi|^{-1}. \tag{3.36}$$

If one takes the αth derivative of r_t^s on $B(0,R)^c$, the quotient rule gives rise to a quotient whose denominator results from raising the denominator of r_t^s to the power $2^{|\alpha|}$ and whose numerator is a product of $\widehat{R}_\delta^s$, $\widehat{R}_\delta^t$ and their derivatives with terms bounded independently of s, t. We therefore obtain in view of (3.36), and again (3.77), that

$$|\partial^\alpha r_t^s(\xi)| \leq C|\xi|^{-2^{|\alpha|}} \leq C|\xi|^{-|\alpha|} \qquad \text{for } |\xi| \geq R. \tag{3.37}$$

The combination of (3.34), (3.35) and (3.37) then yields (3.33) on $B(0,R)^c$.

Step 2: Local bounds. To show that (3.33) holds for $|\xi| \leq R$, we observe first that, as a consequence of Lemma 3.3.1 (ii) and the non-negativity of $\widehat{Q}_\delta^s$ (cf. Remark 3.2.3), there is a constant $c > 0$ such that

$$\widehat{Q}_\delta^s(\xi) \geq c$$

for all $\xi \in \overline{B(0,R)}$ and all $s \in [0,1)$. Moreover, for any $\beta \in \mathbb{N}_0^n$ with $|\beta| \leq n/2+1$, the estimate

$$\|(-2\pi i\,\cdot)^\beta Q_\delta^s\|_{L^1(\mathbb{R}^n)} \leq C\|Q_\delta^s\|_{L^1(\mathbb{R}^n)}\delta^{|\beta|} \leq C\delta^{|\beta|},$$

where the last inequality follows in view of Lemma 3.3.1 (i), implies

$$\left|\partial^\beta \widehat{Q}_\delta^s(\xi)\right| \leq C\delta^{|\beta|} \leq C \qquad \text{for all } \xi \in \mathbb{R}^n \text{ and } s \in [0,1).$$

To conclude, we use again the quotient rule to obtain

$$|\xi|^{|\alpha|}\left|\partial^\alpha m_t^s(\xi)\right| = |\xi|^{|\alpha|}\left|\partial^\alpha\left(\frac{\widehat{Q}_\delta^s(\xi)}{\widehat{Q}_\delta^t(\xi)}\right)\right| \leq \frac{R^{|\alpha|}C}{c^{2^{|\alpha|}}} = C \qquad \text{for } |\xi| \leq R.$$

□

We now obtain the next corollary based on the previous lemma; recall the definitions of $H^{1,p,\delta}(\mathbb{R}^n)$ and $H_0^{1,p,\delta}(\Omega)$ in (3.17) and (3.19).

Corollary 3.3.4. *Let $0 \leq s \leq t \leq 1$ and $p \in (1,\infty)$. If $u \in H^{t,p,\delta}(\mathbb{R}^n)$, then $u \in H^{s,p,\delta}(\mathbb{R}^n)$ and there is a constant $C > 0$ depending only on n, δ and p such that*

$$\|D_\delta^s u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \leq C\|D_\delta^t u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)}. \tag{3.38}$$

If $\Omega \subset \mathbb{R}^n$ is open and bounded and $u \in H_0^{t,p,\delta}(\Omega)$, then $u \in H_0^{s,p,\delta}(\Omega)$ with

$$\|D_\delta^s u\|_{L^p(\Omega;\mathbb{R}^n)} \leq C\|D_\delta^t u\|_{L^p(\Omega;\mathbb{R}^n)}.$$

Proof. The case $t = 1$ is covered by (3.31). For the other cases, we deduce from the previous lemma that the map

$$M_t^s : \mathcal{S}(\mathbb{R}^n;\mathbb{R}^n) \to L^p(\mathbb{R}^n;\mathbb{R}^n), \qquad v \mapsto (m_t^s\widehat{v})^\vee = \left(\frac{\widehat{Q}_\delta^s}{\widehat{Q}_\delta^t}\widehat{v}\right)^\vee$$

can be extended to a bounded linear operator on $L^p(\mathbb{R}^n;\mathbb{R}^n)$ with

$$\|M_t^s v\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \le C\|v\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \tag{3.39}$$

for all $v \in L^p(\mathbb{R}^n;\mathbb{R}^n)$, where $C > 0$ is a constant independent of s, t. For $\varphi \in C_c^\infty(\mathbb{R}^n)$, we also observe using Proposition 3.2.2 that

$$M_t^s D_\delta^t \varphi = M_t^s(Q_\delta^t * \nabla\varphi) = \left(\frac{\widehat{Q}_\delta^s}{\widehat{Q}_\delta^t}\widehat{Q}_\delta^t\widehat{\nabla\varphi}\right)^\vee = \left(\widehat{Q}_\delta^s\widehat{\nabla\varphi}\right)^\vee = D_\delta^s\varphi.$$

With $u \in H^{t,p,\delta}(\mathbb{R}^n)$, one can take an approximating sequence $\{\varphi_j\}_{j\in\mathbb{N}} \subset C_c^\infty(\mathbb{R}^n)$ with $\varphi_j \to u$ in $H^{t,p,\delta}(\mathbb{R}^n)$ and infer from the continuity of the operator M_t^s that $D_\delta^s\varphi_j = M_t^s D_\delta^t\varphi_j \to M_t^s D_\delta^t u$ in $L^p(\mathbb{R}^n)$. This shows that $u \in H^{s,p,\delta}(\mathbb{R}^n)$ with $D_\delta^s u = M_t^s D_\delta^t u \in L^p(\mathbb{R}^n)$. The bound (3.38) follows now from (3.39).

Finally, the statement for $u \in H_0^{t,p,\delta}(\Omega)$ follows by extending u to $\mathbb{R}^n$ by zero, noting that then the nonlocal gradient of u is zero in Ω^c. □

Remark 3.3.5. a) We note that this approach does not extend to $p = 1$, since the Mihlin-Hörmander theorem is not valid in this case. Moreover, this approach does not apply to $u \in H^{t,p,\delta}(\Omega)$ because it requires functions to be defined on all of $\mathbb{R}^n$ for the Fourier transform techniques. In fact, there is no obvious way of how to extend functions in $H^{t,p,\delta}(\Omega)$, as they can be ill-behaved in the strip $\Omega_\delta \setminus \Omega$.

b) An inequality of the type (3.38) does not hold for the fractional gradient, which can be seen from the homogeneity property. Indeed, for $u \in C_c^\infty(\mathbb{R}^n)$ and $0 \le s < t \le 1$, we may define for $\lambda > 0$ the function $u_\lambda := \lambda^{n/p-t}u(\lambda\,\cdot)$. Then, we can calculate that for $x \in \mathbb{R}^n$

$$D^t u_\lambda(x) = \lambda^{n/p} D^t u(\lambda x) \quad \text{and} \quad D^s u_\lambda(x) = \lambda^{n/p-(t-s)} D^s u(\lambda x).$$

This gives $\|D^t u_\lambda\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} = \|D^t u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)}$, whereas $\|D^s u_\lambda\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} = \lambda^{-(t-s)}\|D^s u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)}$. Letting $\lambda \to 0$ shows that (3.38) cannot hold for the fractional gradient. △

As a consequence, we derive the following generalization of the convergence result Lemma 3.3.2 to the nonlocal Sobolev setting.

Theorem 3.3.6. *Let $p \in (1,\infty)$ and let $\{s_j\}_{j\in\mathbb{N}} \subset [0,1]$ be a sequence converging to $s \in [0,1]$ with $\bar{s} := \sup_{j\in\mathbb{N}} s_j$. Then, it holds for every $u \in H^{\bar{s},p,\delta}(\mathbb{R}^n)$ that*

$$D_\delta^{s_j} u \to D_\delta^s u \ \text{ in } L^p(\mathbb{R}^n;\mathbb{R}^n) \text{ as } j \to \infty.$$

If $\Omega \subset \mathbb{R}^n$ is open and bounded and $u \in H_0^{\bar{s},p,\delta}(\Omega)$, then

$$D_\delta^{s_j} u \to D_\delta^s u \ \text{ in } L^p(\Omega;\mathbb{R}^n) \text{ as } j \to \infty.$$

Proof. Take $\varepsilon > 0$ and $\varphi_\varepsilon \in C_c^\infty(\mathbb{R}^n)$ such that $\|u - \varphi_\varepsilon\|_{H^{\bar{s},p,\delta}(\mathbb{R}^n)} \le \varepsilon$, cf. Theorem 3.2.8. Then, due to Corollary 3.3.4,

$$\|D_\delta^{s_j}(u-\varphi_\varepsilon)\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \le C\varepsilon \quad \text{for all } j \in \mathbb{N} \quad \text{and} \quad \|D_\delta^s(u-\varphi_\varepsilon)\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \le C\varepsilon.$$

If we choose j large enough so that $\|D_\delta^s\varphi_\varepsilon - D_\delta^{s_j}\varphi_\varepsilon\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \le \varepsilon$, which is possible by Lemma 3.3.2, we obtain

$$\begin{aligned}\|D_\delta^s u - D_\delta^{s_j} u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} &\le \|D_\delta^s(u-\varphi_\varepsilon)\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} + \|D_\delta^s\varphi_\varepsilon - D_\delta^{s_j}\varphi_\varepsilon\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \\ &\quad + \|D_\delta^{s_j}(u-\varphi_\varepsilon)\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \\ &\le (2C+1)\varepsilon,\end{aligned}$$

and letting $\varepsilon \to 0$ yields the desired convergence. The case $u \in H_0^{\bar{s},p,\delta}(\Omega)$ follows again via extension. □

Remark 3.3.7. For the particular case of localization to the classical gradient, i.e., when $s_j \to 1$ as $j \to \infty$, the convergence $D_\delta^{s_j} u \to \nabla u$ in $L^p(\Omega;\mathbb{R}^n)$ with $u \in W^{1,p}(\Omega_\delta)$ holds without imposing complementary values. Indeed, by Proposition 3.2.2 and Lemma 3.3.1, we can bound $\|D_\delta^s u\|_{L^p(\Omega;\mathbb{R}^n)} \leq C\|\nabla u\|_{L^p(\Omega_\delta;\mathbb{R}^n)}$ uniformly in s, and then, a similar argument to that of the proof of Theorem 3.3.6 applies. △

As another consequence of (3.38), we establish a nonlocal Poincaré inequality with a constant independent of the fractional order s. The proof builds on two pillars, namely the estimate of Corollary 3.3.4, which says that it is enough to prove the inequality for $s = 0$, and in order to achieve the latter, a version of the fundamental theorem of calculus for the case $s = 0$ from Proposition 3.2.9.

Theorem 3.3.8 (Nonlocal Poincaré inequality with uniform constants in s). *Let $s \in [0,1]$, $p \in (1,\infty)$ and $\Omega \subset \mathbb{R}^n$ be open and bounded. Then, there exists a constant $C > 0$ depending only on Ω, δ and p such that for all $u \in H_0^{s,p,\delta}(\Omega)$,*

$$\|u\|_{L^p(\Omega)} \leq C\|D_\delta^s u\|_{L^p(\Omega;\mathbb{R}^n)}. \tag{3.40}$$

Proof. Given Corollary 3.3.4, it suffices to prove (3.40) for $s = 0$. Moreover, we may assume by density (cf. (3.18)) that $u \in C_c^\infty(\Omega_{-\delta})$. Proposition 3.2.9 together with the fact that $\operatorname{supp}(D_\delta^0 u) \subset \Omega$ then implies

$$\|u\|_{L^p(\Omega)} \leq \|RD_\delta^0 u\|_{L^p(\mathbb{R}^n)} + |\Omega|\|W_\delta\|_{L^\infty(\mathbb{R}^n;\mathbb{R}^n)}\|D_\delta^0 u\|_{L^p(\Omega;\mathbb{R}^n)} \leq C\|D_\delta^0 u\|_{L^p(\Omega;\mathbb{R}^n)},$$

where the second inequality uses the L^p-boundedness of the Riesz transform. □

Finally, we present a compactness statement for sequences that are bounded in nonlocal spaces of different order. It will be used later in the proof of the Γ-convergence result in Section 3.6.

Lemma 3.3.9 (Weak compactness of sequences in varying order nonlocal spaces). *Let $p \in (1,\infty)$ and $\Omega \subset \mathbb{R}^n$ be open and bounded with $\Omega_{-\delta}$ a Lipschitz domain. Consider any sequence $\{s_j\}_{j\in\mathbb{N}} \subset [0,1]$ converging to $s \in [0,1]$ and $u_j \in H_0^{s_j,p,\delta}(\Omega)$ for $j \in \mathbb{N}$ with*

$$\sup_{j\in\mathbb{N}} \|D_\delta^{s_j} u_j\|_{L^p(\Omega;\mathbb{R}^n)} < \infty.$$

Then, up to a non-relabeled subsequence, $u_j \rightharpoonup u$ in $L^p(\Omega_\delta)$ with $u \in H_0^{s,p,\delta}(\Omega)$ and as $j \to \infty$,

$$D_\delta^{s_j} u_j \rightharpoonup D_\delta^s u \ \text{ in } L^p(\Omega;\mathbb{R}^n) \quad \text{and} \quad D_\delta^{s_j} u_j(x) \to D_\delta^s u(x) \ \text{ for a.e. } x \in \Omega \setminus \Omega_{-\delta}.$$

Additionally, if $s > 0$ then also $u_j \to u$ in $L^p(\Omega_\delta)$.

Proof. In view of the Poincaré inequality of Theorem 3.3.8, we observe that

$$\sup_{j\in\mathbb{N}} \|u_j\|_{L^p(\Omega)} \leq C \sup_{j\in\mathbb{N}} \|D_\delta^{s_j} u_j\|_{L^p(\Omega;\mathbb{R}^n)} < \infty.$$

Therefore, we can extract a subsequence of $\{u_j\}_{j\in\mathbb{N}}$ (non-relabeled) and find $u \in L^p(\Omega_\delta)$ and $V \in L^p(\Omega;\mathbb{R}^n)$ such that

$$u_j \rightharpoonup u \ \text{ in } L^p(\Omega_\delta) \quad \text{and} \quad D_\delta^{s_j} u_j \rightharpoonup V \ \text{ in } L^p(\Omega;\mathbb{R}^n)$$

as $j \to \infty$. Note that $u = 0$ in $\Omega_\delta \setminus \Omega_{-\delta}$, since the same holds for the functions u_j. To show that $u \in H_0^{s,p,\delta}(\Omega)$ and $V = D_\delta^s u$, take $\varphi \in C_c^\infty(\Omega; \mathbb{R}^n)$ and observe that

$$\int_\Omega V \cdot \varphi \, dx = \lim_{j\to\infty} \int_\Omega D_\delta^{s_j} u_j \cdot \varphi \, dx = -\lim_{j\to\infty} \int_{\Omega_\delta} u_j \operatorname{div}_\delta^{s_j} \varphi \, dx = -\int_{\Omega_\delta} u \operatorname{div}_\delta^s \varphi \, dx,$$

where the last equality results from the weak convergence $u_j \rightharpoonup u$ in $L^p(\Omega_\delta)$ and the uniform convergence $\operatorname{div}_\delta^{s_j} \varphi \to \operatorname{div}_\delta^s \varphi$ by Lemma 3.3.2. Hence, $u \in H^{s,p,\delta}(\Omega)$ with $D_\delta^s u = V$, and Proposition 3.2.10 implies $u \in H_0^{s,p,\delta}(\Omega)$, since $u = 0$ a.e. in $\Omega_\delta \setminus \Omega_{-\delta}$ and $\Omega_{-\delta}$ is Lipschitz.

It remains to prove the pointwise convergence of the nonlocal gradients outside of $\Omega_{-\delta}$. To this end, we observe in view of Remark 3.2.1 that for any $t \in [0,1]$ and $v \in H_0^{t,p,\delta}(\Omega)$,

$$D_\delta^t v(x) = \begin{cases} (d_\delta^t * v)(x) & \text{if } t \in [0,1), \\ 0 & \text{if } t = 1, \end{cases} \tag{3.41}$$

for a.e. $x \in \Omega \setminus \Omega_{-\delta}$; note that $|\partial\Omega_{-\delta}| = 0$, so that this set may be ignored. If $s \neq 1$, it holds for any $\varepsilon > 0$ that $d_\delta^{s_j} \to d_\delta^s$ uniformly on $B_\varepsilon(0)^c$ as $j \to \infty$. Consequently,

$$\lim_{j\to\infty} D_\delta^{s_j} u_j(x) = \lim_{j\to\infty} \int_{\Omega_{-\delta}} u_j(y) d_\delta^{s_j}(x-y) \, dy = \int_{\Omega_{-\delta}} u(y) d_\delta^s(x-y) \, dy = D_\delta^s u(x)$$

for a.e. $x \in \Omega \setminus \Omega_{-\delta}$. In the case $s = 1$, we have $d_\delta^{s_j} \to 0$ uniformly on $B_\varepsilon(0)^c$ as $j \to \infty$ due to the convergence $c_{n,s_j} \to 0$. The same argument then yields the desired pointwise convergence in light of (3.41).

Finally, if $s > 0$, one may assume without loss of generality that $\underline{s} := \inf_{j\in\mathbb{N}} s_j > 0$. We can then exploit the continuous embeddings $H_0^{s_j,p,\delta}(\Omega) \hookrightarrow H_0^{\underline{s},p,\delta}(\Omega)$ for $j \in \mathbb{N}$ with uniform constants, which follow in light of Corollary 3.3.4, to deduce that $u_j \rightharpoonup u$ in $H_0^{\underline{s},p,\delta}(\Omega)$. Then, $u_j \to u$ in $L^p(\Omega_\delta)$ by the compactness result in [31, Theorem 6.1 and 7.3]. □

3.4 Weak lower semicontinuity and existence theory

This section is devoted to characterizing the weak lower semicontinuity of integral functionals depending on the nonlocal gradient, that is, functionals of the form

$$\mathcal{F}(u) = \int_\Omega f(x, u(x), D_\delta^s u(x)) \, dx \quad \text{for } u \in H_g^{s,p,\delta}(\Omega; \mathbb{R}^m), \tag{3.42}$$

where $s \in (0,1)$, $p \in (1,\infty)$, $\Omega \subset \mathbb{R}^n$ is open and bounded, $g \in H^{s,p,\delta}(\Omega; \mathbb{R}^m)$, and $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a suitable integrand with p-growth. Using the connection between the nonlocal gradient and the classical gradient from Theorem 3.2.13, we can employ a translation procedure along the lines of [140] to conclude that the weak lower semicontinuity of $\mathcal{F}$ is equivalent to the quasiconvexity of f in its third argument. In fact, the quasiconvexity is only required in $\Omega_{-\delta}$, which is due to the strong convergence of the nonlocal gradient in $\Omega_\delta \setminus \Omega_{-\delta}$ from Lemma 3.2.12.

Theorem 3.4.1 (Characterization of weak lower semicontinuity). *Let $s \in (0,1)$, $p \in (1,\infty)$, $\Omega \subset \mathbb{R}^n$ be open and bounded with $|\partial\Omega_{-\delta}| = 0$ and $g \in H^{s,p,\delta}(\Omega; \mathbb{R}^m)$. Further, let $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory function satisfying*

$$-C(1 + |z|^p + |A|^q) \leq f(x, z, A) \leq C(1 + |z|^p + |A|^p)$$

for a.e. $x \in \Omega$ and all $(z, A) \in \mathbb{R}^m \times \mathbb{R}^{m\times n}$ with $C > 0$ and $q \in [1, p)$.

Then, $\mathcal{F}$ from (3.42) *is weakly lower semicontinuous on $H_g^{s,p,\delta}(\Omega;\mathbb{R}^m)$ if and only if*

$$f(x,z,\cdot) \quad \text{is quasiconvex for a.e. } x \in \Omega_{-\delta} \text{ and all } z \in \mathbb{R}^m, \tag{3.43}$$

i.e., it holds for a.e. $x \in \Omega_{-\delta}$ and all $z \in \mathbb{R}^m$ with $Y = (0,1)^n$ that

$$f(x,z,A) \leq \int_Y f(x,z,A+\nabla\varphi(y))\,dy \quad \text{for all } \varphi \in W_0^{1,\infty}(Y;\mathbb{R}^m) \text{ and } A \in \mathbb{R}^{m\times n}.$$

Proof. The proof follows the lines of [140, Theorem 4.1 and 4.5], we detail the differences for the reader's convenience.

Step 1: Sufficiency. Assuming (3.43), let $\{u_j\}_{j\in\mathbb{N}} \subset H_g^{s,p,\delta}(\Omega;\mathbb{R}^m)$ be a sequence that converges weakly to u in $H_g^{s,p,\delta}(\Omega;\mathbb{R}^m)$. We divide the proof by splitting the integral functional $\mathcal{F}$ and considering separately the integral contributions over $\Omega_{-\delta}$ and $\Omega \setminus \Omega_{-\delta}$.

Since $u_j \to u$ in $L^p(\Omega_\delta;\mathbb{R}^m)$ by [31, Theorem 7.3] and $\mathcal{Q}_\delta^s u_j \rightharpoonup \mathcal{Q}_\delta^s u$ in $W^{1,p}(\Omega;\mathbb{R}^m)$ by Theorem 3.2.13 (i), we conclude

$$\begin{aligned} \int_{\Omega_{-\delta}} f(x,u,D_\delta^s u)\,dx &= \int_{\Omega_{-\delta}} f(x,u,\nabla \mathcal{Q}_\delta^s u)\,dx \\ &\leq \liminf_{j\to\infty} \int_{\Omega_{-\delta}} f(x,u_j,\nabla \mathcal{Q}_\delta^s u_j)\,dx \\ &= \liminf_{j\to\infty} \int_{\Omega_{-\delta}} f(x,u_j,D_\delta^s u_j)\,dx, \end{aligned} \tag{3.44}$$

where the inequality is due to the quasiconvexity and p-growth of f, with the exact argument of [140, Theorem 4.1] involving Young measures. Note that this requires the negative part of the sequence $\{f(\cdot,u_j,\nabla \mathcal{Q}_\delta^s u_j)\}_{j\in\mathbb{N}}$ to be equi-integrable, which is guaranteed by the lower bound on f.

Secondly, for the integral on $\Omega \setminus \Omega_{-\delta}$, we invoke from Lemma 3.2.12 the convergence

$$D_\delta^s u_j \to D_\delta^s u \in L^p(\Omega \setminus O;\mathbb{R}^{m\times n})$$

for any $O \Supset \Omega_{-\delta}$. Hence, a well-known strong lower semicontinuity result (e.g., [112, Theorem 6.49]) yields

$$\int_{\Omega\setminus O} f(x,u,D_\delta^s u)\,dx \leq \liminf_{j\to\infty} \int_{\Omega\setminus O} f(x,u_j,D_\delta^s u_j)\,dx.$$

Letting $O \downarrow \Omega_{-\delta}$ implies, using once again the equi-integrability of the negative part of the sequence $\{f(\cdot,u_j,D_\delta^s u_j)\}_{j\in\mathbb{N}}$ and the assumption $|\partial\Omega_{-\delta}| = 0$, that

$$\int_{\Omega\setminus\Omega_{-\delta}} f(x,u,D_\delta^s u)\,dx \leq \liminf_{j\to\infty} \int_{\Omega\setminus\Omega_{-\delta}} f(x,u_j,D_\delta^s u_j)\,dx. \tag{3.45}$$

The sufficiency now follows from adding (3.44) and (3.45).

Step 2: Necessity. Analogously to the proof of [140, Theorem 4.5], we may assume without loss of generality that $g = 0$. In order to prove the stated quasiconvexity of f, let us fix $(x_0, z_0, A_0) \in \Omega_{-\delta} \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$. Using Lemma 3.4.2, we may select a $\varphi_0 \in C_c^\infty(\Omega_{-\delta};\mathbb{R}^m)$ such that

$$\varphi_0(x_0) = z_0 \qquad \text{and} \qquad D_\delta^s \varphi_0(x_0) = A_0. \tag{3.46}$$

Consider any $\varphi \in W_0^{1,\infty}(Y;\mathbb{R}^m)$ and assume that $x_0 + Y \Subset \Omega_{-\delta}$; the latter can be done without loss of generality in light of the scaling and translation invariances related to the definition of quasiconvexity, see e.g., [75, Proposition 5.11]. If we fix $\rho \in (0,1)$ and periodically extend φ to $\mathbb{R}^n$, we can define the sequence $\{\varphi_j^\rho\}_{j\in\mathbb{N}} \subset W^{1,\infty}(\mathbb{R}^n;\mathbb{R}^m)$ by

$$\varphi_j^\rho(x) = \begin{cases} \frac{\rho}{j}\varphi\Big(j\frac{(x-x_0)}{\rho}\Big) & \text{for } x \in Y_\rho := x_0 + (0,\rho)^n, \\ 0 & \text{otherwise,} \end{cases} \qquad x \in \mathbb{R}^n.$$

As this is a periodically oscillating sequence that converges to zero essentially uniformly, we find that $\varphi_j^\rho \rightharpoonup 0$ in $W^{1,p}(\mathbb{R}^n;\mathbb{R}^m)$ as $j \to \infty$.

Take a cut-off function $\chi \in C_c^\infty(\Omega_{-\delta};[0,1])$ with $\chi \equiv 1$ on $x_0 + Y$ and define the sequence $\{u_j\}_{j\in\mathbb{N}} \subset H_0^{s,p,\delta}(\Omega;\mathbb{R}^m)$ given by

$$u_j := \varphi_0 + \chi \mathcal{P}_\delta^s \varphi_j^\rho,$$

which converges weakly to φ_0 in $H_0^{s,p,\delta}(\Omega;\mathbb{R}^m)$ in light of the continuity of $\mathcal{P}_\delta^s$ in Theorem 3.2.13 (ii). In particular, we have $u_j \to \varphi_0$ in $L^p(\Omega_\delta;\mathbb{R}^m)$ by [31, Theorem 7.3]. Moreover, it holds by the Leibniz rule in Lemma 3.2.11 and the fact that $D_\delta^s \mathcal{P}_\delta^s \varphi_j^\rho = \nabla \varphi_j^\rho$ that

$$D_\delta^s u_j = D_\delta^s \varphi_0 + \chi \nabla \varphi_j^\rho + K_\chi(\mathcal{P}_\delta^s \varphi_j^\rho).$$

Observe that $K_\chi(\mathcal{P}_\delta^s \varphi_j^\rho) \to 0$ in $L^p(\Omega;\mathbb{R}^{m\times n})$ as $j \to \infty$ due to the boundedness of K_χ and that $\chi\nabla\varphi_j^\rho = \nabla\varphi_j^\rho$ on Ω since φ_j^ρ is zero outside Y_ρ.

Finally, we exploit the weak lower semicontinuity of $\mathcal{F}$ on $H_0^{s,p,\delta}(\Omega;\mathbb{R}^m)$ to derive

$$\begin{aligned} \int_\Omega f(x,\varphi_0,D_\delta^s\varphi_0)\,dx &\le \liminf_{j\to\infty} \int_\Omega f(x,u_j,D_\delta^s u_j)\,dx \\ &= \liminf_{j\to\infty} \int_{Y_\rho} f(x,u_j,D_\delta^s\varphi_0 + \nabla\varphi_j^\rho + K_\chi(\mathcal{P}_\delta^s\varphi_j^\rho))\,dx \\ &\qquad + \int_{\Omega\setminus Y_\rho} f(x,u_j,D_\delta^s\varphi_0 + K_\chi(\mathcal{P}_\delta^s\varphi_j^\rho))\,dx \\ &\le \liminf_{j\to\infty} \int_{Y_\rho} f(x,\varphi_0,D_\delta^s\varphi_0 + \nabla\varphi_j^\rho)\,dx + \int_{\Omega\setminus Y_\rho} f(x,\varphi_0,D_\delta^s\varphi_0)\,dx, \end{aligned}$$

where the last inequality uses [140, Lemma 4.10] to remove all the terms that converge strongly to zero. In view of the p-growth of f, the integral over $\Omega \setminus Y_\rho$ is finite, so that subtracting it from both sides gives

$$\int_{Y_\rho} f(x,\varphi_0,D_\delta^s\varphi_0)\,dx \le \liminf_{j\to\infty} \int_{Y_\rho} f(x,\varphi_0,D_\delta^s\varphi_0 + \nabla\varphi_j^\rho)\,dx.$$

Because φ_0 and $D_\delta^s\varphi_0$ are continuous and satisfy (3.46), the rest of the proof follows by mimicking Steps 2-4 of [140, Theorem 4.5]. □

The following lemma was used in the previous proof and shows that one can construct smooth functions with compact support whose nonlocal gradient has a desired value at a point. The proof is omitted here, as it is nearly identical to [140, Lemma 4.3], given that w_δ is radial by (H1).

Lemma 3.4.2. *Let $s \in (0,1)$ and let $\Omega \subset \mathbb{R}^n$ be open and bounded. For any $x_0 \in \Omega_{-\delta}$, $z \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$, there exists a $\varphi \in C_c^\infty(\Omega_{-\delta};\mathbb{R}^m)$ such that $\varphi(x_0) = z$ and $D_\delta^s\varphi(x_0) = A$.*

With the perspective of Section 3.2.5, there is an alternative approach to proving Theorem 3.4.1 that passes through the characterization of weak lower semicontinuity of functionals depending on Riesz fractional gradients from [140]. For simplicity, we take $g = 0$ and drop the dependence on x and z in f.

Alternative proof of Theorem 3.4.1. Step 1: Sufficiency. Let $\{u_j\}_{j\in\mathbb{N}} \subset H_0^{s,p,\delta}(\Omega;\mathbb{R}^m)$ converge weakly in $H_0^{s,p,\delta}(\Omega;\mathbb{R}^m)$ to the limit function $u \in H_0^{s,p,\delta}(\Omega;\mathbb{R}^m)$. As a quasiconvex function, $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is also rank-one convex and hence, locally Lipschitz continuous in the sense that

$$|f(A) - f(B)| \leq C(1 + |A|^{p-1} + |B|^{p-1})|A - B| \qquad \text{for all } A, B \in \mathbb{R}^{m\times n}$$

with a constant $C > 0$, cf. e.g., [75, Proposition 2.32].

Consider the auxiliary function

$$h_u(x, A) = \mathbb{1}_\Omega(x) f\big(A + (\nabla R_\delta^s * u)(x)\big) \quad \text{ for } x \in \mathbb{R}^n \text{ and } A \in \mathbb{R}^{m\times n},$$

which is Carathéodory, quasiconvex in the second variable, and satisfies the growth bound

$$|h_u(x, A)| \leq C\big(1 + |A|^p + |(\nabla R_\delta^s * u)(x)|^p\big) \leq C\big(1 + |A|^p + \|u\|^p_{L^p(\Omega_\delta;\mathbb{R}^m)}\big),$$

with the last step using the boundedness of ∇R_δ^s. By the local Lipschitz continuity of f, we also find

$$\begin{aligned}&\Big|\int_\Omega f(D_\delta^s u_j)\, dx - \int_{\mathbb{R}^n} h_u(x, D^s u_j)\, dx\Big| \\ &\qquad \leq C\big(1 + \|D_\delta^s u_j\|_{L^p(\Omega;\mathbb{R}^{m\times n})} + \|u_j\|_{L^p(\Omega_\delta;\mathbb{R}^m)} + \|u\|_{L^p(\Omega_\delta;\mathbb{R}^m)}\big)\|u_j - u\|_{L^p(\Omega_\delta;\mathbb{R}^m)} \to 0\end{aligned}$$

as $j \to \infty$. Since $u_j \rightharpoonup u$ in $H_0^{s,p}(\Omega_{-\delta};\mathbb{R}^m)$ by Lemma 3.2.16 and h_u fulfills the requirements of [140, Theorem 4.1], the desired lower semicontinuity results from

$$\begin{aligned}\liminf_{j\to\infty} \mathcal{F}(u_j) = \liminf_{j\to\infty} \int_\Omega f(D_\delta^s u_j)\, dx &= \liminf_{j\to\infty} \int_{\mathbb{R}^n} h_u(x, D^s u_j)\, dx \\ &\geq \int_{\mathbb{R}^n} h_u(x, D^s u)\, dx = \int_\Omega f(D_\delta^s u)\, dx = \mathcal{F}(u).\end{aligned}$$

Step 2: Necessity. Suppose $\mathcal{F}$ is weakly lower semicontinuous on $H_0^{s,p,\delta}(\Omega;\mathbb{R}^m)$. Fix $(x_0, A_0) \in \Omega_{-\delta} \times \mathbb{R}^{m\times n}$ and using Lemma 3.4.2, let $\varphi \in C_c^\infty(\Omega_{-\delta};\mathbb{R}^m)$ be such that $D_\delta^s\varphi(x_0) = A_0$. A similar reasoning as in Step 1 shows for any sequence $\{u_j\}_{j\in\mathbb{N}} \subset H_0^{s,p}(\Omega_{-\delta};\mathbb{R}^m)$ converging weakly in $H_0^{s,p}(\Omega_{-\delta};\mathbb{R}^m)$ to φ and with $\{D^s u_j\}_{j\in\mathbb{N}}$ p-equi-integrable that

$$\liminf_{j\to\infty} \int_{\mathbb{R}^n} h_\varphi(x, D^s u_j)\, dx \geq \liminf_{j\to\infty} \int_\Omega f(D_\delta^s u_j)\, dx \geq \int_{\mathbb{R}^n} h_\varphi(x, D^s\varphi)\, dx,$$

where the first inequality uses the p-equi-integrability of $\{D^s u_j\}_{j\in\mathbb{N}}$ and the strong convergence $\nabla R_\delta^s * u_j \to \nabla R_\delta^s * \varphi$ in $L^p(\Omega;\mathbb{R}^{m\times n})$ to apply a well-known freezing lemma (see e.g., [140, Lemma 4.10]). The proof of [140, Theorem 4.5] then yields for all $v \in W_0^{1,\infty}(Y;\mathbb{R}^m)$,

$$h_\varphi(x_0, A_0) \leq \int_Y h_\varphi(x_0, A_0 + \nabla v)\, dy.$$

If we further suppose that $\mathrm{supp}(\varphi) \subset B(x_0, b_0\delta)$, which is possible by Lemma 3.4.2, then $(\nabla R_\delta^s * \varphi)(x_0) = 0$ since $\nabla R_\delta^s = 0$ in $B(0, b_0\delta)$, see (3.74). Therefore, the inequality turns into

$$f(A_0) \leq \int_Y f(A_0 + \nabla v)\, dy,$$

as desired. □

Let us briefly comment on the role of quasiconvexity in Theorem 3.4.1, especially in relation with a new generalized convexity notion that can be considered natural in our nonlocal setting. For simplicity, we assume that f is constant in the x- and z-variables.

Remark 3.4.3 (D_δ^s-quasiconvexity). Let $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ be a measurable function. We call f D_δ^s-quasiconvex if for every $A \in \mathbb{R}^{m\times n}$,

$$f(A) \le \int_Y f\big(A + D_\delta^s\psi(y)\big)\,dy \qquad \text{for all } \psi \in H_\#^{s,\infty,\delta}(Y;\mathbb{R}^m), \tag{3.47}$$

whenever the integral on the left-hand side exists.

Under consideration of Remark 3.2.14 c) and by using the characterization of quasiconvexity with periodic test functions (see e.g., [75, Proposition 5.13]), it follows immediately that D_δ^s-quasiconvexity is equivalent to the usual quasiconvexity. An analogous result for fractional instead of nonlocal gradients was established in [140], by showing equivalence of quasiconvexity with α-quasiconvexity, where $\alpha = s$.

In fact, opposed to the fractional case, one can show here that for continuous $f : \mathbb{R}^{m\times n} \to \mathbb{R}$, the periodic test functions in (3.47) can equivalently be replaced by test functions in the complementary-value space $H_0^{s,\infty,\delta}(\Omega;\mathbb{R}^m)$ for any open and bounded $\Omega \subset \mathbb{R}^n$. The resulting generalized convexity notion defined by

$$f(A) \le \frac{1}{|\Omega|}\int_\Omega f\big(A + D_\delta^s\psi(y)\big)\,dy \qquad \text{for all } \psi \in H_0^{s,\infty,\delta}(\Omega;\mathbb{R}^m) \text{ and } A \in \mathbb{R}^{m\times n} \tag{3.48}$$

is indeed equivalent to quasiconvexity as well. To see this, we observe first that quasiconvexity implies (3.48) since it holds for any $\psi \in H_0^{s,\infty,\delta}(\Omega;\mathbb{R}^m)$ that

$$\varphi := \mathcal{Q}_\delta^s\psi \in W_0^{1,\infty}(\Omega;\mathbb{R}^m) \quad \text{ with } \nabla\varphi = D_\delta^s\psi,$$

cf. Remark 3.2.14. Conversely, for any given $\varphi \in W_0^{1,\infty}(Y;\mathbb{R}^m)$, one can consider the sequence $\{\psi_j\}_{j\in\mathbb{N}}$ given by

$$\psi_j := \chi\mathcal{P}_\delta^s\varphi_j^\rho \in H_0^{s,\infty,\delta}(\Omega;\mathbb{R}^m),$$

where $\rho \in (0,1)$ is fixed and φ_j^ρ and the cut-off function χ are as in Step 2 of the proof of Theorem 3.4.1. Then, with similar arguments and the compact embedding of $H_0^{s,p,\delta}(\Omega;\mathbb{R}^m)$ into $L^\infty(\Omega_\delta;\mathbb{R}^m)$ for $p > n/s$, see [31, Theorem 7.3], we obtain $D_\delta^s\psi_j - \nabla\varphi_j^\rho \to 0$ in $L^\infty(\Omega;\mathbb{R}^{m\times n})$ as $j \to \infty$. Using the continuity of f and (3.48) then implies

$$\begin{aligned} f(A) &\le \lim_{j\to\infty}\frac{1}{|\Omega|}\int_\Omega f\big(A + D_\delta^s\psi_j(y)\big)\,dy = \lim_{j\to\infty}\frac{1}{|\Omega|}\int_\Omega f\big(A + \nabla\varphi_j^\rho(y)\big)\,dy \\ &= \frac{|Y_\rho|}{|\Omega|}\int_Y f\big(A + \nabla\varphi(y)\big)\,dy + \frac{|\Omega| - |Y_\rho|}{|\Omega|}f(A), \end{aligned}$$

which shows the quasiconvexity of f. △

With the previous findings at hand, the following existence result is now a simple consequence of the direct method.

Corollary 3.4.4. *Let $s \in (0,1)$, $p \in (1,\infty)$, $\Omega \subset \mathbb{R}^n$ be open and bounded with $|\partial\Omega_{-\delta}| = 0$ and $g \in H^{s,p,\delta}(\Omega)$. Suppose that $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a Carathéodory function satisfying*

$$c|A|^p - C \le f(x,z,A) \le C(1 + |z|^p + |A|^p)$$

for a.e. $x \in \Omega$ and all $(z,A) \in \mathbb{R}^m \times \mathbb{R}^{m\times n}$ with constants $c, C > 0$. If $A \mapsto f(x,z,A)$ is quasiconvex for a.e. $x \in \Omega_{-\delta}$ and all $z \in \mathbb{R}^m$, then $\mathcal{F}$ as in (3.42) admits a minimizer in $H_g^{s,p,\delta}(\Omega;\mathbb{R}^m)$.

Proof. If $\{u_j\}_{j\in\mathbb{N}} \subset H^{s,p,\delta}_g(\Omega;\mathbb{R}^m)$ is a minimizing sequence for $\mathcal{F}$, we find by the coercivity bound on f that $\{D^s_\delta u_j\}_{j\in\mathbb{N}}$ is a bounded sequence in $L^p(\Omega;\mathbb{R}^{m\times n})$. By the nonlocal Poincaré inequality in [31, Theorem 6.2], it follows that $\{u_j\}_{j\in\mathbb{N}}$ is a bounded sequence in $H^{s,p,\delta}_g(\Omega;\mathbb{R}^m)$, so that, up to a non-relabeled subsequence, $u_j \rightharpoonup u$ in $H^{s,p,\delta}_g(\Omega;\mathbb{R}^m)$ for some $u \in H^{s,p,\delta}_g(\Omega;\mathbb{R}^m)$. Together with Theorem 3.4.1, this shows that u is a minimizer of $\mathcal{F}$ over $H^{s,p,\delta}_g(\Omega;\mathbb{R}^m)$. □

3.5 Homogenization and relaxation

In the next step, we aim to prove new relaxation and homogenization results for our nonlocal functionals. Both will follow as corollaries of a more general, abstract statement about the Γ-convergence of integral functionals with dependence on nonlocal gradients, which is of independent interest. Our approach relies on the connection between the nonlocal and classical gradient, as established in Section 3.2.4, in order to reduce the problem to a standard setting.

Throughout the section, let $s \in (0,1)$, $p \in (1,\infty)$ and $\Omega \subset \mathbb{R}^n$ be an open and bounded set with $|\partial\Omega_{-\delta}| = 0$, and $g \in H^{s,p,\delta}(\Omega;\mathbb{R}^m)$. Further, we assume that the readers are familiar with the basics of Γ-convergence, and refer to [49, 80] for a comprehensive introduction.

Let us start with some necessary notations in preparation for the announced abstract Γ-convergence result. For any Carathéodory integrand $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ with standard p-growth and p-coercivity, i.e., there are constants $C, c > 0$ such that

$$c|A|^p - C \leq f(x,A) \leq C(|A|^p + 1) \tag{3.49}$$

for a.e. $x \in \Omega$ and all $A \in \mathbb{R}^{m\times n}$, we define the three integral functionals $\mathcal{I}_f : L^p(\Omega_{-\delta};\mathbb{R}^m) \to \mathbb{R}_\infty$, $\mathcal{J}_f : L^p(\Omega \setminus \Omega_{-\delta};\mathbb{R}^{m\times n}) \to \mathbb{R}$ and $\mathcal{F}_f : L^p(\Omega_\delta;\mathbb{R}^m) \to \mathbb{R}_\infty$ as

$$\mathcal{I}_f(v) = \begin{cases} \displaystyle\int_{\Omega_{-\delta}} f(x,\nabla v)\,dx & \text{for } v \in W^{1,p}(\Omega_{-\delta};\mathbb{R}^m), \\ \infty & \text{otherwise,} \end{cases}$$

$$\mathcal{J}_f(V) = \int_{\Omega\setminus\Omega_{-\delta}} f(x,V)\,dx,$$

and

$$\mathcal{F}_f(u) = \begin{cases} \displaystyle\int_{\Omega} f(x,D^s_\delta u)\,dx & \text{for } u \in H^{s,p,\delta}_g(\Omega;\mathbb{R}^m), \\ \infty & \text{otherwise,} \end{cases}$$

respectively.

Theorem 3.5.1 (General Γ-convergence result). *Suppose $f_j, f_\infty : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ for $j \in \mathbb{N}$ are Carathéodory integrands satisfying* (3.49) *uniformly in j and*

$$|f_j(x,A) - f_j(x,B)| \leq M(1 + |A|^{p-1} + |B|^{p-1})|A - B| \tag{3.50}$$

for a.e. $x \in \Omega$, all $A, B \in \mathbb{R}^{m\times n}$ and all $j \in \mathbb{N}$ with a constant $M > 0$. If the sequence $\{\mathcal{I}_{f_j}\}_{j\in\mathbb{N}}$ converges to $\mathcal{I}_{f_\infty}$ in the sense of Γ-convergence regarding the strong topology in $L^p(\Omega_{-\delta};\mathbb{R}^m)$ as $j \to \infty$, in short,

$$\Gamma(L^p)\text{-}\lim_{j\to\infty} \mathcal{I}_{f_j} = \mathcal{I}_{f_\infty} \tag{3.51}$$

and

$$\mathcal{J}_{f_j} \to \mathcal{J}_{f_\infty} \text{ pointwise,} \tag{3.52}$$

then

$$\Gamma(L^p)\text{-}\lim_{j\to\infty} \mathcal{F}_{f_j} = \mathcal{F}_{f_\infty}, \tag{3.53}$$

that is, $\{\mathcal{F}_{f_j}\}_{j\in\mathbb{N}}$ *Γ-converges with respect to the strong topology in* $L^p(\Omega_\delta;\mathbb{R}^m)$ *to* $\mathcal{F}_{f_\infty}$ *as* $j\to\infty$.

Moreover, every sequence $\{u_j\}_{j\in N}\subset L^p(\Omega_\delta;\mathbb{R}^m)$ *with uniformly bounded energy* $\sup_j \mathcal{F}_{f_j}(u_j)<\infty$ *has a converging subsequence in* $L^p(\Omega_\delta;\mathbb{R}^m)$.

Proof. By adding a constant, we may assume without loss of generality that f_j for $j\in\mathbb{N}$ and f_∞ are non-negative. Further, we observe upfront that due to (3.50), the functionals $\mathcal{J}_{f_j}$ for $j\in\mathbb{N}$ are locally Lipschitz on $L^p(\Omega\setminus\Omega_{-\delta};\mathbb{R}^{m\times n})$ with a uniform Lipschitz constant. The pointwise convergence $\mathcal{J}_{f_j}\to\mathcal{J}_{f_\infty}$ in (3.52) is therefore equivalent to locally uniform convergence; in particular, it holds for any sequence $V_j\to V$ in $L^p(\Omega\setminus\Omega_{-\delta};\mathbb{R}^{m\times n})$ that

$$\lim_{j\to\infty}\mathcal{J}_{f_j}(V_j)=\mathcal{J}_{f_\infty}(V). \tag{3.54}$$

The rest of the proof is split into the usual steps, proving first compactness to obtain the add-on and then, the liminf-inequality and a complementary upper bound via the existence of recovery sequences, which in combination yields (3.53).

Step 1: Compactness. In view of the lower bound in (3.49), this is an immediate consequence of the Poincaré inequality and compactness result in $H^{s,p,\delta}_g(\Omega;\mathbb{R}^m)$ (cf. [31, Theorem 6.1 and 7.3]).

Step 2: Liminf-inequality. Let $\{u_j\}_{j\in\mathbb{N}}\subset L^p(\Omega_\delta;\mathbb{R}^m)$ be a convergent sequence for $j\to\infty$ with limit $u\in L^p(\Omega_\delta;\mathbb{R}^m)$. Suppose without loss of generality that $\liminf_{j\to\infty}\mathcal{F}_{f_j}(u_j)=\lim_{j\to\infty}\mathcal{F}_{f_j}(u_j)<\infty$. It follows then from the coercivity bound in (3.49) that $\{u_j\}_{j\in\mathbb{N}}\subset H^{s,p,\delta}_g(\Omega;\mathbb{R}^m)$ is bounded and thus,

$$u_j \rightharpoonup u \quad \text{in } H^{s,p,\delta}_g(\Omega;\mathbb{R}^m).$$

By Theorem 3.2.13 (i), it holds that $\mathcal{Q}^s_\delta u_j\rightharpoonup\mathcal{Q}^s_\delta u$ in $W^{1,p}(\Omega;\mathbb{R}^m)$ with $\nabla\mathcal{Q}^s_\delta u=D^s_\delta u$ and $\nabla\mathcal{Q}^s_\delta u_j=D^s_\delta u_j$ for $j\in\mathbb{N}$. Hence, the liminf-inequality from the Γ-convergence of $\{\mathcal{I}_{f_j}\}_{j\in\mathbb{N}}$ in (3.51) yields

$$\int_{\Omega_{-\delta}} f_\infty(x,D^s_\delta u)\,dx=\mathcal{I}_{f_\infty}(\mathcal{Q}^s_\delta u)\le\liminf_{j\to\infty}\mathcal{I}_{f_j}(\mathcal{Q}^s_\delta u_j)=\liminf_{j\to\infty}\int_{\Omega_{-\delta}} f_j(x,D^s_\delta u_j)\,dx. \tag{3.55}$$

Additionally, for any $O\Subset\mathbb{R}^n$ open with $\Omega_{-\delta}\Subset O$, it holds according to Lemma 3.2.12 that $\mathbb{1}_{\Omega\setminus O}D^s_\delta u_j\to\mathbb{1}_{\Omega\setminus O}D^s_\delta u$ in $L^p(\Omega\setminus\Omega_{-\delta};\mathbb{R}^{m\times n})$. We then find in view of (3.54) and the non-negativity of the functions f_j that

$$\mathcal{J}_{f^\infty}(\mathbb{1}_{\Omega\setminus O}D^s_\delta u)=\lim_{j\to\infty}\mathcal{J}_{f_j}(\mathbb{1}_{\Omega\setminus O}D^s_\delta u_j)\le\liminf_{j\to\infty}\mathcal{J}_{f_j}(D^s_\delta u_j)+\int_{O\setminus\Omega_{-\delta}} f_j(x,0)\,dx.$$

Due to $0\le f_j(\cdot,0)\le C$ for all $j\in\mathbb{N}$ and $|\partial\Omega_{-\delta}|=0$, one may let O tend to $\Omega_{-\delta}$ to conclude

$$\int_{\Omega\setminus\Omega_{-\delta}} f_\infty(x,D^s_\delta u)\,dx=\mathcal{J}_{f_\infty}(D^s_\delta u)\le\liminf_{j\to\infty}\mathcal{J}_{f_j}(D^s_\delta u_j)=\liminf_{j\to\infty}\int_{\Omega\setminus\Omega_{-\delta}} f_j(x,D^s_\delta u_j)\,dx.$$

Finally, combining this with (3.55) yields the desired liminf-inequality $\liminf_{j\to\infty}\mathcal{F}_{f_j}(u_j)\ge\mathcal{F}_{f_\infty}(u)$.

Step 3: Limsup-inequality. Take $u \in H^{s,p,\delta}_g(\Omega;\mathbb{R}^m)$ with $\mathcal{F}_{f_\infty}(u) < \infty$ and define $v = \mathcal{Q}^s_\delta u \in W^{1,p}(\Omega;\mathbb{R}^m)$, which satisfies $\nabla v = D^s_\delta u$ on Ω by Theorem 3.2.13 (i). We need to construct a recovery sequence $(u_j)_j \subset H^{s,p,\delta}_g(\Omega;\mathbb{R}^m)$ that converges to u weakly in $L^p(\Omega;\mathbb{R}^m)$ and satisfies

$$\limsup_{j\to\infty} \mathcal{F}_{f_j}(u_j) \leq \mathcal{F}_{f_\infty}(u). \tag{3.56}$$

To this end, let $\varepsilon > 0$ be fixed. The upper bound from the Γ-convergence of $\{\mathcal{I}_{f_j}\}_{j\in\mathbb{N}}$ to $\mathcal{I}_{f_\infty}$ in combination with an argument to enforce boundary conditions as in [80, Proof of Theorem 21.1] allows us to find a sequence $\{v_j\}_{j\in\mathbb{N}} \subset W^{1,p}(\Omega_{-\delta};\mathbb{R}^m)$ with the properties that $v_j = v$ in $\Omega_{-\delta} \setminus U_\varepsilon$ for all $j \in \mathbb{N}$ with some open $U_\varepsilon \Subset \Omega_{-\delta}$, $v_j \to v$ in $L^p(\Omega_{-\delta};\mathbb{R}^m)$, and

$$\limsup_{j\to\infty} \int_{\Omega_{-\delta}} f_j(x,\nabla v_j)\,dx \leq \int_{\Omega_{-\delta}} f_\infty(x,\nabla v)\,dx + \varepsilon < \infty. \tag{3.57}$$

As a consequence of (3.57) together with the coercivity bound in (3.49) and Poincaré's inequality, the sequence $\{v_j\}_{j\in\mathbb{N}}$ converges not only in L^p, but also weakly in $W^{1,p}$, that is,

$$v_j - v \rightharpoonup 0 \quad \text{in } W^{1,p}(\Omega_{-\delta};\mathbb{R}^m). \tag{3.58}$$

After extending $\{v_j - v\}_{j\in\mathbb{N}}$ by zero to a sequence in $W^{1,p}(\mathbb{R}^n;\mathbb{R}^m)$, we conclude from Theorem 3.2.13 (ii) that

$$\tilde{u}_j := \mathcal{P}^s_\delta(v_j - v) \in H^{s,p,\delta}(\Omega;\mathbb{R}^m)$$

satisfies $D^s_\delta \tilde{u}_j = \nabla(v_j - v)$ on Ω. Hence, under consideration of (3.58) and Lemma 3.2.15,

$$\tilde{u}_j \to 0 \text{ in } L^p(\Omega;\mathbb{R}^m) \quad \text{and} \quad \tilde{u}_j \rightharpoonup 0 \text{ in } H^{s,p,\delta}(\Omega;\mathbb{R}^m) \tag{3.59}$$

as $j \to \infty$. Considering a cut-off function $\chi \in C^\infty_c(\Omega_{-\delta})$ with $\chi \equiv 1$ on U_ε, we define

$$u_j := u + \chi\tilde{u}_j \in H^{s,p,\delta}_g(\Omega;\mathbb{R}^m) \text{ for } j \in \mathbb{N}.$$

Then, by the Leibniz rule in Lemma 3.2.11,

$$D^s_\delta u_j = D^s_\delta u + \chi D^s_\delta \tilde{u}_j + K_\chi(\tilde{u}_j) = \nabla v + \chi\nabla(v_j - v) + K_\chi(\tilde{u}_j)$$

for every $j \in \mathbb{N}$; note that, in particular, $D^s_\delta u_j = \nabla v_j + K_\chi(\tilde{u}_j)$ on U_ε, while $D^s_\delta u_j = D^s_\delta u + K_\chi(\tilde{u}_j)$ on $\Omega \setminus U_\varepsilon$ since $\nabla(v_j - v)$ is zero there. As $j \to \infty$, we have in view of (3.59) that

$$u_j \to u \text{ in } L^p(\Omega;\mathbb{R}^m) \quad \text{and} \quad K_\chi(\tilde{u}_j) \to 0 \text{ in } L^p(\Omega;\mathbb{R}^{m\times n}). \tag{3.60}$$

To show (3.56), we split up the functionals $\mathcal{F}_{f_j}$ into three integrals over U_ε, $\Omega\setminus\Omega_{-\delta}$, and $\Omega_{-\delta}\setminus U_\varepsilon$, and study their asymptotic behavior for $j \to \infty$ separately. First, since the local Lipschitz condition (3.50) in combination with Hölder's inequality, (3.58), and the second convergence in (3.60) shows

$$\lim_{j\to\infty} \int_{U_\varepsilon} |f_j(x,\nabla v_j + K_\chi(\tilde{u}_j)) - f_j(x,\nabla v_j)|\,dx = 0,$$

we can use (3.57) to infer

$$\begin{aligned}
\limsup_{j\to 0} \int_{U_\varepsilon} f_j(x, D^s_\delta u_j)\,dx &= \limsup_{j\to\infty} \int_{U_\varepsilon} f_j(x,\nabla v_j + K_\chi(\tilde{u}_j))\,dx \\
&= \limsup_{j\to\infty} \int_{U_\varepsilon} f_j(x,\nabla v_j)\,dx \\
&\leq \int_{\Omega_{-\delta}} f_\infty(x,\nabla v)\,dx + \varepsilon = \int_{\Omega_{-\delta}} f_\infty(x, D^s_\delta u)\,dx + \varepsilon.
\end{aligned} \tag{3.61}$$

Second, $D^s_\delta u_j = D^s_\delta u + K_\chi(\tilde{u}_j) \to D^s_\delta u$ in $L^p(\Omega \setminus \Omega_{-\delta};\mathbb{R}^{m\times n})$ along with (3.54) implies

$$\lim_{j\to\infty} \int_{\Omega\setminus\Omega_{-\delta}} f_j(x, D^s_\delta u_j)\,dx = \lim_{j\to\infty} \mathcal{J}_{f_j}(D^s_\delta u_j) = \mathcal{J}_{f^\infty}(D^s_\delta u) = \int_{\Omega\setminus\Omega_{-\delta}} f_\infty(x, D^s_\delta u)\,dx. \tag{3.62}$$

For the third integral expression, we find with the upper bound in (3.49) that

$$\begin{aligned}\limsup_{j\to\infty} \int_{\Omega_{-\delta}\setminus U_\varepsilon} f_j(x, D^s_\delta u_j)\,dx &\le \limsup_{j\to\infty} C\big(\|D^s_\delta u + K_\chi(\tilde{u}_j)\|_{L^p(\Omega_{-\delta}\setminus U_\varepsilon)} + |\Omega_{-\delta} \setminus U_\varepsilon|\big)\\ &= C\big(\|D^s_\delta u\|_{L^p(\Omega_{-\delta}\setminus U_\varepsilon)} + |\Omega_{-\delta} \setminus U_\varepsilon|\big).\end{aligned} \tag{3.63}$$

Summing (3.61), (3.62) and (3.63) finally gives

$$\limsup_{j\to\infty} \mathcal{F}_{f_j}(u_j) \le \mathcal{F}_{f_\infty}(u) + C\big(\|D^s_\delta u\|_{L^p(\Omega_{-\delta}\setminus U_\varepsilon)} + |\Omega_{-\delta} \setminus U_\varepsilon|\big) + \varepsilon.$$

Letting $U_\varepsilon \uparrow \Omega_{-\delta}$ and $\varepsilon \downarrow 0$ finishes the proof of (3.56) after choosing an appropriate diagonal sequence. □

As indicated before, the above theorem enables us to carry over well-known results on variational convergence for standard integral-functionals to our nonlocal setting. One example we wish to highlight here lies within the variational theory of homogenization. Given the classical findings in [48, 168], we can derive the Γ-limit of nonlocal functionals with periodic oscillations in the space variable as an immediate consequence of Theorem 3.5.1. It turns out that the homogenized integrand complies with the same (multi-)cell formula as in the classical case when integrating over Ω, while in the strip where complementary-values are prescribed, an averaging of the integrand in the fast variable occurs.

Corollary 3.5.2 (Homogenization). *Let $Y = (0,1)^n$ and let $f : \mathbb{R}^n\times\mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory integrand that is Y-periodic in its first argument and satisfies for a.e. $y \in \mathbb{R}^n$ and all $A, B \in \mathbb{R}^{m\times n}$ that*

$$c|A|^p - C \le f(y, A) \le C(|A|^p + 1)$$

and

$$|f(y, A) - f(y, B)| \le M(1 + |A|^{p-1} + |B|^{p-1})|A - B| \tag{3.64}$$

with constants $c, C, M > 0$. Further, let the functionals $\mathcal{F}_\varepsilon, \mathcal{F}_{\mathrm{hom}} : L^p(\Omega_\delta;\mathbb{R}^m) \to \mathbb{R}_\infty$ with $\varepsilon > 0$ be defined as

$$\mathcal{F}_\varepsilon(u) = \begin{cases} \displaystyle\int_\Omega f\Big(\frac{x}{\varepsilon}, D^s_\delta u\Big)\,dx & \text{for } u \in H^{s,p,\delta}_g(\Omega;\mathbb{R}^m),\\ \infty & \text{otherwise,}\end{cases}$$

and

$$\mathcal{F}_{\mathrm{hom}}(u) = \begin{cases} \displaystyle\int_{\Omega_{-\delta}} f_{\mathrm{hom}}(D^s_\delta u)\,dx + \int_{\Omega\setminus\Omega_{-\delta}} \bar{f}(D^s_\delta u)\,dx & \text{for } u \in H^{s,p,\delta}_g(\Omega;\mathbb{R}^m),\\ \infty & \text{otherwise,}\end{cases}$$

where $\bar{f} := \int_Y f(y,\,\cdot\,)\,dy$ and f_{hom} is the classical homogenized integrand given for $A \in \mathbb{R}^{m\times n}$ by

$$f_{\mathrm{hom}}(A) = \lim_{k\to\infty} \frac{1}{k^n} \inf\left\{\int_{kY} f(y, A + \nabla v(y))\,dy \;:\; v \in W^{1,\infty}_\#(kY;\mathbb{R}^m)\right\}. \tag{3.65}$$

Then, the convergence

$$\Gamma(L^p)\text{-}\lim_{\varepsilon\to 0} \mathcal{F}_\varepsilon = \mathcal{F}_{\mathrm{hom}}$$

holds, along with the corresponding compactness in $L^p(\Omega_\delta;\mathbb{R}^m)$.

Proof. Let $\{\varepsilon_j\}_{j\in\mathbb{N}}$ be a sequence with $\varepsilon_j \to 0$ as $j \to \infty$ and set

$$f_j(x, A) := f\Big(\frac{x}{\varepsilon_j}, A\Big) \text{ for } j \in \mathbb{N} \quad \text{and} \quad f_\infty(x, A) = \mathbb{1}_{\Omega_{-\delta}}(x) f_{\mathrm{hom}}(A) + \mathbb{1}_{\Omega\setminus\Omega_{-\delta}}(x)\bar{f}(A),$$

for $x \in \Omega$ and $A \in \mathbb{R}^{m\times n}$. To conclude the statement from Theorem 3.5.1, it suffices to verify the two convergence conditions (3.51) and (3.52) for these specific choices of f_j and f_∞. Indeed, (3.51) follows from a classical homogenization result, see e.g., [50, Theorem 2.1]. For (3.52), we note that since $\mathcal{J}_{f_j}$ is locally Lipschitz on $L^p(\Omega \setminus \Omega_{-\delta}; \mathbb{R}^{m\times n})$ with a constant uniform in j by (3.64), it is enough to prove the pointwise convergence on a dense set, for example on $C(\overline{\Omega}; \mathbb{R}^{m\times n})$. For $V \in C(\overline{\Omega}; \mathbb{R}^{m\times n})$, the convergence

$$\lim_{j\to\infty} \mathcal{J}_{f_j}(V) = \lim_{j\to\infty} \int_{\Omega\setminus\Omega_{-\delta}} f\Big(\frac{x}{\varepsilon_j}, V\Big)\,dx = \int_{\Omega\setminus\Omega_{-\delta}} \bar{f}(V)\,dx = \mathcal{J}_{f_\infty}(V)$$

follows from the fact that $(y, x) \mapsto f(y, V(x))$ is an admissible two-scale integrand (cf. [9, Corollary 5.4]). □

As a special case of the homogenization result when the integrand does not depend on y, we derive a relaxation result for functionals $\mathcal{F} : L^p(\Omega_\delta; \mathbb{R}^m) \to \mathbb{R}_\infty$ of the form

$$\mathcal{F}(u) = \begin{cases} \displaystyle\int_\Omega f(D^s_\delta u)\,dx & \text{if } u \in H^{s,p,\delta}_g(\Omega; \mathbb{R}^m), \\ \infty & \text{otherwise.} \end{cases} \tag{3.66}$$

Recall that the relaxation of $\mathcal{F}$ with respect to L^p-convergence is given by

$$\mathcal{F}^{\mathrm{rel}}(u) = \inf\Big\{\liminf_{j\to\infty} \mathcal{F}(u_j) \;:\; u_j \to u \text{ in } L^p(\Omega_\delta; \mathbb{R}^m)\Big\},$$

which corresponds to the Γ-limit of the sequence constantly equal to $\mathcal{F}$ (cf. [80, Remark 4.5]). Besides, it is easy to verify in this case that the multi-cell homogenization formula in (3.65) reduces to the quasiconvex envelope

$$f^{\mathrm{qc}}(A) = \inf\Big\{\int_Y f(A + \nabla v)\,dx \;:\; v \in W^{1,\infty}_{\#}(Y; \mathbb{R}^m)\Big\}, \quad A \in \mathbb{R}^{m\times n}.$$

The following statement is now an immediate consequence of Corollary 3.5.2.

Corollary 3.5.3 (Relaxation). *Let $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ be continuous and satisfy for all $A, B \in \mathbb{R}^{m\times n}$*

$$c|A|^p - C \leq f(A) \leq C(|A|^p + 1)$$

and

$$|f(A) - f(B)| \leq M(1 + |A|^{p-1} + |B|^{p-1})|A - B|$$

with constants $c, C, M > 0$. Then, the relaxation of $\mathcal{F}$ in (3.66) is given by

$$\mathcal{F}^{\mathrm{rel}}(u) = \begin{cases} \displaystyle\int_{\Omega_{-\delta}} f^{\mathrm{qc}}(D^s_\delta u)\,dx + \int_{\Omega\setminus\Omega_{-\delta}} f(D^s_\delta u)\,dx & \text{if } u \in H^{s,p,\delta}_g(\Omega; \mathbb{R}^m), \\ \infty & \text{otherwise.} \end{cases} \tag{3.67}$$

Remark 3.5.4. Note that the three Γ-convergence statements in this section can be rephrased equivalently for functionals defined on $H^{s,p,\delta}_g(\Omega; \mathbb{R}^m)$, if the latter is endowed with the weak topology (cf. e.g., [80, Proposition 8.16]). △

3.6 Γ-convergence for varying fractional parameter

Finally, we study the asymptotic behavior of the nonlocal integral functionals in (3.68) as the fractional parameter s varies. Of particular interest is the critical regime $s \to 1$, which leads to localization, meaning a local limit functional, as we prove below.

The set-up in this section is similar to the previous one. Let $s \in [0,1]$, $p \in (1,\infty)$ and let $\Omega \subset \mathbb{R}^n$ be open and bounded such that $\Omega_{-\delta}$ is a Lipschitz domain. Further, let $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory function with (uniform) p-growth and p-coercivity in the second variable, i.e., there are constants $c, C > 0$ such that

$$c|A|^p - C \leq f(x,A) \leq C(1+|A|^p) \qquad \text{for a.e. } x \in \Omega \text{ and all } A \in \mathbb{R}^{m\times n}.$$

We define the functionals $\mathcal{F}_s : L^p(\Omega_\delta;\mathbb{R}^m) \to \mathbb{R}_\infty$ as

$$\mathcal{F}_s(u) = \begin{cases} \displaystyle\int_\Omega f(x, D^s_\delta u(x))\,dx & \text{for } u \in H^{s,p,\delta}_0(\Omega;\mathbb{R}^m), \\ \infty & \text{otherwise.} \end{cases} \tag{3.68}$$

Recalling that D^1_δ is defined to coincide with the classical weak gradient, i.e., $D^1_\delta u = \nabla u$, and the identification of $H^{1,p,\delta}_0(\Omega;\mathbb{R}^m)$ in (3.19), we have for $s = 1$ the local integral functional,

$$\mathcal{F}_1(u) = \int_\Omega f(x, \nabla u(x))\,dx \quad \text{for } u \in W^{1,p}_0(\Omega_\delta;\mathbb{R}^m) \text{ with } u = 0 \text{ a.e. in } \Omega_\delta \setminus \Omega_{-\delta},$$

and $\mathcal{F}_1 = \infty$ otherwise in $L^p(\Omega_\delta;\mathbb{R}^m)$.

The next theorem establishes the variational convergence of the functionals $\{\mathcal{F}_s\}_s$. The proof combines the preparations and tools from the earlier sections, such as the compactness result in Lemma 3.3.9 and the translation mechanism between nonlocal and local gradients of Theorem 3.2.13.

Theorem 3.6.1 (Γ-limits for $s \to s' \in [0,1]$). *Let $\mathcal{F}_s$ for $s \in [0,1]$ be as in* (3.68) *with the additional property that $f(x,\cdot)$ is quasiconvex for a.e. $x \in \Omega_{-\delta}$. Then, the family $\{\mathcal{F}_s\}_s$ converges for $s \to s'$ to $\mathcal{F}_{s'}$ in the sense of Γ-convergence, both regarding the weak and strong topology in $L^p(\Omega_\delta;\mathbb{R}^m)$, that is,*

$$\Gamma(L^p)\text{-}\lim_{s\to s'} \mathcal{F}_s = \mathcal{F}_{s'} = \Gamma(w\text{-}L^p)\text{-}\lim_{s\to s'} \mathcal{F}_s. \tag{3.69}$$

Sequential compactness of sequences with uniformly bounded energy holds with respect to the strong topology in $L^p(\Omega_\delta;\mathbb{R}^m)$ if $s' \in (0,1]$ and the weak topology in $L^p(\Omega;\mathbb{R}^m)$ if $s' = 0$.

Proof. Let $\{s_j\}_{j\in\mathbb{N}} \subset [0,1]$ be a sequence converging to $s' \in [0,1]$ as $j \to \infty$.

Step 1: Compactness. Let $\{u_j\}_{j\in\mathbb{N}} \subset L^p(\Omega_\delta;\mathbb{R}^m)$ with $\sup_{j\in\mathbb{N}} \mathcal{F}_{s_j}(u_j) < \infty$. This implies $u_j \in H^{s_j,p,\delta}_0(\Omega;\mathbb{R}^m)$ for all $j \in \mathbb{N}$ and by the coercivity bound on f, also

$$\sup_{j\in\mathbb{N}} \|D^{s_j}_\delta u_j\|_{L^p(\Omega;\mathbb{R}^{m\times n})} < \infty.$$

We can therefore use Lemma 3.3.9 to deduce the strong and weak sequential compactness of the sequence $\{u_j\}_j$ in $L^p(\Omega_\delta;\mathbb{R}^m)$ when $s' \in (0,1]$ and $s' = 0$, respectively.

Step 2: Liminf-inequality for weakly converging sequences. Let $u \in L^p(\Omega_\delta;\mathbb{R}^m)$ and $\{u_j\}_{j\in\mathbb{N}} \subset L^p(\Omega_\delta;\mathbb{R}^m)$ with $u_j \rightharpoonup u$ in $L^p(\Omega_\delta;\mathbb{R}^m)$. Assuming without loss of generality that

$$\liminf_{j\to\infty} \mathcal{F}_{s_j}(u_j) = \lim_{j\to\infty} \mathcal{F}_{s_j}(u_j) < \infty$$

yields $u_j \in H_0^{s_j,p,\delta}(\Omega;\mathbb{R}^m)$ for $j \in \mathbb{N}$ and by the coercivity bound on f, also

$$\sup_{j\in\mathbb{N}} \|D_\delta^{s_j} u_j\|_{L^p(\Omega;\mathbb{R}^{m\times n})} < \infty.$$

Hence, Lemma 3.3.9 applies, which shows that $u \in H_0^{s',p,\delta}(\Omega;\mathbb{R}^m)$ with

$$D_\delta^{s_j} u_j \rightharpoonup D_\delta^{s'} u \ \text{ in } L^p(\Omega;\mathbb{R}^n) \quad \text{and} \quad D_\delta^{s_j} u_j \to D_\delta^{s'} u \ \text{ a.e. in } \Omega \setminus \Omega_{-\delta} \text{ as } j \to \infty. \tag{3.70}$$

Defining

$$v_j = \begin{cases} Q_\delta^{s_j} * u_j & \text{if } s_j \neq 1, \\ u_j & \text{if } s_j = 1 \end{cases} \quad \text{for } j \in \mathbb{N} \quad \text{and} \quad v = \begin{cases} Q_\delta^{s'} * u & \text{if } s' \neq 1, \\ u & \text{if } s' = 1, \end{cases}$$

we conclude from Theorem 3.2.13 (i) that $\{v_j\}_{j\in\mathbb{N}} \subset W^{1,p}(\Omega;\mathbb{R}^m)$ and $v \in W^{1,p}(\Omega;\mathbb{R}^m)$ with

$$\nabla v_j = D_\delta^{s_j} u_j \text{ on } \Omega \text{ for } j \in \mathbb{N} \quad \text{and} \quad \nabla v = D_\delta^{s'} u \text{ on } \Omega. \tag{3.71}$$

Moreover, as $\sup_{t\in[0,1)} \|Q_\delta^t\|_{L^1(\mathbb{R}^n)} < \infty$ by Lemma 3.3.1, the sequence $\{v_j\}_{j\in\mathbb{N}}$ is bounded in $W^{1,p}(\Omega;\mathbb{R}^m)$. In account of (3.70) and (3.71), one can find a non-relabeled subsequence with $v_j \rightharpoonup v$ in $W^{1,p}(\Omega;\mathbb{R}^m)$, after a suitable choice of translations. The quasiconvexity (in $\Omega_{-\delta}$) and p-growth of f then allow us to invoke a well-known weak lower semicontinuity result (cf. e.g., [75, Theorem 8.11]) to infer

$$\begin{aligned} \int_{\Omega_{-\delta}} f(x, D_\delta^{s'} u)\, dx &= \int_{\Omega_{-\delta}} f(x, \nabla v)\, dx \\ &\leq \liminf_{j\to\infty} \int_{\Omega_{-\delta}} f(x, \nabla v_j)\, dx = \liminf_{j\to\infty} \int_{\Omega_{-\delta}} f(x, D_\delta^{s_j} u_j)\, dx. \end{aligned} \tag{3.72}$$

On the other hand, in view of the pointwise convergence from (3.70) and the fact that f is Carathéodory and bounded from below by a constant, we may use Fatou's lemma to deduce

$$\liminf_{j\to\infty} \int_{\Omega\setminus\Omega_{-\delta}} f(x, D_\delta^{s_j} u_j)\, dx \geq \int_{\Omega\setminus\Omega_{-\delta}} f(x, D_\delta^{s'} u)\, dx. \tag{3.73}$$

Summing (3.72) and (3.73) shows

$$\liminf_{j\to\infty} \mathcal{F}_{s_j}(u_j) \geq \mathcal{F}_{s'}(u),$$

as desired.

Step 3: Strongly converging recovery sequences. Our construction relies on the uniform convergence of the nonlocal gradients in Lemma 3.3.2. The rest follows then via a standard density and diagonalization argument.

To be precise, let us consider $u \in H_0^{s',p,\delta}(\Omega;\mathbb{R}^m)$, otherwise the limsup-inequality is immediate due to $\mathcal{F}_{s'}(u) = \infty$. By the definition of $H_0^{s',p,\delta}(\Omega;\mathbb{R}^m)$, there is a sequence $\{u_k\}_{k\in\mathbb{N}} \subset C_c^\infty(\Omega_{-\delta};\mathbb{R}^m)$ with

$$u_k \to u \quad \text{in } H^{s',p,\delta}(\Omega;\mathbb{R}^m) \text{ as } k \to \infty.$$

For each $k \in \mathbb{N}$, Lemma 3.3.2 shows $D_\delta^{s_j} u_k \to D_\delta^{s'} u_k$ in $L^p(\Omega;\mathbb{R}^{m\times n})$ as $j \to \infty$, and we conclude from Lebesgue's dominated convergence theorem combined with the growth bound on f that

$$\lim_{j\to\infty} \mathcal{F}_{s_j}(u_k) = \lim_{j\to\infty} \int_\Omega f(x, D_\delta^{s_j} u_k)\, dx = \int_\Omega f(x, D_\delta^{s'} u_k)\, dx = \mathcal{F}_{s'}(u_k).$$

Since $D_\delta^{s'} u_k \to D_\delta^{s'} u$ in $L^p(\Omega;\mathbb{R}^{m\times n})$ as $k\to\infty$, an analogous reasoning gives $\lim_{k\to\infty}\mathcal{F}_{s'}(u_k) = \mathcal{F}_{s'}(u)$. Altogether, we have that $u_k \to u$ in $L^p(\Omega_\delta;\mathbb{R}^m)$ and

$$\lim_{k\to\infty}\lim_{j\to\infty}\mathcal{F}_{s_j}(u_k) = \mathcal{F}_{s'}(u).$$

Extracting a suitable diagonal sequence $\{u_{k_j}\}_{j\in\mathbb{N}}$ via Attouch's lemma finishes the proof. □

Remark 3.6.2. a) We remark that the two Γ-convergence statements in (3.69) are equivalent to the L^p-Mosco-convergence of the family $\{\mathcal{F}_s\}_s$ to $\mathcal{F}_{s'}$.

b) Note that one cannot expect strong L^p-compactness for $\{\mathcal{F}_s\}_s$ as $s\to 0$, considering that $H^{0,p,\delta}(\Omega;\mathbb{R}^m) = L^p(\Omega_\delta;\mathbb{R}^m)$ with equivalent norms (cf. Remark 3.2.6).

c) Throughout this paper, we work with the sequential definition of Γ-limits, which may differ in general from the topological definition for non-metric spaces. However, the equi-coerciveness of the family $\{\mathcal{F}_s\}_s$ in $L^p(\Omega_\delta;\mathbb{R}^m)$ (in fact, $\mathcal{F}_s(u) \geq c'\|u\|_{L^p(\Omega_\delta;\mathbb{R}^m)} - C$ for all $u\in L^p(\Omega_\delta;\mathbb{R}^m)$ due to Theorem 3.3.8) and the metrizability of the weak L^p-topology on norm bounded sets guarantee that the sequential $\Gamma(w\text{-}L^p)$-limit coincides with the topological one, see e.g., [80, Proposition 8.10].

d) It is not hard to see that an analogous statement to Theorem 3.6.1 holds also for more general complementary values other than zero, e.g., for $g\in H^{1,p,\delta}(\Omega;\mathbb{R}^m)$.

e) Under the additional assumptions required in the relaxation result of Corollary 3.5.3, we can prove Γ-convergence for $\{\mathcal{F}_s\}_s$ as $s\to s'\in(0,1]$ also in the case when f is a homogeneous integrand that is not necessarily quasiconvex. Indeed, by first relaxing the functionals (cf. [80, Proposition 6.11]), we find

$$\Gamma(L^p)\text{-}\lim_{s\to s'}\mathcal{F}_s = \Gamma(L^p)\text{-}\lim_{s\to s'}\mathcal{F}_s^{\mathrm{rel}} = \mathcal{F}_{s'}^{\mathrm{rel}};$$

here, $\mathcal{F}_s^{\mathrm{rel}}$ is given by the relaxation formula (3.67) for $s\in(0,1)$, which extends also to the case $s=1$ because of classical relaxation theory. △

3.7 Conclusion

One of the two ingredients in the direct method of calculus of variations for proving the existence of minimizers is weak lower semicontinuity of the functional. In the case of integral functionals, it is closely linked to notions of convexity of the integrand. This paper shows that quasiconvexity, the classical convexity notion in vectorial variational calculus, also characterizes weak lower semicontinuity of functionals depending on nonlocal gradients D_δ^s (Theorem 3.4.1). As a consequence, we could establish a general existence theory for this class of energy functionals (Corollary 3.4.4).

The technical foundation for our new findings lies in suitable translation operators that allow switching between nonlocal and classical gradients, as well as fractional ones, as stated in Theorem 3.2.13. This result has interest on its own and can potentially be useful in various problems involving nonlocal gradients. However, this does not imply that the nonlocal framework automatically inherits every result from the local case; in particular, this translation technique does not preserve boundary conditions.

As an application, we derive a general statement about variational convergence, which allows us to carry over established Γ-convergence results from the local to the nonlocal setting. We illustrate the flexibility of this method with the example of homogenization and relaxation (Corollaries 3.5.2 and 3.5.3). In both cases, the integrand of the limit functional shows the same form in the classical case, i.e., a multi-cell formula for the homogenized integrand and a quasiconvex envelope for the relaxed one, except for the collar region.

Finally, we considered the asymptotics for varying fractional order, proving the continuity of the energy functional with respect to s, Theorem 3.6.1. The two main ingredients for the proof are a Poincaré inequality independent of s (Theorem 3.3.8) and an extension of the nonlocal version of the fundamental theorem of calculus to $s = 0$ (Proposition 3.2.9). As a particular case, we obtained a localization result when s goes to 1, recovering a classical limit problem. An alternative localization of interest for future work is to study the limit of vanishing horizon $\delta \to 0$ under the assumption of a normalized kernel.

Appendix 3.A Comparison with the Riesz potential kernel

To provide the technical basis for quantitative comparisons between the convolution kernel that can be used to represent the nonlocal gradient and the Riesz potential kernel, which plays the analogous role for the Riesz fractional gradient, we collect here several useful properties about the quantity

$$R_\delta^s = Q_\delta^s - I_{1-s}$$

with $s \in [0, 1)$.

Recalling the definitions of Q_δ^s and I_{1-s} in (3.16) and (3.10), (3.11), respectively, we can represent R_δ^s as

$$R_\delta^s(x) = \begin{cases} c_{n,s} \displaystyle\int_{|x|}^{\infty} \frac{\overline{w}_\delta(t) - 1}{t^{n+s}}\, dt & \text{if } n + s - 1 > 0, \\ c_{1,0} \displaystyle\int_{|x|}^{\infty} \frac{\overline{w}_\delta(t)}{t}\, dt + \frac{1}{\pi} \log(|x|) & \text{if } n = 1 \text{ and } s = 0 \end{cases}$$

for $x \in \mathbb{R}^n \setminus \{0\}$; note that $c_{n,s} = \frac{n+s-1}{\gamma_{n,1-s}}$ and $c_{1,0} = \frac{1}{\pi}$. As a consequence,

$$\nabla R_\delta^s(x) = c_{n,s}(1 - w_\delta(x)) \frac{x}{|x|^{n+s+1}} \quad \text{for } x \in \mathbb{R}^n, \tag{3.74}$$

for all $n \geq 1$ and $s \in [0, 1)$. Observe that $\nabla R_\delta^s \in L^1(\mathbb{R}^n; \mathbb{R}^n) \cap C^\infty(\mathbb{R}^n; \mathbb{R}^n)$ for $s \in (0, 1)$ and

$$D_\delta^s \varphi - D^s \varphi = \nabla((Q_\delta^s - I_{1-s}) * \varphi) = \nabla R_\delta^s * \varphi \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^n). \tag{3.75}$$

Since $1 - w_\delta$ is zero near the origin by (H3), R_δ^s is constant near the origin.

The Fourier transform of R_δ^s for any $s \in [0, 1)$ satisfies

$$\widehat{R_\delta^s}(\xi) = \widehat{Q_\delta^s}(\xi) - \frac{1}{|2\pi\xi|^{1-s}} \qquad \text{for } |\xi| \geq 1; \tag{3.76}$$

if $n + s - 1 > 0$, this follows directly from the well-known formula for $\widehat{I_{1-s}}$ (see e.g., [122, Theorem 2.4.6]), and also the case $n = 1$ and $s = 0$ is standard; one can argue via the fact that the distributional derivative of $-\frac{1}{\pi} \log(|\cdot|)$ corresponds to the distribution

$$\eta \mapsto \lim_{r \downarrow 0} \int_{(-r,r)^c} \frac{\eta(x)}{x}\, dx \quad \text{for } \eta \in \mathcal{S}(\mathbb{R}),$$

whose Fourier transform equals $i\mathrm{sgn}$ (see e.g., [122, Eq. (5.1.12)]); note that in this case $\widehat{R_\delta^s}$ is only a tempered distribution on $\mathbb{R}^n$, but for convenience we view it as a function outside $B(0, 1)$.

The following auxiliary result establishes estimates on the decay behavior of the Fourier transform of R_δ^s and its derivatives.

Lemma 3.A.1. *Let $s \in [0,1)$ and let $\beta, \omega \in \mathbb{N}_0^n$ be multi-indices. Then, there exists a constant $C > 0$ independent of s such that*

$$|\xi|^{|\beta|}\big|\partial^\omega \widehat{R}^s_\delta(\xi)\big| \leq C \quad \textit{for all } |\xi| \geq 1. \tag{3.77}$$

Proof. Throughout the proof, we use C to denote possibly different constants that do not depend on s; note in particular, that we may ignore the constant $c_{n,s}$, since it is bounded for $s \in [0,1)$. Additionally, we may restrict to the case $|\beta| \geq |\omega| + 2$ since $|\xi| \geq 1$.

We observe first that by the boundedness of the Fourier transform from $L^1(\mathbb{R}^n;\mathbb{C})$ to $C_0(\mathbb{R}^n;\mathbb{C})$ in combination with standard properties of the interaction between the Fourier transform with derivatives (see e.g., [122, Proposition 2.3.22 (8)-(9)]), the claim follows as soon as

$$\big\|\partial^\beta\big((-2\pi i\,\cdot)^\omega R^s_\delta\big)\big\|_{L^1(\mathbb{R}^n;\mathbb{C})} \leq C < \infty, \tag{3.78}$$

is established. The argument, which is detailed below, relies on repeated use of the Leibniz rule and exploits the representation (3.74).

Let $\gamma, \gamma', \gamma'', \tau \in \mathbb{N}_0^n$ in the following be multi-indices not exceeding the order of β. A straightforward calculation shows that

$$\left|\partial^{\gamma'}\left(\frac{x}{|x|^{n+s+1}}\right)\right| \leq \frac{C}{|x|^{n+s+|\gamma'|}}$$

for $x \in \mathbb{R}^n \setminus \{0\}$, and we have due to (H2) and (H3) that $\partial^{\gamma''} w_\delta = 0$ outside of the annulus $A_\delta := B(0,\delta) \setminus B(0, b_0\delta)$ if $\gamma'' \neq 0$. Hence,

$$\begin{aligned}\left|\partial^{\gamma''} w_\delta(x)\partial^{\gamma'}\left(\frac{x}{|x|^{n+s+1}}\right)\right| &\leq \frac{C}{(b_0\delta)^{n+s+|\gamma'|}}\mathbb{1}_{A_\delta}(x)\\ &\leq C\big((b_0\delta)^{-n-1-|\beta|}+1\big)\mathbb{1}_{A_\delta}(x) \leq C\mathbb{1}_{A_\delta}(x).\end{aligned}$$

This allows us to infer in view of (3.74), the Leibniz rule, and again (H3), that

$$\big|\partial^\gamma R^s_\delta(x)\big| \leq C\Big(|1-w_\delta(x)|\frac{1}{|x|^{n+s+|\gamma|-1}} + \mathbb{1}_{A_\delta}(x)\Big) \leq C\left(\frac{\mathbb{1}_{B(0,b_0\delta)^c}(x)}{|x|^{n+s+|\gamma|-1}} + \mathbb{1}_{A_\delta}(x)\right) \tag{3.79}$$

for $\gamma \neq 0$. Moreover, if $|\tau| \leq |\omega|$, we have

$$|\partial^\tau(-2\pi i x)^\omega| \leq C|x|^{|\omega|-|\tau|}, \tag{3.80}$$

and $\partial^\tau(-2\pi i x)^\omega = 0$ for $|\tau| > |\omega|$.

Another application of Leibniz' rule together with (3.79) and (3.80) finally yields for $x \in \mathbb{R}^n \setminus \{0\}$ that

$$\big|\partial^\beta\big((-2\pi i x)^\omega R^s_\delta(x)\big)\big| \leq C\left(\frac{\mathbb{1}_{B(0,b_0\delta)^c}(x)}{|x|^{n+s+|\beta|-|\omega|-1}} + \mathbb{1}_{A_\delta}(x)\right).$$

It follows now via integration and under consideration of $s + |\beta| - |\omega| - 1 \geq 1$ that

$$\big\|\partial^\beta\big((-2\pi i\,\cdot)^\omega R^s_\delta\big)\big\|_{L^1(\mathbb{R}^n;\mathbb{C})} \leq C\left(\frac{(b_0\delta)^{-s-|\beta|+|\omega|+1}}{s+|\beta|-|\omega|-1} + \delta^n\right) \leq C\big((b_0\delta)^{|\omega|-|\beta|} + 1 + \delta^n\big),$$

which gives (3.78). □

Appendix 3.B Proof of density results

This part of the appendix is devoted to proving the density result stated in Theorem 3.2.8. We begin with a lemma on the Leibniz rule for the distributionally defined spaces $H^{s,p,\delta}(\Omega)$. It serves as a technical tool for proving the approximate extension and retraction results stated afterwards.

Lemma 3.B.1. *Let $s \in [0,1)$, $p \in [1,\infty)$ and $\Omega \subset \mathbb{R}^n$ be open and bounded. Further, let $u \in H^{s,p,\delta}(\Omega)$, identified with its extension by zero, and $\chi \in C_c^\infty(\mathbb{R}^n)$. If $\Omega' \subset \mathbb{R}^n$ is an open and bounded set such that*

$$(\Omega' \setminus \Omega) \cap \operatorname{supp}(\chi) = \emptyset, \tag{3.81}$$

then $\chi u \in H^{s,p,\delta}(\Omega')$.

Proof. Clearly, $\chi u \in L^p(\Omega'_\delta)$. To determine the weak nonlocal gradient, we calculate for any $\varphi \in C_c^\infty(\Omega';\mathbb{R}^n)$ that

$$\begin{aligned}\int_{\Omega'_\delta} (\chi u)\operatorname{div}_\delta^s \varphi\,dx &= \int_{\Omega'_\delta} u\big(\operatorname{div}_\delta^s(\chi\varphi) - K_\chi(\varphi^\intercal)\big)\,dx\\ &= -\int_{\Omega'} D_\delta^s u \cdot (\chi\varphi)\,dx - \int_{\Omega'_\delta} uK_\chi(\varphi^\intercal)\,dx\\ &= -\int_{\Omega'} \chi D_\delta^s u\cdot\varphi + K_\chi(u)\cdot\varphi\,dx.\end{aligned}$$

Indeed, the first line exploits the Leibniz rule for the nonlocal divergence in (3.20), while the second line follows directly from the formula defining the weak nonlocal gradient, which is valid here since $\chi\varphi \in C_c^\infty(\Omega;\mathbb{R}^n)$ in light of the assumption (3.81). For the third equality, we have used Fubini's theorem and the boundedness of $K_\chi : L^p(\Omega'_\delta) \to L^p(\Omega';\mathbb{R}^n)$ according to Lemma 3.2.11.

The calculation above shows that $D_\delta^s(\chi u) = \chi D_\delta^s u + K_\chi(u)$ on Ω', and hence, $u \in H^{s,p,\delta}(\Omega')$. □

The next auxiliary results will be useful in the proofs of Theorem 3.2.8 and Proposition 3.2.10 to generate room for mollification arguments. The techniques are similar to the proof of [189, Theorem 3.9].

Lemma 3.B.2 (Approximate extension and retraction). *Let $s \in [0,1)$, $p \in [1,\infty)$, and let $\Omega \subset \mathbb{R}^n$ be an open and bounded set.*

(*i*) *If Ω is Lipschitz, then for any $\varepsilon > 0$ and $u \in H^{s,p,\delta}(\Omega)$ there exists $\Omega' \Supset \Omega$ and $u_\varepsilon \in H^{s,p,\delta}(\Omega')$ such that*

$$\|u - u_\varepsilon\|_{H^{s,p,\delta}(\Omega)} < \varepsilon.$$

(*ii*) *If $\Omega_{-\delta}$ is Lipschitz, then for any $\varepsilon > 0$ and $u \in H_0^{s,p,\delta}(\Omega)$ there exists $u_\varepsilon \in H_0^{s,p,\delta}(\Omega)$ with $\operatorname{supp}(u_\varepsilon) \Subset \Omega_{-\delta}$ and*

$$\|u - u_\varepsilon\|_{H^{s,p,\delta}(\Omega)} < \varepsilon.$$

Proof. (*i*) Given that the boundary of Ω is locally the graph of a Lipschitz function, we can find a partition of unity $\chi_0, \chi_1, \dots, \chi_{N+1} \subset C_c^\infty(\mathbb{R}^n)$ and translation vectors $\zeta_1, \dots, \zeta_N \in \mathbb{R}^n$ such that

$$\sum_{i=0}^{N+1} \chi_i = 1 \;\text{ on } \Omega_\delta, \quad \chi_0 \in C_c^\infty(\Omega), \quad \chi_{N+1} \in C_c^\infty(\Omega^c)$$

and

$$(\operatorname{supp}(\chi_i) \cap \Omega^c) + \lambda\zeta_i \Subset \Omega^c \quad \text{for } i = 1, \ldots, N. \tag{3.82}$$

for all $\lambda > 0$ small enough. For these λ, we define

$$v_\lambda := \chi_0 u + \chi_{N+1} u + \sum_{i=1}^{N} \tau_{\lambda\zeta_i}(\chi_i u)$$

where $\tau_\zeta(v) := v(\cdot - \zeta)$ denotes translation by $\zeta \in \mathbb{R}^n$ of a function $v : \mathbb{R}^n \to \mathbb{R}$. Note that by construction, $v_\lambda \in H^{s,p,\delta}(\Omega)$ according to Lemma 3.B.1.

Next, we exploit continuity of the translation operator on L^p and the translation invariance of the nonlocal gradient to find $\lambda_\varepsilon > 0$ such that $u_\varepsilon := v_{\lambda_\varepsilon}$ satisfies

$$\|u - u_\varepsilon\|^p_{H^{s,p,\delta}(\Omega)} \leq \sum_{i=1}^{N} \|\chi_i u - \tau_{\lambda_\varepsilon\zeta_i}(\chi_i u)\|^p_{L^p(\Omega_\delta)} + \sum_{i=1}^{N} \|D^s_\delta(\chi_i u) - \tau_{\lambda_\varepsilon\zeta_i} D^s_\delta(\chi_i u)\|^p_{L^p(\Omega;\mathbb{R}^n)} < \varepsilon^p.$$

Finally, if $\Omega' \Supset \Omega$ is chosen such that

$$(\operatorname{supp}(\chi_i) \cap \Omega^c) + \lambda_\varepsilon\zeta_i \Subset (\Omega')^c \quad \text{for } i = 1, \ldots, N \quad \text{and} \quad \operatorname{supp}(\chi_{N+1}) \Subset (\Omega')^c,$$

where the first condition is achievable in view of (3.82), Lemma 3.B.1 implies that even $u_\varepsilon \in H^{s,p,\delta}(\Omega')$, as desired.

(ii) A similar argument to that in (i) applies here as well, with the main difference in the choice of the partition of unity, which is now considered for $\Omega_{-\delta}$ and translated inwards instead of outwards as in (3.82). □

With these tools at hand, one can now deduce the alternative characterizations for $H^{s,p,\delta}(\Omega)$ and $H^{s,p,\delta}_g(\Omega)$ from Sections 3.2.2 and 3.2.3, respectively.

Proof of Theorem 3.2.8. Case 1: $\Omega = \mathbb{R}^n$. Via a mollification argument we may suppose that $u \in C^\infty(\mathbb{R}^n) \cap H^{s,p,\delta}(\mathbb{R}^n)$. Take $\chi \in C^\infty_c(\mathbb{R}^n)$ with $\chi \equiv 1$ on $B(0,1)$ and define $\chi_j := \chi(\cdot/j)$ for $j \in \mathbb{N}$. We then find that $\{\chi_j u\}_{j\in\mathbb{N}} \subset C^\infty_c(\mathbb{R}^n)$, $\chi_j u \to u$ in $L^p(\mathbb{R}^n)$ and

$$\|D^s_\delta u - D^s_\delta(\chi_j u)\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \leq \|(1-\chi_j) D^s_\delta u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} + C\operatorname{Lip}(\chi_j)\|u\|_{L^p(\mathbb{R}^n)} \to 0 \text{ as } j \to \infty,$$

where we have used Lemma 3.2.11 and the fact that $\operatorname{Lip}(\chi_j) \leq \operatorname{Lip}(\chi)/j$.

Case 2: Ω *a bounded Lipshitz domain.* Lemma 3.B.2 (i) implies for every $j \in \mathbb{N}$ that there is $u_j \in H^{s,p,\delta}(\Omega'_j)$ with some appropriately chosen $\Omega \Subset \Omega'_j$ such that

$$\|u - u_j\|_{H^{s,p,\delta}(\Omega)} < \frac{1}{2j}. \tag{3.83}$$

We are now in the position to use a standard mollification procedure on u_j, identified with its extension to $\mathbb{R}^n$ by zero, with mollifying radius smaller than $d(\partial\Omega, \partial\Omega'_j)$ to find a $\varphi_j \in C^\infty_c(\mathbb{R}^n)$ with

$$\|u_j - \varphi_j\|_{H^{s,p,\delta}(\Omega)} < \frac{1}{2j}, \tag{3.84}$$

so that the result follows from (3.83) and (3.84) along with the triangle inequality. □

Proof of Proposition 3.2.10. Without loss of generality, consider $g = 0$. Utilizing a similar strategy as above, one can apply Lemma 3.B.2 (ii) and suitably mollify the resulting function $u_\varepsilon \in H^{s,p,\delta}_0(\Omega)$ with support compactly contained in $\Omega_{-\delta}$. □

Chapter 4

Non-constant functions with zero nonlocal gradient and their role in nonlocal Neumann-type problems

This chapter is available as the preprint

[141] C. Kreisbeck and H. Schönberger. Non-constant functions with zero nonlocal gradients and their role in nonlocal Neumann-type problems. Preprint arXiv:2402.11308, 2024.

4.1 Introduction

It is well-known that differentiable functions with zero gradient are exactly the constant functions, that is, for any open and connected set $\Omega \subset \mathbb{R}^n$ and $u \in C^1(\Omega)$ it holds that

$$\nabla u = 0 \text{ in } \Omega \qquad \text{if and only if} \qquad u \text{ is constant on } \Omega, \tag{4.1}$$

and the same is true (almost everywhere) for Sobolev functions with weak gradients. One may wonder if this fundamental observation carries over when considering fractional and nonlocal derivatives instead of classical derivatives. As intriguingly basic as the question may sound, a universal answer is not easily available and depends on the specific setting, as we will demonstrate in the following.

In fractional and nonlocal calculus, the study of gradient operators has attained increasing attention in recent years, see e.g., [28, 30, 66, 92, 140, 161, 193, 208]. The nonlocal gradient of a function $u : \mathbb{R}^n \to \mathbb{R}$ is of the form

$$\mathcal{G}_\rho u(x) = \int_{\mathbb{R}^n} \frac{u(x) - u(y)}{|x - y|} \frac{x - y}{|x - y|} \rho(x - y)\, dy \tag{4.2}$$

with a suitable kernel function ρ, whenever the integral is defined.

Especially the Riesz fractional gradient, that is, $D^s = \mathcal{G}_{\rho^s}$ with $\rho^s \propto |\cdot|^{-(n+s-1)}$ for $s \in (0,1)$, has been popular. Not only does D^s have unique natural invariance and homogeneity properties [208], it also lends itself to a distributional approach towards fractional function spaces [66]; in fact, the function spaces associated with D^s in analogy to the classical Sobolev spaces coincide with the Bessel potential spaces $H^{s,p}(\mathbb{R}^n)$, as observed in [193]. The combination of these features make D^s a good choice of fractional derivative, both from the applied point of view and in the context of variational theories and PDEs.

In contrast, a compactly supported, radial kernel ρ in (4.2) reduces the nonlocal interaction between all of $\mathbb{R}^n$ to points within an interaction range $\delta > 0$, commonly referred to as horizon. By cutting-off the Riesz potential kernel, one obtains the finite-horizon fractional gradient defined as

$$D_\delta^s = \mathcal{G}_{\rho_\delta^s} \quad \text{with} \quad \rho_\delta^s \propto \frac{w_\delta}{|\cdot|^{n+s-1}}$$

where $w_\delta : \mathbb{R}^n \to [0,1]$ is a radial cut-off function supported in a ball of radius δ; for further properties, we refer to Section 4.2.2. These gradients D_δ^s, which we simply call nonlocal gradients in the following, are the key objects in this paper. They were first considered in [30] by Bellido, Cueto & Mora Corral (see also [72]), motivated by applications in materials science. Since the nonlocal gradients inherit the desirable properties from the Riesz fractional gradients, while being suitable for variational problems on bounded domains, they have become the core of a newly proposed model for nonlocal elasticity.

On a more technical note, we remark that in order to properly determine $D_\delta^s u$ in Ω, the function u needs to be defined in the set $\Omega_\delta = \{x \in \mathbb{R}^n : \operatorname{dist}(x, \Omega) < \delta\}$ enlarged by the horizon variable, and in particular, in the collar $\Gamma_\delta := \Omega_\delta \setminus \overline{\Omega}$ of thickness $\delta > 0$ around Ω, cf. Figure 4.1. The values of u in Γ_δ can be viewed as nonlocal boundary values. One defines the space $H^{s,p,\delta}(\Omega)$ as the functions in $L^p(\Omega_\delta)$ with $D_\delta^s u \in L^p(\Omega)$; see Section 4.2.2 for more details. As a powerful tool, we wish to point out the translation mechanism established in [30,72] (cf. Section 4.2.3). It relates the nonlocal and local setting in the sense that nonlocal gradients can be expressed as classical ones and vice versa, allowing for statements to be carried over; in formulas, we have

$$D_\delta^s = \nabla \circ (Q_\delta^s * \cdot) \qquad \text{and} \qquad \nabla = \mathcal{P}_\delta^s \circ D_\delta^s, \tag{4.3}$$

where Q_δ^s is an integrable, compactly supported kernel function and $\mathcal{P}_\delta^s$ corresponds to the inverse of the convolution with Q_δ^s. In comparison with the analogous results for the Riesz fractional gradient [140, 193], the operator $\mathcal{P}_\delta^s$ replaces the fractional Laplacian of order $\frac{1-s}{2}$.

Let us now return to and specify the question raised earlier:

Is (4.1) still true when ∇ is replaced with $\mathcal{G}_\rho$?

In the case of the Riesz fractional gradient $\mathcal{G}_\rho = D^s$ on $H^{s,p}(\mathbb{R}^n)$ for $p \in (1,\infty)$, the answer is affirmative as a consequence of the fractional Poincaré-type inequalites [193, Theorem 1.8, 1.10 and 1.11]; in fact, the functions with vanishing Riesz fractional gradient must even be zero due to their integrability properties. The same is true when $\mathcal{G}_\rho = D_\delta^s$ is considered for functions in the complementary-value space $H_0^{s,p,\delta}(\Omega) := \{u \in H^{s,p,\delta}(\mathbb{R}^n) : u = 0 \text{ a.e. in } \Omega^c\}$. Here as well, a Poincaré inequality is available, see [30, Theorem 6.2]. If the complementary-value is dropped, however, and one considers nonlocal gradients $\mathcal{G}_\rho = D_\delta^s$ on $H^{s,p,\delta}(\Omega)$ with bounded Ω, the picture changes substantially.

This paper revolves around the class of functions with zero nonlocal gradient

$$N^{s,p,\delta}(\Omega) := \{h \in H^{s,p,\delta}(\Omega) : D_\delta^s h = 0 \text{ a.e. in } \Omega\},$$

which turns out to be non-trivial. Indeed, we show that there exist functions in $N^{s,p,\delta}(\Omega)$ that are non-constant in any open subset of Ω (Proposition 4.3.1) and establish that they are numerous in the sense that $N^{s,p,\delta}(\Omega)$ forms an infinite-dimensional space (Proposition 4.3.3).

Knowing that the set $N^{s,p,\delta}(\Omega)$ consists of more than just constant functions stirs up interesting new issues for further investigation. We first give a characterization of all the elements of $N^{s,p,\delta}(\Omega)$, which provides a deeper understanding of its properties. With this at hand, we then highlight the role of the functions with zero nonlocal gradients in the theory of the spaces $H^{s,p,\delta}(\Omega)$ and discuss applications in nonlocal differential inclusion problems and variational problems on $H^{s,p,\delta}(\Omega)$ with Neumann-type boundary conditions. Here is a more detailed overview of our main new findings.

Characterization of $N^{s,p,\delta}(\Omega)$. While the set of functions in $W^{1,p}(\Omega)$ with zero gradient corresponds to the set of constant functions and can thus, be identified with $\mathbb{R}$ by taking mean values, the characterization of $N^{s,p,\delta}(\Omega)$ involves an additional feature due to boundary effects of the nonlocal interactions. Roughly speaking, two ingredients are necessary to uniquely identify the elements of $N^{s,p,\delta}(\Omega)$, that is, an average or mean-value condition on Ω and boundary values in the collar region $\Gamma_\delta := \Omega_\delta \setminus \overline{\Omega}$.

We start from the observation that $N^{s,p,\delta}(\Omega)$ consists of all functions $h \in L^p(\Omega_\delta)$ satisfying

$$Q_\delta^s * h = c \text{ a.e. in } \Omega \qquad \text{and} \qquad h = g \text{ a.e. in } \Gamma_\delta, \tag{4.4}$$

for a given $g \in L^p(\Gamma_\delta)$ and $c \in \mathbb{R}$. This is a consequence of the translation mechanism (4.3). Hence, the problem reduces to finding the solutions of (4.4). Since $\mathcal{P}_\delta^s$ from (4.3) is in fact a pseudo-differential operator, our proof strategy is to rewrite (4.4) equivalently as a pseudo-differential Dirichlet problem and to exploit the recent progress in their existence, uniqueness, and regularity theory. Precisely, the properties of $\mathcal{P}_\delta^s$ make it fit into the setting of the works by Grubb [125] and by Abels & Grubb [2]. Given that the regularity results are sensitive to the relation between the fractional and integrability parameters s and p, there are two qualitatively different regimes to distinguish.

Our main characterization result (see Theorem 4.3.8 and Proposition 4.3.12) states the following:

(i) If $p \in (1, \frac{2}{1-s})$ (including the case $p = 2$), then $N^{s,p,\delta}(\Omega)$ consists of the unique solutions to (4.4), which exist for every constant $c \in \mathbb{R}$ and given boundary values $g \in L^p(\Gamma_\delta)$.

(ii) For $p \in [\frac{2}{1-s}, \infty)$, only those (unique) solutions to (4.4) that lie also in $L^p(\Omega_\delta)$ constitute $N^{s,p,\delta}(\Omega)$.

An alternative way of phrasing (i) is to say that

$$N^{s,p,\delta}(\Omega) \text{ is isomorphic to } \mathbb{R} \times L^p(\Gamma_\delta),$$

with the isomorphism $N^{s,p,\delta}(\Omega) \ni h \mapsto (\int_\Omega Q_\delta^s * h \, dx, h|_{\Gamma_\delta})$. This formalizes the statement that an average condition on Ω and the boundary values in a boundary layer of thickness δ are the characteristics for any function with zero nonlocal gradient. Besides, we show that $h \mapsto (\int_\Omega h \, dx, h|_{\Gamma_\delta})$ is a isomorphism between $N^{s,p,\delta}(\Omega)$ and $\mathbb{R} \times L^p(\Gamma_\delta)$ as well, which indicates that even a simple mean-value condition along with the values in Γ_δ suffices to pin down the elements of $N^{s,p,\delta}(\Omega)$. Part (ii) implies that the previous identifications with $\mathbb{R} \times L^p(\Gamma_\delta)$ remain injective when $p \in [\frac{2}{1-s}, \infty)$, however, surjectivity generally fails (Remark 4.3.13).

Technical tools in $H^{s,p,\delta}(\Omega)$ modulo functions of zero nonlocal gradient. The set $N^{s,p,\delta}(\Omega)$ can be used to develop new functional analytic tools for the spaces $H^{s,p,\delta}(\Omega)$ without complementary-values. Unlike for $H^{s,p,\delta}(\mathbb{R}^n)$ and $H_0^{s,p,\delta}(\Omega)$, however, analogues of the relevant tools and estimates in classical Sobolev spaces only hold in the quotient space $H^{s,p,\delta}(\Omega)/N^{s,p,\delta}(\Omega)$, meaning modulo elements in $N^{s,p,\delta}(\Omega)$. With that in mind, we obtain the following:

(a) *Refined translation mechanism for functions on bounded domains:* We show in Theorem 4.4.1 that the quotient space $H^{s,p,\delta}(\Omega)/N^{s,p,\delta}(\Omega)$ is isometrically isomorphic to $W^{1,p}(\Omega)$ modulo constants, meaning that one can identify $H^{s,p,\delta}(\Omega)$ and $W^{1,p}(\Omega)$ up to functions with zero (nonlocal) gradient. The isomorphism turns nonlocal gradients into gradients.

(b) *Extension of functions from $H^{s,p,\delta}(\Omega)$ to $H^{s,p,\delta}(\mathbb{R}^n)$ up to $N^{s,p,\delta}(\Omega)$:* Even though an exact extension of functions in $H^{s,p,\delta}(\Omega)$ to $\mathbb{R}^n$ is generally not possible (cf. Example 4.3.4), there exists a bounded linear operator $\mathcal{E}^s_\delta : H^{s,p,\delta}(\Omega) \to H^{s,p,\delta}(\mathbb{R}^n)$, such that $\mathcal{E}^s_\delta u$ differs from u in Ω_δ merely by a function with zero nonlocal gradient.

(c) *Nonlocal Poincaré-type inequalities:* As a major tool, we prove different nonlocal versions of Poincaré inequalities. If $p \in (1, \frac{2}{1-s})$, there exists a constant $C > 0$ such that

$$\|u\|_{L^p(\Omega_\delta)} \leq C\|D^s_\delta u\|_{L^p(\Omega;\mathbb{R}^n)}$$

for all $u \in H^{s,p,\delta}(\Omega)$ satisfying $u = 0$ in Γ_δ and one of the averaging conditions $\int_\Omega u\,dx = 0$ or $\int_\Omega Q^s_\delta * u\,dx = 0$. The same estimate holds for all $p \in (1,\infty)$ and $u \in H^{s,p,\delta}(\Omega)$ whose metric projection onto $N^{s,p,\delta}(\Omega)$ vanishes.

(d) *L^p-compactness modulo $N^{s,p,\delta}(\Omega)$.* Based on (c), we derive the following Rellich-Kondrachov-type compactness: If $(u_j)_j \subset H^{s,p,\delta}(\Omega)$ is a bounded sequence such that the metric projection of u_j onto $N^{s,p,\delta}(\Omega)$ vanishes for all j, then $(u_j)_j$ is precompact in $L^p(\Omega_\delta)$.

We remark that for the complementary-value spaces $H^{s,p,\delta}_0(\Omega)$ the analogues of (a), (c), and (d) have recently been established in [30, 72]. The approach there relies on Fourier techniques, given that the functions in $H^{s,p,\delta}_0(\Omega)$ are defined on the whole of $\mathbb{R}^n$.

Variational problems on $N^{s,p,\delta}(\Omega)^\perp$ and nonlocal boundary-value problems. A significant application of the aforementioned tools are the existence theory and asymptotic analysis of nonlocal PDEs subject to Neumann-type boundary conditions. Precisely, we adopt a variational viewpoint and prove the existence of solutions to the problem

$$\text{Minimize} \quad \frac{1}{2}\int_\Omega |D^s_\delta u|^2\,dx - \int_{\Omega_\delta} Fu\,dx \quad \text{over } N^{s,2,\delta}(\Omega)^\perp \subset H^{s,2,\delta}(\Omega), \tag{4.5}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain, $F \in L^2(\Omega_\delta)$, and $N^{s,2,\delta}(\Omega)^\perp$ denotes the orthogonal complement of $N^{s,2,\delta}(\Omega)$; note that $N^{s,2,\delta}(\Omega)$ plays the same role as the set of constant functions in the variational formulation of the Neumann problem with classical gradients. In fact, a remarkable aspect of our framework is that one can also handle more general vector-valued nonlinear problems with $p \in (1,\infty)$ and energy densities that are either quasiconvex or polyconvex, see Theorem 4.6.1 and Remark 4.6.2.

To draw the connection between (4.5) and PDEs with Neumann-type boundary conditions, one assumes the nonlocal compatibility condition

$$\int_{\Omega_\delta} Fh\,dx = 0 \quad \text{for all } h \in N^{s,2,\delta}(\Omega),$$

under which the solutions to (4.5) weakly satisfy Euler-Lagrange equations with a nonlocal boundary operator $\mathcal{N}^s_\delta$ featured in the collar regions, see (4.65). In fact, this boundary operator was recently introduced by Bellido, Cueto, Foss & Radu [26], where the authors derive, amongst others, a new integration by parts formula in the spirit of [100].

Our second main result regarding (4.5) confirms the expectation that these problems localize as the fractional parameter s tends to 1, that is, they converge to their classical counterparts with usual gradients. Working in the framework of variational convergence, we obtain that the Γ-limit of the functional in (4.5) with respect to strong convergence in $L^2(\Omega_\delta)$ is

$$\frac{1}{2}\int_\Omega |\nabla u|^2\,dx - \int_\Omega Fu\,dx \qquad \text{for } u \in W^{1,2}(\Omega) \text{ with } \int_\Omega u\,dx = 0,$$

see Theorem 4.6.4; again, this also holds in the more general setting mentioned before. In the case that F satisfies the classical compatibility condition $\int_\Omega F\,dx = 0$, we obtain, in particular, that the minimizers of (4.5) converge in $L^2(\Omega)$ as $s \uparrow 1$ to the unique mean-zero solution of the standard Neumann problem

$$\begin{cases} -\Delta u = F & \text{in } \Omega, \\ \frac{\partial u}{\partial \nu} = 0 & \text{on } \partial\Omega. \end{cases}$$

Localisation via rigorous limit passages is a general theme in the study of fractional and nonlocal calculus, not least because they serve as important consistency checks for new problems and models; we refer e.g., to [29,72] for $s \uparrow 1$, and [35,158,161] for limits with vanishing horizon $\delta \to 0$.

Let us close by pointing out some literature on Neumann problems in other fractional and nonlocal set-ups, involving the fractional Laplacian and more general integral and integro-differential operators, see e.g., [22, 45, 89, 100, 125, 167] and also the references therein. One of the works we wish to highlight is [100], where Dipierro, Ros-Oton & Valdinoci introduce a Neumann problem for the fractional Laplacian by a natural notion of normal nonlocal derivative. These results have been refined, expanded and generalized in various directions, e.g., in [22,98,99,111]. Closely related are also the recent results on nonlocal trace spaces [105,126,127,203]. The distinguishing factor in our work, is the central role of a nonlocal gradient object, which enables us to handle a broad variety of nonlinearities.

Outline. We have organized this paper as follows. After introducing notations and providing theoretical background and useful auxiliary results in Section 4.2, Section 4.3 is centered around a solid understanding of the functions with vanishing finite-horizon fractional gradient. Our analysis includes proofs that non-constant functions with vanishing nonlocal gradient exist and that $N^{s,p,\delta}(\Omega)$ is an infinite-dimensional space, see Section 4.3.1. The main theorems about the characterization of $N^{s,p,\delta}(\Omega)$ are stated and proven in Section 4.3.2. We round off this section with a discussion of regularity properties of functions with zero nonlocal gradient and give illustrative examples in Section 4.3.3.

The second part of the paper presents different implications and applications involving $N^{s,p,\delta}(\Omega)$. In Section 4.4, we establish the technical tools (a)-(d) for working in the nonlocal function spaces $H^{s,p,\delta}(\Omega)$. The previous findings are then used in Section 4.5 to contribute to the theory of nonlocal differential inclusions. We show that rigidity statements as well as existence results for approximate solutions can be carried over from the classical setting via the translation mechanism. Section 4.6 features the new class of variational problems on $N^{s,p,\delta}(\Omega)^\perp$, which relates to nonlocal Neumann-type problems. A proof of well-posedness for these problems is contained in Section 4.6.1, while Section 4.6.2 establishes the rigorous link with the classical local problems through a localization result via Γ-convergence.

4.2 Preliminaries

In this section, we introduce the relevant notations and collect the necessary background on nonlocal gradients and function spaces along with some useful technical tools.

4.2.1 Notation

Unless specified otherwise in the following, we take $s \in (0,1)$, $p \in [1,\infty]$, and $\delta > 0$. The Euclidean norm of $x \in \mathbb{R}^n$ is denoted by $|x|$ and

$$\langle x \rangle := \sqrt{1 + |x|^2}.$$

We use the notation l_A with $A \in \mathbb{R}^{m\times n}$ for the linear function $l_A(x) = Ax$ with $x \in \mathbb{R}^n$.

Moreover, $E^c := \mathbb{R}^n \setminus E$ is the complement of a set $E \subset \mathbb{R}^n$, $\overline{E}$ is its closure, and $|E|$ is its Lebesgue measure, provided E is measurable. We use the notation $\mathbb{1}_E$ for the indicator function of a set $E \subset \mathbb{R}^n$, i.e., $\mathbb{1}_E(x) = 1$ if $x \in E$ and $\mathbb{1}_E(x) = 0$ otherwise. Whenever convenient, we identify a function on a subset of $\mathbb{R}^n$ with its trivial extension by zero without explicit mention. If we wish to highlight the trivial extension, we use an extra indicator function, writing e.g., $\mathbb{1}_E f : \mathbb{R}^n \to \mathbb{R}$ for the zero extension of $f : E \to \mathbb{R}$. The restriction of any $f : E \to \mathbb{R}$ to a subset $E' \subset E$ is denoted by $f|_{E'}$.

By $B_\rho(x) = \{y \in \mathbb{R}^n : |x - y| < \rho\}$, we denote the ball centered at $x \in \mathbb{R}^n$ with radius $\rho > 0$, and $\operatorname{dist}(x, E)$ is the distance between a point $x \in \mathbb{R}^n$ and a set $E \subset \mathbb{R}^n$. For a domain $\Omega \subset \mathbb{R}^n$, i.e., open and connected set, we introduce its expansion and reduction by thickness δ as

$$\Omega_\delta := \Omega + B_\delta(0) = \{x \in \mathbb{R}^n : \operatorname{dist}(x, \Omega) < \delta\} \quad \text{and} \quad \Omega_{-\delta} := \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \delta\},$$

where $\partial\Omega$ is the boundary of Ω, and define

$$\Gamma_\delta := \Omega_\delta \setminus \overline{\Omega} \quad \text{and} \quad \Gamma_{-\delta} := \Omega \setminus \overline{\Omega_{-\delta}}$$

as the inner and outer collars of Ω, respectively. Further, let $\Gamma_{\pm\delta} := \Gamma_\delta \cup \Gamma_{-\delta} \cup \partial\Omega$ be the double layer around the boundary of Ω. For an illustration of this geometric set-up, see Figure 4.1.

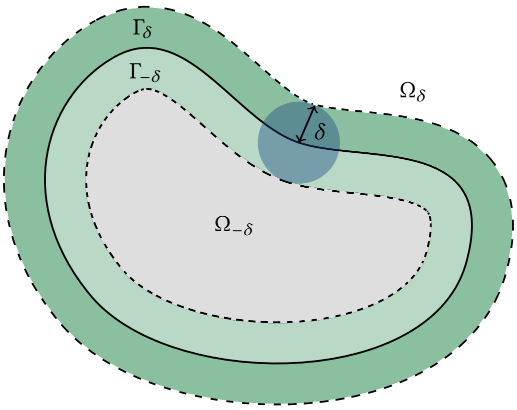

Figure 4.1: Illustration of a set $\Omega \subset \mathbb{R}^n$ with its expansion Ω_δ, the outer and inner collar regions Γ_δ (green) and $\Gamma_{-\delta}$ (light green), and the reduced set $\Omega_{-\delta}$ (gray).

Let $U \subset \mathbb{R}^n$ be an open set. The notation $C_c^\infty(U)$ stands for the space of smooth functions $U \mapsto \mathbb{R}$ with compact support, which will often be identified with their trivial extensions to $\mathbb{R}^n$ by zero, and $\operatorname{Lip}(\psi)$ is the Lipschitz constant of a function $\psi : \mathbb{R}^n \to \mathbb{R}$. Throughout the paper, we use the standard notation for Lebesgue- and Sobolev-spaces $L^p(U)$ and $W^{1,p}(U)$ with $p \in [1, \infty]$. For the inner product on $L^2(U)$, we write $\langle \cdot, \cdot \rangle_{L^2(U)}$. Notice that each of the function spaces defined above, as well as those introduced later, can be extended componentwise to vector-valued functions; the target set is then reflected in the notation, for example, $L^p(U; \mathbb{R}^m)$. Moreover, the restriction of a function space is denoted, for example, as $C^\infty(\mathbb{R}^n)|_U := \{u|_U : u \in C^\infty(\mathbb{R}^n)\}$.

For an integrable function $f \in L^1(\mathbb{R}^n)$, the Fourier transform is defined as

$$\widehat{f}(\xi) := \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i x \cdot \xi}\, dx, \quad \xi \in \mathbb{R}^n.$$

It is well-known that the Fourier transform is an isomorphism from the Schwartz space $\mathcal{S}(\mathbb{R}^n; \mathbb{C})$ onto itself, which can be extended to the spaces $L^2(\mathbb{R}^n; \mathbb{C})$ and the space of tempered distributions $\mathcal{S}'(\mathbb{R}^n; \mathbb{C})$ by density and duality, respectively. The inverse Fourier transform of f, denoted $f^\vee$, corresponds to $x \mapsto \widehat{f}(-x)$. For more background on Fourier analysis, see e.g., [106, 122].

Lastly, C denotes a generic constant, which may change from one estimate to the next without further mention. To indicate that a constant depends on specific quantities, they are added in brackets.

4.2.2 Nonlocal gradients and Sobolev spaces

Let us now introduce in detail the key objects in this paper, namely, a class of fractional gradients with finite horizon, and the associated nonlocal Sobolev spaces. Our presentation follows along the lines of [30, 72] (see also [32]), where we also refer to for more details.

The truncated Riesz fractional gradient, simply referred to as nonlocal gradient, and the corresponding divergence for smooth functions are defined as follows: For $s \in (0,1)$ and $\delta > 0$, the nonlocal gradient of $\varphi \in C^\infty(\mathbb{R}^n)$ is

$$D^s_\delta \varphi(x) = \int_{\mathbb{R}^n} \frac{\varphi(x) - \varphi(y)}{|x-y|} \frac{x-y}{|x-y|} \rho^s_\delta(x-y)\, dy \quad \text{for } x \in \mathbb{R}^n, \tag{4.6}$$

and the nonlocal divergence of $\psi \in C^\infty(\mathbb{R}^n; \mathbb{R}^n)$ is

$$\operatorname{div}^s_\delta \psi(x) = \int_{\mathbb{R}^n} \frac{\psi(x) - \psi(y)}{|x-y|} \cdot \frac{x-y}{|x-y|} \rho^s_\delta(x-y)\, dy \quad \text{for } x \in \mathbb{R}^n;$$

here, the kernel function ρ^s_δ is given by

$$\rho^s_\delta(z) = c_{n,s,\delta} \frac{w_\delta(z)}{|z|^{n+s-1}} \qquad \text{for } z \in \mathbb{R}^n \setminus \{0\},$$

with $w_\delta : \mathbb{R}^n \to [0,\infty)$ a non-negative cut-off function satisfying the hypotheses

(H1) w_δ is radial, i.e., $w_\delta = \overline{w}_\delta(|\cdot|)$ with a function $\overline{w}_\delta : \mathbb{R} \to [0,\infty)$;

(H2) w_δ is smooth and compactly supported in $B_\delta(0)$, i.e., $w_\delta \in C^\infty_c(B_\delta(0))$;

(H3) w_δ is normalized around the origin, i.e., $w_\delta = 1$ on $B_{\mu\delta}(0)$ for some $\mu \in (0,1)$;

(H4) w_δ is radially decreasing, i.e., $w_\delta(z) \geq w_\delta(\tilde{z})$ if $|z| \leq |\tilde{z}|$,

and the scaling constant $c_{n,s,\delta} > 0$ is such that

$$c_{n,s,\delta} \int_{B_\delta(0)} \frac{w_\delta(z)}{|z|^{n+s-1}}\, dz = n. \tag{4.7}$$

Remark 4.2.1. a) Note that the scaling factor $c_{n,s,\delta}$, determined by (4.7), is the same as in [32] in order to ensure that the nonlocal derivative of any linear map l_A with $A \in \mathbb{R}^{m\times n}$ is equal to A, see [32, Proposition 4.1]. This choice is slightly different from the scaling in [72], but provides no substantial issues for the application of these results, as discussed in Remark 4.2.3 below.

b) An alternative way of expressing the nonlocal gradient in (4.6) is as a principle value integral. In view of the radial symmetry of w_δ from (H1), one has for $x \in \mathbb{R}^n$ that

$$D^s_\delta \varphi(x) = \text{p.v.} \int_{B(x,r)^c} \varphi(y) d^s_\delta(x-y)\, dy := \lim_{r \downarrow 0} \int_{B(x,r)^c} \varphi(y) d^s_\delta(x-y)\, dy \tag{4.8}$$

with $d^s_\delta(z) = -c_{n,s,\delta} \frac{z w_\delta(z)}{|z|^{n+s+1}}$ for $z \in \mathbb{R}^n \setminus \{0\}$. This shows, in particular, that $D^s_\delta \varphi(x)$ can be written as the convolution of d^s_δ with φ, when $x \notin \operatorname{supp}(\varphi)$. △

The above definitions can be extended to locally integrable functions via a distributional approach. Indeed, for $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\psi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$, the integration by parts formula

$$\int_{\mathbb{R}^n} D_\delta^s \varphi \cdot \psi \, dx = -\int_{\mathbb{R}^n} \varphi \operatorname{div}_\delta^s \psi \, dx$$

holds. Based on this, we then define $v \in L^1_{\mathrm{loc}}(\Omega;\mathbb{R}^n)$ as the weak nonlocal gradient of $u \in L^1_{\mathrm{loc}}(\Omega_\delta)$, written as $v = D_\delta^s u$, if

$$\int_\Omega v \cdot \psi \, dx = -\int_{\Omega_\delta} u \operatorname{div}_\delta^s \psi \, dx \qquad \text{for all } \psi \in C_c^\infty(\Omega;\mathbb{R}^n); \tag{4.9}$$

the weak nonlocal divergence is defined analogously. In parallel to classical Sobolev spaces, one can introduce nonlocal Sobolev spaces as follows.

Definition 4.2.2 (Nonlocal Sobolev spaces). *Let $s \in (0,1)$, $\delta > 0$, $p \in [1,\infty]$, and let $\Omega \subset \mathbb{R}^n$ be open. The nonlocal Sobolev space $H^{s,p,\delta}(\Omega)$ is defined as*

$$H^{s,p,\delta}(\Omega) := \{u \in L^p(\Omega_\delta) \,:\, D_\delta^s u \in L^p(\Omega;\mathbb{R}^n)\},$$

endowed with the norm

$$\|u\|_{H^{s,p,\delta}(\Omega)} = \left(\|u\|^p_{L^p(\Omega_\delta)} + \|D_\delta^s u\|^p_{L^p(\Omega;\mathbb{R}^n)}\right)^{\frac{1}{p}}.$$

These spaces can be equivalently defined via density if Ω is a bounded Lipschitz domain or $\Omega = \mathbb{R}^n$, see [72, Theorem 1]. A more detailed study of these spaces, including results such as Leibniz rules, Poincaré inequalities and compact embeddings can be found in [30, 72].

When working with functions on the full space $\mathbb{R}^n$, we will often exploit the connection between the nonlocal Sobolev spaces of Definition 4.2.2 and the well-known Bessel potential spaces, which are defined for any $t \in \mathbb{R}$ and $p \in (1,\infty)$ as

$$H^{t,p}(\mathbb{R}^n) = \left\{u \in \mathcal{S}'(\mathbb{R}^n) \,:\, \left(\langle\,\cdot\,\rangle^t \widehat{u}\right)^\vee \in L^p(\mathbb{R}^n)\right\}, \tag{4.10}$$

with the norm $\|u\|_{H^{t,p}(\mathbb{R}^n)} = \|\left(\langle\,\cdot\,\rangle^t \widehat{u}\right)^\vee\|_{L^p(\mathbb{R}^n)}$ and the notation $H^t := H^{t,2}$; for more on the theory of Bessel potential spaces, see e.g., [123, Chapter 1.3.1] or [204]. Indeed, it holds for all $p \in (1,\infty)$ and $s \in (0,1)$ that

$$H^{s,p,\delta}(\mathbb{R}^n) = H^{s,p}(\mathbb{R}^n),$$

with equivalent norms. This follows from the observation that $H^{s,p}(\mathbb{R}^n)$ coincides with the space of functions in $L^p(\mathbb{R}^n)$ with a weak Riesz fractional gradient in $L^p(\mathbb{R}^n;\mathbb{R}^n)$ (cf. [193, Theorem 1.7] together with the density result in [54, Theorem A.1]), along with the fact that the latter is again the same as $H^{s,p,\delta}(\mathbb{R}^n)$ due to [72, Lemma 5].

We mention here some additional properties of the Bessel potential spaces that we need. First of all, for each $t > 0$, there is a $f_t \in L^1(\mathbb{R}^n)$ with $\|f_t\|_{L^1(\mathbb{R}^n)} = 1$ and $\widehat{f_t} = \langle\cdot\rangle^{-t}$ (f_t is a rescaled version of the Bessel potential function, see [123, Chapter 1.2.2]). Therefore, for any $t_1 < t_2$ we find with Young's convolution inequality

$$\begin{aligned}\|u\|_{H^{t_1,p}(\mathbb{R}^n)} &= \|\left(\langle\,\cdot\,\rangle^{t_1}\widehat{u}\right)^\vee\|_{L^p(\mathbb{R}^n)} = \|f_{t_2-t_1} * \left(\langle\,\cdot\,\rangle^{t_2}\widehat{u}\right)^\vee\|_{L^p(\mathbb{R}^n)} \\ &\leq \|\left(\langle\,\cdot\,\rangle^{t_2}\widehat{u}\right)^\vee\|_{L^p(\mathbb{R}^n)} = \|u\|_{H^{t_2,p}(\mathbb{R}^n)}.\end{aligned} \tag{4.11}$$

Secondly, if $(u_j)_j \subset H^{t,p}(\mathbb{R}^n)$ is a bounded sequence and $t > 0$, then we find that

$$u_j = f_t * \big(\langle \cdot \rangle^t \widehat{u_j}\big)^\vee,$$

which is the convolution of an L^1-function with a bounded sequence in L^p, and hence, precompact in $L^p_{\mathrm{loc}}(\mathbb{R}^n)$ by the Fréchet-Kolmogorov criterion (see e.g., [52, Corollary 4.28]). As such, $H^{t,p}(\mathbb{R}^n)$ is compactly embedded into $L^p_{\mathrm{loc}}(\mathbb{R}^n)$. In fact, when $tp > n$, we find that $f_t \in L^{p'}(\mathbb{R}^n)$ with p' the dual exponent of p (cf. [123, Theorem 1.3.5 (c)]), so that one can even deduce the compact embedding of $H^{t,p}(\mathbb{R}^n)$ into $C_{\mathrm{loc}}(\mathbb{R}^n)$ due to the Arzelà-Ascoli theorem.

In the following, we also use the complementary value space of $H^{t,p}(\mathbb{R}^n)$ for $t \geq 0$, which consists of functions with zero values outside of an open set $V \subset \mathbb{R}^n$ and is denoted by

$$H^{t,p}_0(V) = \{u \in H^{t,p}(\mathbb{R}^n) \,:\, u = 0 \text{ a.e. in } V^c\}.$$

4.2.3 Translation operators

In this section, we present a method that will be frequently used, and further refined (see Section 4.4.1) in this paper, namely, a translation procedure that allows switching between nonlocal gradients and classical gradients. The following auxiliary results are mainly taken from [72]; related statements about the Riesz fractional gradient have been established in [140].

Our starting point is the following finite-horizon analogue of the Riesz potential from [30], defined by

$$Q^s_\delta : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}, \quad Q^s_\delta(x) = c_{n,s,\delta} \int_{|x|}^{\delta} \frac{\overline{w}_\delta(r)}{r^{n+s}}\, dr. \tag{4.12}$$

It holds that Q^s_δ is integrable with compact support in $B_\delta(0)$ and, a simple calculation yields that, due to the choice of scaling,

$$\|Q^s_\delta\|_{L^1(\mathbb{R}^n)} = 1. \tag{4.13}$$

Remark 4.2.3. a) With the scaling constant $c_{n,s}$ used in [72], one obtains instead of (4.13) that $[0,1) \ni s \mapsto \|Q^s_\delta\|_{L^1(\mathbb{R}^n)}$ is continuous with $\lim_{s\to 1}\|Q^s_\delta\|_{L^1(\mathbb{R}^n)} = 1$, see [72, Lemma 6]. This shows that the two different scaling regimes are comparable uniformly in s.

b) The Fourier transform of $\widehat{Q}^s_\delta$ is a smooth, positive and radial function. Moreover, the difference between $\widehat{Q}^s_\delta$ and $\xi \mapsto |2\pi\xi|^{-(1-s)}$ is a Schwartz function for $|\xi| \geq 1$, see [30] and [72, Remark 2 and Lemma 11]. △

An essential observation about the kernel function Q^s_δ regards its relation with the nonlocal gradient D^s_δ, that is,

$$D^s_\delta \varphi = \nabla(Q^s_\delta * \varphi) = Q^s_\delta * \nabla\varphi \qquad \text{for any } \varphi \in C^\infty_c(\mathbb{R}^n).$$

This identity can be extended to the Sobolev spaces in a weak sense, as shown in [72, Theorem 2 (i)].

Lemma 4.2.4 (From nonlocal to local gradients). *Let $s \in (0,1)$, $\delta > 0$, $p \in [1,\infty]$, and $\Omega \subset \mathbb{R}^n$ be open. Then, the linear map $Q^s_\delta : H^{s,p,\delta}(\Omega) \to W^{1,p}(\Omega)$, $u \mapsto Q^s_\delta * u$ is bounded (uniformly with respect to s) with*

$$(\nabla \circ Q^s_\delta)u = \nabla(Q^s_\delta u) = D^s_\delta u \quad \textit{for every } u \in H^{s,p,\delta}(\Omega).$$

The convolution with the kernel Q_δ^s enables us to pass from the nonlocal Sobolev space to the classical one. To go back, we consider the operator from [72], given by

$$\mathcal{P}_\delta^s : \mathcal{S}(\mathbb{R}^n) \mapsto \mathcal{S}(\mathbb{R}^n), \quad \widehat{\mathcal{P}_\delta^s \varphi} = \frac{\widehat{\varphi}}{\widehat{Q_\delta^s}}. \tag{4.14}$$

It is proven in [72, Theorem 2 (ii)] that this operator can be extended to the Sobolev space as the inverse of convolution with Q_δ^s.

Lemma 4.2.5 (From local to nonlocal gradients). *Let $s \in (0,1)$, $\delta > 0$, $p \in [1,\infty]$. Then, $\mathcal{P}_\delta^s$ in* (4.14) *can be extended to a isomorphism between $W^{1,p}(\mathbb{R}^n)$ and $H^{s,p,\delta}(\mathbb{R}^n)$ with satisfies $(\mathcal{P}_\delta^s)^{-1} = \mathcal{Q}_\delta^s$. In particular,*

$$(D_\delta^s \circ \mathcal{P}_\delta^s)v = D_\delta^s(\mathcal{P}_\delta^s v) = \nabla v \quad \textit{for every } v \in W^{1,p}(\mathbb{R}^n). \tag{4.15}$$

We mention a useful alternative representation of $\mathcal{P}_\delta^s$ involving the kernel function of the nonlocal fundamental theorem of calculus in [30, Theorem 4.5]. It holds that

$$\mathcal{P}_\delta^s \varphi(x) = \int_{\mathbb{R}^n} V_\delta^s(x-y) \cdot \nabla\varphi(y)\, dy =: (V_\delta^s * \nabla\varphi)(x) \quad \text{for } \varphi \in C_c^\infty(\mathbb{R}^n), \tag{4.16}$$

where $V_\delta^s \in C^\infty(\mathbb{R}^n \setminus \{0\}; \mathbb{R}^n) \cap L^1_{\mathrm{loc}}(\mathbb{R}^n; \mathbb{R}^n)$ is a vector-radial function, i.e., $V_\delta^s(x) = x f_\delta^s(|x|)$ for $x \in \mathbb{R}^n \setminus \{0\}$ with $f_\delta^s : (0,\infty) \to \mathbb{R}$, cf. [72, Remark 4 d)] as well as [30, Theorem 5.9] for more properties of V_δ^s.

When $p \in (1,\infty)$, one can deduce some more general properties for $\mathcal{P}_\delta^s$ using Fourier methods. To this end, we recall that by Remark 4.2.3 b), there are $R_\delta^s, S_\delta^s \in \mathcal{S}(\mathbb{R}^n)$ such that

$$\widehat{Q}_\delta^s(\xi) = \frac{1}{|2\pi\xi|^{1-s}} + R_\delta^s(\xi) \quad \text{and} \quad \frac{1}{\widehat{Q}_\delta^s(\xi)} = |2\pi\xi|^{1-s} + S_\delta^s(\xi) \qquad \text{for } |\xi| \geq 1. \tag{4.17}$$

As a consequence of the Mihlin-Hörmander theorem (e.g., [122, Theorem 6.2.7]), using the smoothness and positivity of $\widehat{Q}_\delta^s$ locally and the decay of R_δ^s, S_δ^s for large ξ to obtain the desired estimates similarly to [72, Lemma 8], it follows that both

$$\langle\cdot\rangle^{1-s}\widehat{Q}_\delta^s \quad \text{and} \quad \frac{1}{\langle\cdot\rangle^{1-s}\widehat{Q}_\delta^s} \tag{4.18}$$

are L^p-multipliers. We finally infer from this observation (along with the definition of Bessel potential spaces in (4.10)) that for $t \geq 1-s$ and $p \in (1,\infty)$,

$$\mathcal{P}_\delta^s : H^{t,p}(\mathbb{R}^n) \to H^{t-(1-s),p}(\mathbb{R}^n), \tag{4.19}$$

is a isomorphism with inverse $(\mathcal{P}_\delta^s)^{-1} = \mathcal{Q}_\delta^s$.

Moreover, since the decay of R_δ^s is uniform in s, a similar argument as in [72, Lemma 8] shows that the operator norm of $\mathcal{P}_\delta^s$ is uniformly bounded in $s \in (0,1)$; in particular, using (4.11) and $H^{0,p} = L^p$, there is a $C > 0$ independent of s such that

$$\|\mathcal{P}_\delta^s v\|_{L^p(\mathbb{R}^n)} \leq \|\mathcal{P}_\delta^s v\|_{H^{s,p}(\mathbb{R}^n)} \leq C\|v\|_{W^{1,p}(\mathbb{R}^n)} \quad \text{for all } v \in W^{1,p}(\mathbb{R}^n) \text{ and } s \in (0,1). \tag{4.20}$$

4.2.4 Pseudo-differential operators and Dirichlet problems

The recent existence and uniqueness theory from [125] together with the regularity results in [2] for boundary-value problems involving pseudo-differential operators play an important role for our work. We collect here the statements that we will need, while keeping the presentation accessible, and refer to e.g., [125, Section 2.2] for precise definitions and properties of pseudo-differential operators.

For a suitable pseudo-differential operator $\mathcal{P}$, an open subset $V \subset \mathbb{R}^n$, and a function $g : V \to \mathbb{R}$, the associated Dirichlet problem reads

$$\begin{cases} \mathcal{P}w = g & \text{on } V \\ w = 0 & \text{in } V^c. \end{cases} \tag{4.21}$$

By combining the results of [2, 125], we obtain into the following statement tailored to our needs.

Theorem 4.2.6 (Existence and uniqueness for pseudo-differential Dirichlet problems). *Let $p \in (1, \infty)$, $a \in (0, 1/2)$, $V \subset \mathbb{R}^n$ be an open and bounded set with $C^{1,1}$-boundary, and let $\mathcal{P}$ be a strongly elliptic, even, classical pseudo-differential operator of order $2a$ satisfying*

$$\langle \mathcal{P}\varphi, \varphi \rangle_{L^2(\mathbb{R}^n)} \geq C\|\varphi\|^2_{H^a(\mathbb{R}^n)} \quad \textit{for all } \varphi \in C_c^\infty(\mathbb{R}^n) \tag{4.22}$$

with some constant $C > 0$. Then, there exists for every $g \in L^p(V)$ a unique $w_g \in H_0^{a,p}(V)$ with

$$\mathcal{P}w_g = g \quad \textit{in } V.$$

If $p \in (1, \frac{1}{a})$, it even holds that $w_g \in H_0^{2a,p}(V)$ and there is a $c > 0$ such that

$$\|w_g\|_{H^{2a,p}(\mathbb{R}^n)} \leq c\|g\|_{L^p(V)} \quad \textit{for all } g \in L^p(V).$$

Proof. We define the operator $\mathcal{P}_V$ with domain

$$\mathrm{dom}(\mathcal{P}_V) := \{w \in H_0^{a,p}(V) \,:\, (\mathcal{P}w)|_V \in L^p(V)\}$$

via restriction of $\mathcal{P}$ to V, i.e., $\mathcal{P}_V w = (\mathcal{P}w)|_V$ for $w \in \mathrm{dom}(\mathcal{P}_V)$. Due to (4.22), we may apply [125, Theorem 4.2] with $\beta = 0$, and then also [125, Theorem 4.16 2°], to deduce that $\mathcal{P}_V : \mathrm{dom}(\mathcal{P}_V) \to L^p(V)$ is a bijection. This shows the first part of the statement.

For the case $p \in (1, \frac{1}{a})$, we note that in [2] (see also [125, Theorem 3.2]), the domain $\mathrm{dom}(\mathcal{P}_V)$ has been characterized as the so-called a-transmission space, which agrees with $H_0^{2a,p}(V)$ when $p \in (1, \frac{1}{a})$ (cf. [125, Eq. (2.20)]). Consequently, $\mathcal{P}_V : H_0^{2a,p}(V) \to L^p(V)$ is a bijective bounded linear operator. In particular, it is invertible with bounded inverse, which implies the full statement. □

Remark 4.2.7. a) We note that connectedness is not part of the definition of a domain in [2, 125], in contrast to our definition, and hence, Theorem 4.2.6 is valid for non-connected sets V as well. Moreover, the regularity of the domain V in Theorem 4.2.6 can even be reduced to $V \subset \mathbb{R}^n$ that have $C^{1,\tau}$-boundaries with $\tau \in (2a, 1)$, see [2, 125]. Only for simplicity of the presentation, we work here with a stronger assumption.

b) Note that the range of p such that w_g lies in $H_0^{2a,p}(V)$ is sharp, which is due to the fact that the solutions to the pseudo-differential problem in (4.21) contain a factor of $\mathrm{dist}(\cdot, \partial V)^a$ near the boundary. To give a precise example, we can take any smooth, bounded domain V and any function $w : \mathbb{R}^n \to \mathbb{R}$ that is smooth in V, equal to $\mathrm{dist}(\cdot, \partial V)^a$ near the boundary ∂V and zero in V^c. It follows then by [124, Theorem 4] that $(\mathcal{P}w)|_V \in C^\infty(\mathbb{R}^n)|_V$, which implies, in particular, that $w \in H_0^{a,p}(V)$ is a solution to (4.21) with a smooth right-hand side. However, we have that $w\,\mathrm{dist}(\cdot, \partial V)^{-2a}$ is equal to $\mathrm{dist}(\cdot, \partial V)^{-a}$ near the boundary of V, which is not in $L^p(V)$ for $p \geq \frac{1}{a}$, so that $w \notin H_0^{2a,p}(V)$ in view of the Hardy-type inequality [205, Proposition 5.7]. △

The translation operator $\mathcal{P}^s_\delta$ from the previous section is in fact a pseudo-differential operator that fits exactly into the abstract framework of Theorem 4.2.6, which is the content of the following lemma. This observation will be crucial for our characterization result of $N^{s,p,\delta}(\Omega)$ (cf. Theorem 4.3.8 and Lemma 4.3.6).

Lemma 4.2.8 ($\mathcal{P}^s_\delta$ as pseudo-differential operator). *The operator $\mathcal{P}^s_\delta$ defined in (4.14) is a strongly elliptic, even classical pseudo-differential operator of order $1-s$ and there is a $C>0$ such that*

$$\langle \mathcal{P}^s_\delta \varphi, \varphi\rangle_{L^2(\mathbb{R}^n)} \geq C\|\varphi\|^2_{H^{\frac{1-s}{2}}(\mathbb{R}^n)} \quad \textit{for all } \varphi \in C_c^\infty(\mathbb{R}^n).$$

Proof. The properties can be deduced from the fact that the symbol of $\mathcal{P}^s_\delta$ is smooth, radial and positive, and for large frequencies only differs from the symbol of the fractional Laplacian $(-\Delta)^{\frac{1-s}{2}}$ up to a Schwartz function (see (4.17)). For the reader's convenience, we work out the details below, referring to [125, Section 2.2] for the precise definitions of the properties of pseudo-differential operators.

It is easy to check in light of (4.17) that for any $\alpha \in \mathbb{N}_0^n$,

$$\left|\partial^\alpha \big(1/\widehat{Q}^s_\delta\big)\right| \leq C_\alpha \langle\cdot\rangle^{1-s-|\alpha|},$$

which means that $\mathcal{P}^s_\delta$ has order $1-s$. Defining $p_0 : \mathbb{R}^n \to \mathbb{R}$ to be a smooth function with $p_0(\xi) = |2\pi\xi|^{1-s}$ for $|\xi|\geq 1$, we obtain again from (4.17) that

$$\left|\partial^\alpha \big(1/\widehat{Q}^s_\delta - p_0\big)\right| \leq C_\alpha \langle\cdot\rangle^{1-s-|\alpha|-J},$$

for any $J\in\mathbb{N}_0$. This means that $\mathcal{P}^s_\delta$ is classical (where in the expansion $p_j=0$ for $j\geq 1$) and, since $p_0(-\xi)=p_0(\xi)$ for $|\xi|\geq 1$, $\mathcal{P}^s_\delta$ is even. Finally, because $p_0(\xi)\geq C|\xi|^{1-s}$ for $|\xi|\geq 1$, the operator $\mathcal{P}^s_\delta$ is strongly elliptic, and since $1/\widehat{Q}^s_\delta \geq C\langle\cdot\rangle^{1-s}$, we have by the Plancherel identity that

$$\langle \mathcal{P}^s_\delta\varphi,\varphi\rangle_{L^2(\mathbb{R}^n)} = \langle \widehat{\varphi}/\widehat{Q}^s_\delta, \widehat{\varphi}\rangle_{L^2(\mathbb{R}^n)} \geq C\|\langle\cdot\rangle^{\frac{1-s}{2}}\widehat{\varphi}\|_{L^2(\mathbb{R}^n)} = C\|\varphi\|_{H^{\frac{1-s}{2}}(\mathbb{R}^n)}$$

for all $\varphi\in C_c^\infty(\mathbb{R}^n)$, which finishes the proof. □

4.3 Discussion and characterization of functions with zero nonlocal gradient

This section revolves around the study of the functions in the nonlocal Sobolev space with vanishing finite-horizon fractional gradient. Our analysis examines different facets of the set

$$N^{s,p,\delta}(\Omega) = \{h\in H^{s,p,\delta}(\Omega)\ :\ D^s_\delta h = 0 \text{ a.e. in } \Omega\},$$

including the existence and construction of non-trivial functions, characterization results, a discussion of regularity properties, and illustrative examples. Throughout Sections 4.3-4.6, Ω is assumed to be a bounded Lipschitz domain, unless mentioned otherwise.

4.3.1 Non-constant elements of $N^{s,p,\delta}(\Omega)$

In contrast to the functions with zero classical gradient, the set $N^{s,p,\delta}(\Omega)$ encompasses strictly more than constant functions. In fact, one can find functions in $N^{s,p,\delta}(\Omega)$ that are non-constant on any subset of Ω, which is the content of the following proposition.

Proposition 4.3.1 (Existence of non-constant functions in $N^{s,p,\delta}(\Omega)$). *Let $p \in [1, \infty]$. It holds for any open, non-empty $U \subset \Omega$ that*

$$N^{s,p,\delta}(\Omega) \not\subset \{u \in H^{s,p,\delta}(\Omega) \,:\, u \text{ is constant a.e. on } U\}.$$

Proof. Suppose to the contrary that every $u \in N^{s,p,\delta}(\Omega)$ is constant a.e. on U. The reasoning that will lead to the desired contradiction is organized in three steps.

Step 1: Representation of V_δ^s away from the origin. We deduce from the above assumption that the kernel V_δ^s in (4.16) then has to satisfy

$$V_\delta^s(x) = \frac{1}{\sigma_{n-1}} \frac{x}{|x|^n} \quad \text{for all } x \in \overline{B_\rho(0)}^c, \tag{4.23}$$

where $\rho = \operatorname{Diam}(\Omega)$ is the diameter of Ω and σ_{n-1} denotes the surface area of the unit sphere in $\mathbb{R}^n$.[1] To see this, we split the argument in three sub-steps, showing first that $\operatorname{div} V_\delta^s$ is constant outside of $\overline{B_\rho(0)}$. Next, we exploit the radiality of V_δ^s (cf. (H1)) and its boundedness away from the origin, which yields a representation of V_δ^s on $\overline{B_\rho(0)}^c$ up to constants. The latter are then determined explicitly in the final step.

Step 1a. For every $\varphi \in C_c^\infty(\overline{\Omega}^c)$, we infer from (4.15) that $D_\delta^s(\mathcal{P}_\delta^s \varphi) = \nabla\varphi = 0$ in Ω, and hence, $\mathcal{P}_\delta^s \varphi \in N^{s,p,\delta}(\Omega)$. By our initial assumption, $\mathcal{P}_\delta^s\varphi$ is then constant on U, which, in view of the identity $\mathcal{P}_\delta^s \varphi = V_\delta^s * \nabla\varphi$ in (4.16), is equivalent to

$$\int_{\mathbb{R}^n} \big(V_\delta^s(x-z) - V_\delta^s(y-z)\big) \cdot \nabla\varphi(z)\, dz = 0 \quad \text{for all } x, y \in U.$$

Since this holds for any $\varphi \in C_c^\infty(\overline{\Omega}^c)$, the fundamental lemma of the calculus of variations in combination with integration by parts allows us to deduce

$$\operatorname{div} V_\delta^s(x-z) = \operatorname{div} V_\delta^s(y-z) \quad \text{for all } x, y \in U \text{ and all } z \in \overline{\Omega}^c. \tag{4.24}$$

Let us fix $x \in U$ and consider $w \in \mathbb{R}^n$ with $|w| > \rho = \operatorname{Diam}(\Omega)$. It follows then that $x - w \in \overline{\Omega}^c$, and we obtain with $z = x - w \in \overline{\Omega}^c$ that

$$\operatorname{div} V_\delta^s(w) = \operatorname{div} V_\delta^s(w + y - x)$$

for all $y \in U$. Taking $y \in B_\varepsilon(x) \subset U$, with $\varepsilon > 0$ sufficiently small, yields

$$\operatorname{div} V_\delta^s(w) = \operatorname{div} V_\delta^s(w') \quad \text{for all } w' \in B_\varepsilon(w).$$

For $n > 1$, this shows that the divergence of V_δ^s is locally constant on the connected set $\overline{B_\rho(0)}^c$, and thus, constant outside of $\overline{B_\rho(0)}$; the case $n = 1$ yields the same conclusion by also utilizing the vector-radiality of V_δ^s.

Step 1b. Recall that V_δ^s is smooth on $\mathbb{R}^n \setminus \{0\}$ and vector-radial, meaning that there is a smooth function $f_\delta^s : (0, \infty) \to \mathbb{R}$ with $V_\delta^s(x) = x f_\delta^s(|x|)$. We can then rewrite the divergence of V_δ^s as

$$\operatorname{div} V_\delta^s(x) = n f_\delta^s(|x|) + |x| (f_\delta^s)'(|x|)$$

[1] Note that the function $x \mapsto \frac{1}{\sigma_{n-1}} \frac{x}{|x|^n}$ for $x \in \mathbb{R}^n$ corresponds exactly to the kernel function appearing in the fundamental theorem of calculus for the classical gradient, see [179, Proposition 4.14].

for $x \in \mathbb{R}^n \setminus \{0\}$. Since this expression is constant on the complement of $\overline{B_\rho(0)}$ by Step 1, the auxiliary function f_δ^s satisfies for all $r > \rho$ the equation

$$nf_\delta^s(r) + r(f_\delta^s)'(r) = c \tag{4.25}$$

with some constant $c \in \mathbb{R}$. The family of solutions to the ordinary differential equation (4.25) is given by $r \mapsto \frac{c}{n} + \frac{k}{r^n}$ with $k \in \mathbb{R}$. Consequently, there is a $k \in \mathbb{R}$ such that

$$V_\delta^s(x) = c\frac{x}{n} + k\frac{x}{|x|^n} \quad \text{for all } x \in \overline{B_\rho(0)}^c.$$

Step 1c. The boundedness of V_δ^s on $\overline{B_\rho(0)}^c$ according to [30, Theorem 5.9 *b*)] implies $c = 0$. To determine k, consider the compactly supported, integrable function

$$x \mapsto V_\delta^s(x) - k\frac{x}{|x|^n},$$

whose Fourier transform is continuous and can be calculated to be

$$\xi \mapsto \frac{-i\xi}{2\pi|\xi|^2}\Big(\frac{1}{\widehat{Q_\delta^s}(\xi)} - k\sigma_{n-1}\Big), \tag{4.26}$$

see [30, Theorem 5.9]. As the first factor in (4.26) has a pole at the origin, the second factor needs to vanish in 0 because of continuity. Due to $\widehat{Q_\delta^s}(0) = \|Q_\delta^s\|_{L^1(\mathbb{R}^n)} = 1$, this eventually yields $k = \sigma_{n-1}^{-1}$, confirming (4.23).

Step 2: Entire extension of $\widehat{Q_\delta^s}$ is zero-free. Let us introduce the auxiliary function $Z_\delta^s \in C^\infty(\mathbb{R}^n \setminus \{0\}; \mathbb{R}^n) \cap L^1(\mathbb{R}^n; \mathbb{R}^n)$ defined by

$$Z_\delta^s(x) = V_\delta^s(x) - \frac{1}{\sigma_{n-1}}\frac{x}{|x|^n}.$$

As Z_δ^s has compact support owing to Step 1, the Paley-Wiener theorem (see e.g., [122, Theorem 2.3.21]) implies that the Fourier transform $\widehat{Z_\delta^s}$ with

$$\widehat{Z_\delta^s}(\xi) = \frac{-i\xi}{2\pi|\xi|^2}\Big(\frac{1}{\widehat{Q_\delta^s}(\xi)} - 1\Big) \quad \text{for } \xi \in \mathbb{R}^n \setminus \{0\}. \tag{4.27}$$

is real analytic and allows for a unique entire extension to a function $\mathbb{C}^n \to \mathbb{C}^n$. An analogous argument gives that the Fourier transform of the kernel function Q_δ^s (see (4.12) and recall $\operatorname{supp}(Q_\delta^s) \subset B_\delta(0)$) is extendable (uniquely) to a holomorphic function $\mathbb{C}^n \to \mathbb{C}$. In the following, we write $\widehat{Z_\delta^s}$ and $\widehat{Q_\delta^s}$ for both the Fourier transforms of Z_δ^s and Q_δ^s, as well as for their extended versions defined on $\mathbb{C}^n$.

The goal in this step is to show that

$$\widehat{Q_\delta^s} : \mathbb{C}^n \to \mathbb{C} \text{ is zero-free.} \tag{4.28}$$

Suppose for the sake of contradiction that this is not the case, and let $\zeta_0 \in \mathbb{C}^n \setminus \{0\}$ be a zero of $\widehat{Q_\delta^s}$ with minimal norm $r := |\zeta_0|$; note that $\widehat{Q_\delta^s}(0) = \|Q_\delta^s\|_{L^1(\mathbb{R}^n)} = 1$. Applying the identity theorem of complex analysis, in each variable separately, to $\widehat{Z_\delta^s}$ as in (4.27) yields that

$$\widehat{Z_\delta^s}(\zeta) = \frac{-i\zeta}{2\pi|\zeta|^2}\Big(\frac{1}{\widehat{Q_\delta^s}(\zeta)} - 1\Big) \quad \text{for } \zeta \in \mathbb{C}^n \setminus \{0\} \text{ with } |\zeta| < r.$$

We now find that $\lim_{r'\uparrow 1} |\widehat{Z}_\delta^s(r'\zeta_0)| = \infty$, which contradicts the continuity of the holomorphic extension of $\widehat{Z}_\delta^s$. Thus, (4.28) is proven.

Step 3: Section of $\widehat{Q}_\delta^s$ coincides with exponential of a polynomial. Consider $q_\delta^s : \mathbb{C} \to \mathbb{C}, z \mapsto \widehat{Q_\delta^s}(ze_1)$ with $e_1 = (1, 0, \dots, 0) \in \mathbb{C}^n$. From $\|Q_\delta^s\|_{L^1(\mathbb{R}^n)} = 1$ and $\operatorname{supp}(Q_\delta^s) \subset B_\delta(0)$, we conclude that

$$\left|q_\delta^s(z)\right| = \left|\widehat{Q}_\delta^s(ze_1)\right| \leq \int_{\mathbb{R}^n} |Q_\delta^s(x)||e^{-2\pi i x_1 z}|\, dx \leq e^{2\pi\delta|z|}$$

for all $z \in \mathbb{C}$, showing that q_δ^s is an entire function of order at most 1. As a consequence of Step 2, this function is never zero, so that the Hadamard factorization theorem (see e.g., [147, Corollary XII.3.3]) yields that

$$q_\delta^s(z) = e^{az+b} \quad \text{for all } z \in \mathbb{C} \text{ with } a, b \in \mathbb{C}.$$

However, this contradicts the fact that $q_\delta^s : z \mapsto \widehat{Q}_\delta^s(ze_1)$ is non-constant and even (cf. Section 4.2.3), as the section of the Fourier transform of the radial kernel Q_δ^s, which proves the statement. □

We point out that the previous result is not true when $\Omega = \mathbb{R}^n$. In fact, $N^{s,p,\delta}(\mathbb{R}^n)$ for $p < \infty$ contains only the zero function, which can be deduced with the help of the translation operators of Section 4.2.3 as follows: Let $u \in N^{s,p,\delta}(\mathbb{R}^n)$, then $v := \mathcal{Q}_\delta^s u \in W^{1,p}(\mathbb{R}^n)$ satisfies $\nabla v = D_\delta^s u = 0$, and hence, $v = 0$, so that $u = \mathcal{P}_\delta^s(\mathcal{Q}_\delta^s u) = \mathcal{P}_\delta^s v = 0$. A similar argument for $p = \infty$, shows that $N^{s,\infty,\delta}(\mathbb{R}^n)$ only consists of constant functions. Nevertheless, there are unbounded sets $\Omega \subset \mathbb{R}^n$ for which $N^{s,p,\delta}(\Omega)$ is non-trivial.

Remark 4.3.2 (Generalization to unbounded sets). Proposition 4.3.1 holds more generally for open sets $\Omega \subset \mathbb{R}^n$ such that $\overline{\Omega}^c$ contains the trace of an unbounded continuous curve $\gamma : [0, \infty) \to \mathbb{R}^n$.

The proof can easily be adjusted, with only a minor modification in Step 1a. After showing (4.24) as above, let us fix $x \in U$ and consider $w = x - z \in \mathbb{R}^n$ for some $z \in \gamma([0, \infty))$. It follows then from (4.24) that $\operatorname{div} V_\delta^s(w) = \operatorname{div} V_\delta^s(w')$ for all $w' \in B_\varepsilon(w)$ with $\varepsilon > 0$ such that $B_\varepsilon(x) \subset U$. By applying this for all $z \in \gamma([0, \infty))$ and exploiting the radial symmetry of the divergence of V_δ^s, we find that $\operatorname{div} V_\delta^s$ is constant in the complement of $\overline{B_\rho(0)}$ with $\rho := \operatorname{dist}(x, \gamma([0, \infty)))$. △

The following result confirms that there exist, in fact, a great many functions with vanishing nonlocal gradients, by showing that they form an infinite-dimensional space.

Proposition 4.3.3 ($N^{s,p,\delta}(\Omega)$ is infinite-dimensional). *Let $p \in [1, \infty]$, then $N^{s,p,\delta}(\Omega)$ is an infinite-dimensional closed subspace of $H^{s,p,\delta}(\Omega)$.*

Proof. The proof idea relies on a semi-explicit construction procedure in three steps: Starting with $u \in H^{s,p,\delta}(\Omega)$, there is a $v \in W^{1,p}(\Omega)$ with $\nabla v = D_\delta^s u$ on Ω by Lemma 4.2.4. Next, we extend v in an arbitrary way to a compactly supported function in $W^{1,p}(\mathbb{R}^n)$. In view of Lemma 4.2.5, it holds that $\mathcal{E}_\delta^s u := \mathcal{P}_\delta^s v \in H^{s,p,\delta}(\mathbb{R}^n)$ satisfies $D_\delta^s(\mathcal{E}_\delta^s u) = D_\delta^s u$ on Ω, or equivalently,

$$h = (\mathcal{E}_\delta^s u)|_{\Omega_\delta} - u \in N^{s,p,\delta}(\Omega). \tag{4.29}$$

Note that $\mathcal{E}_\delta^s u$ can be viewed as an extension operator of u modulo a function with vanishing nonlocal gradient, see Section 4.4.2 for more details.

Let $m \in \mathbb{N}$. With the aim of constructing m linearly independent functions in $N^{s,p,\delta}(\Omega)$, we take $u_1, \dots, u_m \in H^{s,p,\delta}(\Omega)$ with $m \in \mathbb{N}$ such that no (non-trivial) linear combination of them can be extended to a function in $H^{s,p,\delta}(\mathbb{R}^n) = H^{s,p}(\mathbb{R}^n)$. This can be achieved, for instance, if each function u_j has a suitable singularity in different places in the collar Γ_δ.

To give more details, choose $x_1, \dots, x_m$ as distinct points in Γ_δ and let $\varepsilon > 0$ be such that the balls $B_{2\varepsilon}(x_j)$ are pairwise disjoint and compactly contained in Γ_δ. Further, define

$$u_j(x) = \mathbb{1}_{B_\varepsilon(x_j)}(x)u\Big(\frac{x - x_j}{\varepsilon}\Big) \quad \text{for } x \in \Omega_\delta \text{ and } j = 1, \dots, m,$$

where $u \in L^p(\mathbb{R}^n) \setminus H^{s,p,\delta}(\mathbb{R}^n)$ with $\operatorname{supp}(u)$ compactly contained in $\overline{B_1(0)}$ (cf. Example 4.3.4). Note that any function in $L^p(\Omega_\delta)$ that is zero in a compact set containing Ω lies in $H^{s,p,\delta}(\Omega)$. Indeed, the nonlocal gradient is then given by a convolution and defines an L^p-function due to Young's convolution inequality, see Remark 4.2.1 b); the integration by parts formula in (4.9) can be verified via Fubini's theorem. We conclude that $u_j \in H^{s,p,\delta}(\Omega)$ for all j. On the other hand, by construction no u_j has an extension to $H^{s,p,\delta}(\mathbb{R}^n)$, and thus, nor does any linear combination of $u_1, \dots, u_m$.

According to (4.29), we now set $h_j = (\mathcal{E}^s_\delta u_j)|_{\Omega_\delta} - u_j \in N^{s,p,\delta}(\Omega)$ for $j = 1, \dots, m$. If these functions h_j were linearly dependent, then one could find a non-trivial linear combination of $u_1, \dots, u_m$ that can be extended to $\mathbb{R}^n$ via the operator $\mathcal{E}^s_\delta$, which contradicts the assumption. Hence, $h_1, \dots, h_m$ are linearly independent. Because $m \in \mathbb{N}$ is arbitrary, this shows that $N^{s,p,\delta}(\Omega)$ must be infinite-dimensional. □

For the reader's convenience, we give here explicit examples of functions $u \in L^p(\mathbb{R}^n)\setminus H^{s,p,\delta}(\mathbb{R}^n)$ for all $1 \leq p \leq \infty$ with compact support in $\overline{B_1(0)}$, as they were used in the previous proof.

Example 4.3.4. Let $p \in [1, \infty)$. Defining for some $0 < \nu < \min\{\frac{n}{p}, s\}$,

$$u = \mathbb{1}_{B_1(0)} |\cdot|^{-\frac{n}{p}+\nu} \in L^p(\mathbb{R}^n),$$

gives a function with the desired properties if ν is sufficiently small. That $u \notin H^{s,p,\delta}(\mathbb{R}^n)$, follows by contradiction with the estimate in [30, Proposition 7.2], once we have shown the existence of a constant $c > 0$ such that

$$\|u - u(\cdot + h)\|_{L^p(B_1(0))} \geq c\,|h|^\nu \qquad \text{for all } h \in B_1(0). \tag{4.30}$$

Indeed, we have that

$$\begin{aligned}
\|u - u(\cdot + h)\|_{L^p(B_1(0))} &\geq \|u\|_{L^p(B_{|h|/2}(0))} - \|u(\,\cdot + h)\|_{L^p(B_{|h|/2}(0))} \\
&\geq \left(\int_{B_{|h|/2}(0)} |x|^{-n+\nu p}\, dx\right)^{1/p} - \left(\left|B_{|h|/2}(0)\right| (|h|/2)^{-n+\nu p}\right)^{1/p} \\
&\geq \left(\frac{c_1}{(\nu p)^{1/p}} - c_2\right) |h|^\nu,
\end{aligned}$$

with $c_1, c_2 > 0$, so that (4.30) follows for small ν.

For larger p there are also more elementary examples, such as the indicator function of a ball when $1 < sp < \infty$, or any discontinuous function when $n < sp < \infty$ (see e.g., [28, Section 2.1]). For the case $p = \infty$, we can take u to be any discontinuous function with support in $B_1(0)$. Indeed, if it were true that $u \in H^{s,\infty,\delta}(\mathbb{R}^n)$, then we also find that $u \in H^{s,q,\delta}(\mathbb{R}^n)$ for all $q \in [1, \infty)$ given its compact support. This would yield by [30, Theorem 6.3], that u is Hölder continuous up to order s, which gives a contradiction.

4.3.2 Characterization of $N^{s,p,\delta}(\Omega)$

Now that the presence of non-constant functions with zero nonlocal gradient is confirmed, the next task is to understand - and eventually, characterize - all functions in $N^{s,p,\delta}(\Omega)$.

We start by observing that a function $u \in L^p(\Omega_\delta)$ lies in $N^{s,p,\delta}(\Omega)$ if and only if $\mathcal{Q}^s_\delta u = Q^s_\delta * u$ is constant (in Ω). Indeed, if $Q^s_\delta * u$ is constant, then

$$\int_{\Omega_\delta} u \operatorname{div}^s_\delta \varphi \, dx = \int_{\mathbb{R}^n} u \, Q^s_\delta * \operatorname{div} \varphi \, dx = \int_\Omega (Q^s_\delta * u) \operatorname{div} \varphi \, dx = 0 \quad \text{for all } \varphi \in C_c^\infty(\Omega; \mathbb{R}^n),$$

which shows that $u \in H^{s,p,\delta}(\Omega)$ with $D^s_\delta u = 0$. Conversely, if $u \in N^{s,p,\delta}(\Omega)$, it follows from Lemma 4.2.4 that $\mathcal{Q}^s_\delta u \in W^{1,p}(\Omega)$ has zero gradient, and is thus, constant. The desired characterization of $N^{s,p,\delta}(\Omega)$ therefore comes down to identifying all solutions $h \in L^p(\Omega_\delta)$ for convolution equations of the form

$$Q^s_\delta * h = c \ \text{ a.e. in } \Omega$$

for any $c \in \mathbb{R}$.

Our strategy for characterizing $N^{s,p,\delta}(\Omega)$ starts at a suitable boundary-value problem involving the convolution operator $\mathcal{Q}^s_\delta$. Via inversion of $\mathcal{Q}^s_\delta$ (based on the translation tools from Section 4.2.3), we then rewrite the latter equivalently as a pseudo-differential equation featuring $\mathcal{P}^s_\delta$ subject to Dirichlet boundary conditions, for which a solution theory can be achieved. Let us make the mentioned equivalence precise.

Lemma 4.3.5 (Equivalence between (C) and (P)). *Let $p \in (1, \infty)$, $c \in \mathbb{R}$ and $g \in L^p(\Gamma_\delta)$. Further, let $\Omega' \subset \mathbb{R}^n$ be any smooth and bounded set with $\Omega_{2\delta} \subset \Omega'$ and $\Gamma' := \Omega' \setminus \overline{\Omega}$. Then,*

$$\text{(C)} \quad \begin{cases} \mathcal{Q}^s_\delta h = c & \text{a.e. in } \Omega, \\ h = g & \text{a.e. in } \Gamma_\delta, \end{cases}$$

has a solution $h \in L^p(\Omega_\delta)$ if and only if there exists a $w \in H^{1-s,p}(\mathbb{R}^n)$ solving

$$\text{(P)} \quad \begin{cases} \mathcal{P}^s_\delta w = \mathbb{1}_{\Gamma_\delta} g & \text{a.e. in } \Gamma', \\ w = c & \text{a.e. in } \overline{\Omega}, \\ w = 0 & \text{a.e. in } (\Omega')^c. \end{cases}$$

Specifically, the following holds:

(*i*) *If $h \in L^p(\Omega_\delta)$ is a solution to* (C), *then $w := \mathcal{Q}^s_\delta(\mathbb{1}_{\Omega_\delta} h) \in H^{1-s,p}(\mathbb{R}^n)$ solves* (P).

(*ii*) *If $w \in H^{1-s,p}(\mathbb{R}^n)$ satisfies* (P), *then $h := (\mathcal{P}^s_\delta w)|_{\Omega_\delta} \in L^p(\Omega_\delta)$ is a solution to* (C).

Proof. The main ingredient of this proof is (4.19) with $\beta = 1 - s$, according to which $\mathcal{P}^s_\delta$ is a isomorphism from $H^{1-s,p}(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$ with inverse $\mathcal{Q}^s_\delta$.

The implication (i) follows then immediately from the observation that

$$\mathcal{P}^s_\delta w = \mathcal{P}^s_\delta \mathcal{Q}^s_\delta(\mathbb{1}_{\Omega_\delta} h) = \mathbb{1}_{\Omega_\delta} h = \mathbb{1}_{\Gamma_\delta} g \quad \text{ a.e. in } \Omega^c,$$

along with the property that $\operatorname{supp}(Q^s_\delta) \subset B_\delta(0)$, which implies $\operatorname{supp}(w) \subset \Omega_\delta + B_\delta(0) \subset \Omega_{2\delta} \subset \Omega'$ as well as $w = \mathcal{Q}^s_\delta h = c$ a.e. in $\overline{\Omega}$ given $|\partial\Omega| = 0$.

On the other hand, (ii) holds since $\mathcal{Q}^s_\delta h = \mathcal{Q}^s_\delta(\mathcal{P}^s_\delta w) = w = c$ a.e. in Ω and $h = \mathcal{P}^s_\delta w = g$ a.e. in Γ_δ, again using that $|\partial\Omega| = 0$. □

Based on the previous lemma, we can express $N^{s,p,\delta}(\Omega)$ in terms of the solution sets of the boundary-value problems (C) and (P). To be precise, let

$$\mathfrak{C}^{s,p,\delta}(\Omega) := \bigcup_{c\in\mathbb{R},g\in L^p(\Gamma_\delta)} \mathfrak{C}^{s,p,\delta}(c,g) \qquad \text{and} \qquad \mathfrak{P}^{s,p,\delta}(\Omega) := \bigcup_{c\in\mathbb{R},g\in L^p(\Gamma_\delta)} \mathfrak{P}^{s,p,\delta}(c,g), \tag{4.31}$$

where $\mathfrak{C}^{s,p,\delta}(c,g)$ for $c \in \mathbb{R}$ and $g \in L^p(\Gamma_\delta)$ denotes the set of all solutions in $L^p(\Omega_\delta)$ to (C), and $\mathfrak{P}^{s,p,\delta}(c,g)$ comprises the functions in $H^{1-s,p}(\mathbb{R}^n)$ solving (P). Then,

$$N^{s,p,\delta}(\Omega) = \mathfrak{C}^{s,p,\delta}(\Omega) = \mathcal{P}^s_\delta\big(\mathfrak{P}^{s,p,\delta}(\Omega)\big)|_{\Omega_\delta}. \tag{4.32}$$

We take (4.32) as motivation to turn our attention to (P) and address the question of its solvability. It turns out that the recent existence and uniqueness results by Grubb [125] in combination with the regularity theory in [2] by Abels & Grubb for general pseudo-differential operators (see Theorem 4.2.6 for a version tailored to our setting) provides an answer. Even though the next lemma is a direct application of this abstract framework, we have included for illustration also a hands-on alternative proof in the case $p = 2$.

Lemma 4.3.6 (Existence and uniqueness for (P)). *Let $p \in (1,\infty)$ and Ω be a bounded $C^{1,1}$-domain. Then, for every $c \in \mathbb{R}$ and $g \in L^p(\Gamma_\delta)$, the problem* (P) *admits a unique solution $w_{c,g} \in H^{\frac{1-s}{2},p}(\mathbb{R}^n)$. If $p \in (1, \frac{2}{1-s})$, then $w_{c,g} \in H^{1-s,p}(\mathbb{R}^n)$.*

Proof. We may assume without loss of generality that $c = 0$, since for every $w \in \mathfrak{P}^{s,p,\delta}(0, g-c)$ it holds that $w + \mathcal{Q}^s_\delta(\mathbb{1}_{\Omega_\delta} c) \in \mathfrak{P}^{s,p,\delta}(c,g)$. Hence, we can focus our attention to the pseudo-differential equation

$$\begin{cases} \mathcal{P}^s_\delta w = \mathbb{1}_{\Gamma_\delta} g & \text{a.e. in } \Gamma', \\ w = 0 & \text{a.e. in } (\Gamma')^c. \end{cases} \tag{4.33}$$

The statement now follows immediately by applying Theorem 4.2.6 for the pseudo-differential operator $\mathcal{P} = \mathcal{P}^s_\delta$ and the set $V = \Gamma'$. Indeed, $\mathcal{P}^s_\delta$ satisfies all the required assumptions according to Lemma 4.2.8 and $\Gamma' = \Omega' \setminus \overline{\Omega}$ is a bounded open set with $C^{1,1}$-boundary, given that $\partial\Gamma' = \partial\Omega' \cup \partial\Omega$.

Our alternative proof for $p = 2$ relies on a familiar variational argument and exploits regularity results for the fractional Laplacian. Let us consider the operator

$$\mathcal{L}^s_\delta : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n), \quad \widehat{\mathcal{L}^s_\delta \varphi} = \frac{\widehat{\varphi}}{\sqrt{\widehat{Q^s_\delta}}},$$

which can be extended to a bounded linear operator $H^t(\mathbb{R}^n) \to H^{t-\frac{1-s}{2}}(\mathbb{R}^n)$ for any $t \in \mathbb{R}$. Then, $(\mathcal{L}^s_\delta)^2 = \mathcal{L}^s_\delta \circ \mathcal{L}^s_\delta = \mathcal{P}^s_\delta$, and we observe that $\|\mathcal{L}^s_\delta \cdot\|_{L^2(\mathbb{R}^n)}$ is a norm on $H^{\frac{1-s}{2}}(\mathbb{R}^n)$ that is equivalent to $\|\cdot\|_{H^{\frac{1-s}{2}}(\mathbb{R}^n)}$ in view of (4.18).

As a consequence of the generalized Dirichlet principle, the functional

$$w \mapsto \|\mathcal{L}^s_\delta w\|^2_{L^2(\mathbb{R}^n)} - \int_{\mathbb{R}^n} \mathbb{1}_{\Gamma_\delta} g\, w\, dx$$

over all functions $w \in H_0^{\frac{1-s}{2}}(\Gamma')$ has a unique minimizer w_*, which is also the unique solution to

$$\int_{\mathbb{R}^n} \mathcal{L}^s_\delta w\, \mathcal{L}^s_\delta \varphi - \mathbb{1}_{\Gamma_\delta} g \varphi\, dx = 0 \quad \text{for all } \varphi \in C_c^\infty(\Gamma'). \tag{4.34}$$

Since (4.34) is a weak formulation of (4.33) for $p = 2$, the function $w_* \in H_0^{\frac{1-s}{2}}(\Gamma')$ is indeed the only candidate for the sought solution.

It remains to prove that $w_* \in H^{1-s}(\mathbb{R}^n)$. To this end, we compare the operator $\mathcal{L}_\delta^s$ with the fractional Laplacian $(-\Delta)^{\frac{1-s}{4}} : H^{\frac{1-s}{2}}(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$, showing that they differ by a bounded linear operator on $L^2(\mathbb{R}^n)$. Indeed, $\mathcal{K}_\delta^s := \mathcal{L}_\delta^s - (-\Delta)^{\frac{1-s}{4}}$ is an L^2-Fourier multiplier operator with multiplier

$$m_\delta^s(\xi) = \frac{1}{\sqrt{\widehat{Q}_\delta^s(\xi)}} - |2\pi\xi|^{\frac{1-s}{2}} \qquad \text{for } \xi \in \mathbb{R}^n.$$

The boundedness of m_δ^s follows from the smoothness and positivity of $\widehat{Q}_\delta^s$ together with the observation that for $|\xi| > 1$,

$$m_\delta^s(\xi) = \left(\frac{1}{\sqrt{1 + |2\pi\xi|^{1-s} R_\delta^s(\xi)}} - 1\right)|2\pi\xi|^{\frac{1-s}{2}},$$

which is bounded since R_δ^s (cf. (4.17)) is a Schwartz function. A particular consequence is that $\mathcal{K}_\delta^s(-\Delta)^{\frac{1-s}{4}} = (-\Delta)^{\frac{1-s}{4}}\mathcal{K}_\delta^s : H^{\frac{1-s}{2}}(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a bounded linear operator.

Then, (4.34) turns into

$$\int_{\mathbb{R}^n} (-\Delta)^{\frac{1-s}{4}} w \, (-\Delta)^{\frac{1-s}{4}} \varphi + \big(2\mathcal{K}_\delta^s(-\Delta)^{\frac{1-s}{4}} w + (\mathcal{K}_\delta^s)^2 w - \mathbb{1}_{\Gamma_\delta} g\big)\varphi \, dx = 0 \quad \text{for all } \varphi \in C_c^\infty(\Gamma'),$$

which implies that w_* weakly satisfies

$$\begin{cases} (-\Delta)^{\frac{1-s}{2}} w = 2\mathcal{K}_\delta^s(-\Delta)^{\frac{1-s}{4}} w + (\mathcal{K}_\delta^s)^2 w - \mathbb{1}_{\Gamma_\delta} g & \text{in } \Gamma', \\ w = 0 & \text{in } (\Gamma')^c. \end{cases}$$

Since the right-hand side of the fractional differential equation lies in $L^2(\Gamma')$ and the boundary $\partial\Gamma' = \partial\Omega' \cup \partial\Omega$ is $C^{1,1}$, we obtain from established regularity results for the fractional Laplacian (see, e.g., [124, 151, 207] for smooth domains and [2, Theorem 1.1] for $C^{1,1}$-domains) that $w \in H^{1-s}(\mathbb{R}^n)$, as desired. □

Remark 4.3.7. The range of p for which $w_g \in H^{1-s,p}(\mathbb{R}^n)$ is sharp, since even for smooth Ω, one can find $g \in L^p(\Gamma_\delta)$ such that $w_g \notin H^{1-s,p}(\mathbb{R}^n)$ when $p \in [\frac{2}{1-s}, \infty)$. To see this, let us take w equal to $\mathrm{dist}(\cdot, \partial\Gamma')^{\frac{1-s}{2}}$ near the boundary of Γ' with $\mathcal{P}_\delta^s w|_{\Gamma'} \in C^\infty(\mathbb{R}^n)|_{\Gamma'}$ as in Remark 4.2.7 b). Then, we define

$$\tilde{w} := \mathcal{Q}_\delta^s(\mathbb{1}_{\Omega_\delta}\mathcal{P}_\delta^s w),$$

which coincides with w in a neighborhood of Ω, given (H2). Therefore, $\tilde{w}$ equals $\mathrm{dist}(\cdot, \partial\Omega)^{\frac{1-s}{2}}$ near the boundary of Ω, which yields $\tilde{w} \notin H^{1-s,p}(\mathbb{R}^n)$. However, since $\tilde{w} = w = 0$ in Ω and $\mathrm{supp}(\tilde{w}) \subset \Omega_{2\delta}$, we deduce that

$$\begin{cases} \mathcal{P}_\delta^s \tilde{w} = \mathbb{1}_{\Gamma_\delta}\mathcal{P}_\delta^s w & \text{a.e. in } \Gamma', \\ \tilde{w} = 0 & \text{a.e. in } (\Gamma')^c. \end{cases}$$

With $g := \mathcal{P}_\delta^s w|_{\Gamma_\delta} \in L^p(\Gamma_\delta)$, the claim follows. △

Lemma 4.3.6 in combination with Lemma 4.3.5 provide useful information about the solution sets for the boundary-value problems (P) and (C). Since the statement of Lemma 4.3.6 is qualitatively different depending on whether p is smaller or larger than the critical value $\frac{2}{1-s}$, we discuss these two cases separately.

Suppose first that $p \in (1, \frac{2}{1-s})$. Then, for any $c \in \mathbb{R}$ and $g \in L^p(\Gamma_\delta)$, the sets $\mathfrak{P}^{s,p,\delta}(c, g)$ and $\mathfrak{C}^{s,p,\delta}(c, g)$ are singletons and can be represented as

$$\mathfrak{P}^{s,p,\delta}(c, g) = \{w_{c,g}\} \qquad \text{and} \qquad \mathfrak{C}^{s,p,\delta}(c, g) = \{h_{c,g}\}, \tag{4.35}$$

where $w_{c,g} \in H^{1-s,p}(\mathbb{R}^n)$ and $h_{c,g} := (\mathcal{P}^s_\delta w_{c,g})|_{\Omega_\delta} \in L^p(\Omega_\delta)$ are the unique solutions to (P) and (C), respectively.

Summarizing these findings, we are now in the position to state the main result of this section concerning the characterization of $N^{s,p,\delta}(\Omega)$.

Theorem 4.3.8 (Characterization of $N^{s,p,\delta}(\Omega)$ for $p \in (1, \frac{2}{1-s})$). *Let $p \in (1, \frac{2}{1-s})$ and let Ω be a bounded $C^{1,1}$-domain. Then,*

$$N^{s,p,\delta}(\Omega) = \bigcup_{c\in\mathbb{R}, g\in L^p(\Gamma_\delta)} \mathfrak{C}^{s,p,\delta}(c, g) = \bigcup_{c\in\mathbb{R}, g\in L^p(\Gamma_\delta)} \{h_{c,g}\}, \tag{4.36}$$

where $h_{c,g}$ for $c \in \mathbb{R}$ and $g \in L^p(\Gamma_\delta)$ is the unique solution of (C).

Proof. This follows immediately from (4.31), (4.32) and (4.35). □

As a consequence of Theorem 4.3.8, we obtain that the bounded linear map

$$\Phi^s_\delta : N^{s,p,\delta}(\Omega) \to \mathbb{R} \times L^p(\Gamma_\delta), \quad h \mapsto \Big(\int_\Omega Q^s_\delta * h \, dx, h|_{\Gamma_\delta}\Big) \tag{4.37}$$

is bijective. The inverse $(\Phi^s_\delta)^{-1} : (c, g) \mapsto h_{c,g}$ for $(c, g) \in \mathbb{R} \times L^p(\Gamma_\delta)$ is then bounded as well by the Banach isomorphism theorem, i.e., there is a constant $C > 0$ such that

$$\|h_{c,g}\|_{L^p(\Omega_\delta)} \leq C\big(\|g\|_{L^p(\Gamma_\delta)} + |c|\big) \quad \text{for all } g \in L^p(\Gamma_\delta) \text{ and } c \in \mathbb{R}.$$

The discussion above implies that the functions with zero nonlocal gradient are uniquely determined by their values in the single collar Γ_δ and an averaging condition involving the kernel function Q^s_δ. Besides, one can also observe a one-to-one correspondence between functions in $N^{s,p,\delta}(\Omega)$ and these two basic characteristics: the boundary values in Γ_δ and the mean value in Ω. Indeed, another isomorphism between $N^{s,p,\delta}(\Omega)$ and $\mathbb{R} \times L^p(\Gamma_\delta)$ is given by

$$\Psi^s_\delta : N^{s,p,\delta}(\Omega) \to \mathbb{R} \times L^p(\Gamma_\delta), \quad h \mapsto \Big(\int_\Omega h \, dx, h|_{\Gamma_\delta}\Big), \tag{4.38}$$

which follows essentially from the next proposition.

Proposition 4.3.9 (Uniqueness in $N^{s,p,\delta}(\Omega)$ with vanishing mean value). *Let $p \in (1, \infty)$ and Ω be a bounded $C^{1,1}$-domain. If $h \in N^{s,p,\delta}(\Omega)$ satisfies*

$$h = 0 \text{ a.e. in } \Gamma_\delta \quad \text{and} \quad \int_\Omega h \, dx = 0,$$

then $h = 0$.

Proof. If $h \in N^{s,p,\delta}(\Omega)$ with $h = 0$ a.e. in Γ_δ, then (4.36) implies the existence of a $c \in \mathbb{R}$ such that $h = h_{c,0} \in L^p(\Omega_\delta)$. If $p \in [\frac{2}{1-s}, \infty) > 2$, it is clear that $h \in L^2(\Omega_\delta)$. Otherwise, this property follows from

$$h_{c,0} \in \mathfrak{C}^{s,p,\delta}(c,0) = \mathfrak{C}^{s,2,\delta}(c,0) \subset L^2(\Omega_\delta),$$

see (4.39) below. We exploit $\int_\Omega h\,dx = 0$ and $\mathrm{supp}(Q_\delta^s) \subset B_\delta(0)$ to find

$$\begin{aligned}\|\widehat{\mathbb{1}_{\Omega_\delta} h}\|_{L^2(\mathbb{R}^n)} = \|h\|_{L^2(\Omega)} &= \left\| h - Q_\delta^s * h - |\Omega|^{-1}\int_\Omega h - Q_\delta^s * h\,dx \right\|_{L^2(\Omega)} \\ &\leq \|h - Q_\delta^s * h\|_{L^2(\Omega)} \leq \|\mathbb{1}_{\Omega_\delta} h - Q_\delta^s * (\mathbb{1}_{\Omega_\delta} h)\|_{L^2(\mathbb{R}^n)} \leq \|(1-\widehat{Q_\delta^s})\widehat{\mathbb{1}_{\Omega_\delta} h}\|_{L^2(\mathbb{R}^n)}.\end{aligned}$$

Since $0 \leq 1 - \widehat{Q_\delta^s} < 1$ (see (4.13) and Remark 4.2.3 b)), we deduce that $\widehat{\mathbb{1}_{\Omega_\delta} h} = 0$, and hence, $h = 0$, as stated. □

Remark 4.3.10 (Uniqueness in $N^{s,p,\delta}(\Omega)$ with enlarged single layer). The mean value condition in the previous proposition can be removed in exchange for replacing the Dirichlet condition in the single layer Γ_δ by a Dirichlet condition in $\Omega_\delta \setminus \overline{O}$ for any $O \Subset \Omega$. This means, if $h \in N^{s,p,\delta}(\Omega)$ satisfies $h = 0$ a.e. in $\Omega_\delta \setminus \overline{O}$, then $h = 0$.

Indeed, let $O' \subset \Omega$ be smooth with $O \Subset O' \Subset \Omega$ and take $\eta \in L^1(\mathbb{R}^n)$ with $\|\eta\|_{L^1(\mathbb{R}^n)} = 1$ and $\mathrm{supp}(\eta) \subset B_\varepsilon(0)$ for $\varepsilon > 0$ sufficiently small. Since there is a $c \in \mathbb{R}$ such that $Q_\delta^s * h = c$ a.e. in Ω with $h = 0$ a.e. in $\Omega_\delta \setminus \overline{O}$, we find that the convolution $h_\eta := \eta * h$ satisfies

$$Q_\delta^s * h_\eta = c \text{ a.e. in } O' \text{ with } h_\eta = 0 \text{ a.e. in } O'_\delta \setminus \overline{O'}.$$

By the uniqueness statements in Theorem 4.3.8 and Proposition 4.3.12 (applied to the set O'), it follows that $h = h_\eta$ for all such η, which is only possible for $h = 0$. △

Remark 4.3.11 (Equivalent norms on $N^{s,p,\delta}(\Omega)$). Based on the isomorphisms Ψ_δ^s and Φ_δ^s we conclude that defining

$$|||h|||_{N^{s,p,\delta}(\Omega)} := \|h\|_{L^p(\Gamma_\delta)} + \left|\int_\Omega Q_\delta^s * h\,dx\right|$$

and

$$|||h|||_{N^{s,p,\delta}(\Omega)} := \|h\|_{L^p(\Gamma_\delta)} + \left|\int_\Omega h\,dx\right|$$

for $h \in N^{s,p,\delta}(\Omega)$ yields two norms on $N^{s,p,\delta}(\Omega)$ that are equivalent to $\|\cdot\|_{L^p(\Omega_\delta)}$ when $p \in (1, \frac{2}{1-s})$. △

Finally, we study the case $p \in [\frac{2}{1-s}, \infty)$. Before stating the corresponding representation result for $N^{s,p,\delta}(\Omega)$, we collect a few further observations about the solutions to (C) and (P). Suppose in the following that $c \in \mathbb{R}$ and $g \in L^p(\Gamma_\delta)$. By Lemma 4.3.6, the solution set $\mathfrak{P}^{s,p,\delta}(c,g)$ has at most cardinality 1; specifically,

$$\mathfrak{P}^{s,p,\delta}(c,g) = \{w_{c,g}\} \cap H^{1-s,p}(\mathbb{R}^n).$$

Since $L^p(\Omega_\delta) \subset L^q(\Omega_\delta)$ for all $1 < q \leq p$, Lemma 4.3.6 implies also that $w_{c,g} \in H^{1-s,q}(\mathbb{R}^n)$ for all $q \in (1, \frac{2}{1-s})$. In particular, this shows that the solution sets of (P) are independent of the integrability parameters, in the sense that

$$\mathfrak{P}^{s,q,\delta}(c,g) = \mathfrak{P}^{s,2,\delta}(c,g) \quad \text{for all } q \in (1, \tfrac{2}{1-s}),$$

so that we can conclude

$$\mathfrak{P}^{s,p,\delta}(c,g) = \mathfrak{P}^{s,2,\delta}(c,g) \cap H^{1-s,p}(\mathbb{R}^n).$$

Considering the one-to-one relation between the solutions of (C) and (P) (see Lemma 4.3.5), the above properties of $\mathfrak{P}^{s,p,\delta}(c,g)$ carry over to $\mathfrak{C}^{s,p,\delta}(c,g)$. Hence,

$$\mathfrak{C}^{s,q,\delta}(c,g) = \mathfrak{C}^{s,2,\delta}(c,g) \quad \text{for all } q \in (1, \tfrac{2}{1-s}), \tag{4.39}$$

and along with (4.19),

$$\mathfrak{C}^{s,p,\delta}(c,g) = \mathfrak{C}^{s,2,\delta}(c,g) \cap L^p(\Omega_\delta). \tag{4.40}$$

The latter gives rise to the next result.

Proposition 4.3.12 (Characterization of $N^{s,p,\delta}(\Omega)$ for $p \in [\frac{2}{1-s}, \infty)$). *Let $p \in [\frac{2}{1-s}, \infty)$ and let Ω be a bounded $C^{1,1}$-domain. Then,*

$$N^{s,p,\delta}(\Omega) = N^{s,2,\delta}(\Omega) \cap L^p(\Omega_\delta),$$

where $N^{s,2,\delta}(\Omega)$ can be characterized as in Theorem 4.3.8.

Proof. The combination of (4.31), (4.32), and (4.40) proves the claim. □

Remark 4.3.13. The maps Φ^s_δ and Ψ^s_δ from (4.37) and (4.38) can be defined analogously when $p \in [\frac{2}{1-s}, \infty)$. While Proposition 4.3.12 shows that they are still injective, surjectivity generally fails in view of Remark 4.3.7. This shows that not all boundary values in $L^p(\Gamma_\delta)$ can be attained by functions in $N^{s,p,\delta}(\Omega)$, in contrast to the case $p \in (1, \frac{2}{1-s})$. △

4.3.3 Regularity properties of functions with zero nonlocal gradient and examples

In this section, we dive deeper into some of the properties of functions with zero nonlocal gradient, such as their regularity, and we will show some numerical examples to illustrate how they generally behave.

We start off by showing that all functions in $N^{s,p,\delta}(\Omega)$ are smooth inside Ω.

Corollary 4.3.14 (Functions with vanishing nonlocal gradient are smooth in Ω). *Let $p \in (1,\infty)$, then every $h \in N^{s,p,\delta}(\Omega)$ satisfies*

$$h|_\Omega \in C^\infty(\Omega).$$

Proof. It suffices to prove the statement for $h \in \mathfrak{C}^{s,p,\delta}(0,g)$ with $g \in L^p(\Gamma_\delta)$, given (4.31) and the fact that $\mathfrak{C}^{s,p,\delta}(c,g) = \mathfrak{C}^{s,p,\delta}(0,g-c) + c$ for all $c \in \mathbb{R}$. If $h \in \mathfrak{C}^{s,p,\delta}(0,g)$, we deduce from Lemma 4.3.5 that

$$h = (\mathcal{P}^s_\delta w)|_{\Omega_\delta} \quad \text{for a } w \in \mathfrak{P}^{s,p,\delta}(0,g) \subset H^{1-s,p}_0(\Gamma').$$

To see that the restriction $(\mathcal{P}^s_\delta w)|_\Omega$ is smooth, we argue as follows. For any $\varepsilon > 0$ sufficiently small and $\psi \in C^\infty(\mathbb{R}^n)$ with $\psi = \operatorname{div} V^s_\delta$ on $B_\varepsilon(0)^c$, it holds that

$$\mathcal{P}^s_\delta w = \operatorname{div} V^s_\delta * w = \psi * w \quad \text{on } \mathbb{R}^n \setminus (\operatorname{supp}(w))_\varepsilon;$$

this follows from (4.16) via integration by parts, along with an approximation argument. Consequently, $\mathcal{P}^s_\delta w$ can be expressed as the convolution of a compactly supported L^p-function with a smooth function on $\mathbb{R}^n \setminus (\Omega^c)_\varepsilon$ for any ε, and is thus smooth on the union of all these sets, which is Ω. □

In general, we do not expect that functions in $N^{s,p,\delta}(\Omega)$ will be regular on the larger domain Ω_δ given Remark 4.3.7, see also Figure 4.2 and 4.3 below. However, there do exist smooth non-constant functions in $N^{s,p,\delta}(\Omega)$ (cf. Proposition 4.3.1) and they are exactly those that can be obtained from the translation mechanism.

Proposition 4.3.15 (Functions in $N^{s,p,\delta}(\Omega)$ with extra regularity). *Let $p \in [1,\infty]$, then it holds that*

$$N^{s,p,\delta}(\Omega) \cap H^{s,p,\delta}(\mathbb{R}^n)|_{\Omega_\delta} = \{\mathcal{P}^s_\delta v|_{\Omega_\delta} \,:\, v \in W^{1,p}(\mathbb{R}^n) \text{ with } \nabla v = 0 \text{ a.e. on } \Omega\},$$

and

$$N^{s,p,\delta}(\Omega) \cap C_c^\infty(\mathbb{R}^n)|_{\Omega_\delta} = \{\mathcal{P}^s_\delta v|_{\Omega_\delta} \,:\, v \in C_c^\infty(\mathbb{R}^n) \text{ with } \nabla v = 0 \text{ on } \Omega\}.$$

Proof. Let $u \in N^{s,p,\delta}(\Omega) \cap H^{s,p,\delta}(\mathbb{R}^n)|_{\Omega_\delta}$ and consider an extension $\tilde{u} \in H^{s,p,\delta}(\mathbb{R}^n)$ of u. Then, we find by Lemma 4.2.4 that $v := \mathcal{Q}^s_\delta \tilde{u} \in W^{1,p}(\mathbb{R}^n)$ with $\nabla v = D^s_\delta u = 0$ a.e. on Ω. Moreover, by Lemma 4.2.5 we find that

$$\mathcal{P}^s_\delta v|_{\Omega_\delta} = \tilde{u}|_{\Omega_\delta} = u,$$

as desired. On the other hand, if $u = \mathcal{P}^s_\delta v|_{\Omega_\delta}$ for $v \in W^{1,p}(\mathbb{R}^n)$ with $\nabla v = 0$ a.e. on Ω, then $u \in H^{s,p,\delta}(\mathbb{R}^n)|_{\Omega_\delta}$ and $D^s_\delta u = \nabla \mathcal{Q}^s_\delta u = \nabla v = 0$ a.e. on Ω. This proves the first identification.

For the smooth case we can argue in the same way, by also using that $\mathcal{Q}^s_\delta$ and $\mathcal{P}^s_\delta$ map smooth functions to smooth functions (cf. (4.16)). Note that $\mathcal{P}^s_\delta$ might not preserve the compact support, but this is not an issue since $C_c^\infty(\mathbb{R}^n)|_{\Omega_\delta} = C^\infty(\mathbb{R}^n)|_{\Omega_\delta}$. □

We close this section with an illustration of selected one-dimensional examples of functions with zero nonlocal gradient. Figure 4.2 depicts a numerical approximation of the unique function $h_{c,g} \in \mathfrak{C}^{s,2,\delta}(c,g)$ with $c = 0$ and $g \equiv -1$ on Γ_δ. While $h_{c,g} \in L^q(\Omega_\delta)$ for all $q \in (1, \frac{2}{1-s})$ according to (4.39), we see in the first plot that this function has a jump singularity at the boundary of the domain $\Omega = (-3,3)$. This indicates that $h_{c,g}$ might not lie in $L^p(\Omega_\delta)$ for all $p \in (1,\infty)$, reflecting the observations from Remark 4.3.7 and 4.3.13. Moreover, while one may expect from the first illustration that the function is constant on a sub-interval of $(-3,3)$, the enlarged plots show that this is not the case. Indeed, $h_{c,g}$ seems to be displaying oscillations with decreasing amplitude, which is in line with the fact that functions in $N^{s,p,\delta}(\Omega)$ need not be constant on any subset of Ω (cf. Proposition 4.3.1). It is an interesting topic for further study to understand these oscillatory patterns better, and to see if all non-constant functions in $N^{s,p,\delta}(\Omega)$ have a similar behavior.

In Figure 4.3, there are two further examples of functions with zero nonlocal gradient. The left-hand example is similar to the one from Figure 4.2, but with different boundary values. It still features jump singularities at the boundary, and is nearly constant away from the boundary. The right-hand example in Figure 4.3 shows a function with zero nonlocal gradient constructed via the characterization in Proposition 4.3.15. In contrast to the other examples, this one does not have a jump singularity at the boundary. By construction, it is smooth and an element of $N^{s,p,\delta}(\Omega)$ for all $p \in [1,\infty]$.

4.4 Technical tools involving functions with zero nonlocal gradient

In this section, we present several results regarding the function spaces $H^{s,p,\delta}(\Omega)$ in which the set $N^{s,p,\delta}(\Omega)$ plays an important role. We start off with a bounded-domain analogue of the isomorphism between $H^{s,p,\delta}(\mathbb{R}^n)$ and $W^{1,p}(\mathbb{R}^n)$ from [72, Section 2.4] that turns nonlocal gradients into gradients. Subsequently, we study extensions of functions in $H^{s,p,\delta}(\Omega)$ to the whole space $\mathbb{R}^n$ and prove new nonlocal Poincaré and Poincaré-Wirtinger inequalities and compactness results.

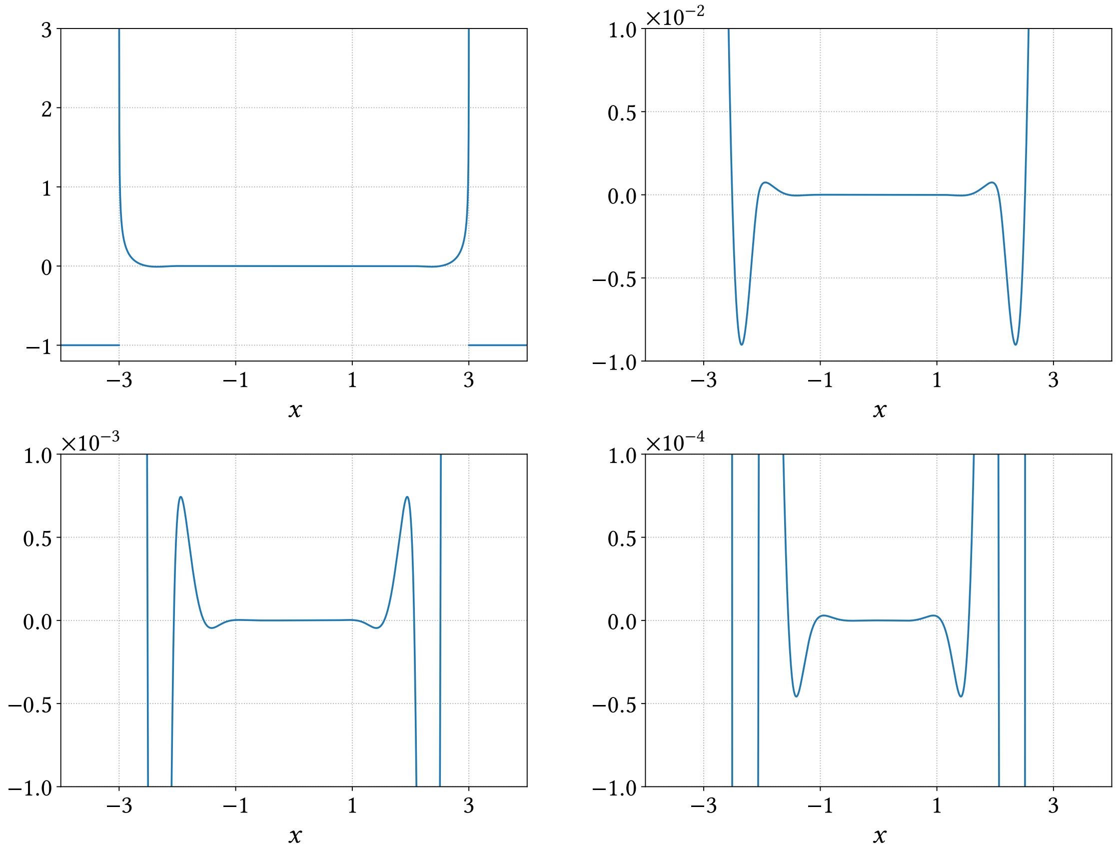

Figure 4.2: Numerical approximation of the function $h_{c,g} \in N^{s,2,\delta}(\Omega)$ for $c = 0$ and $g \equiv -1$ on Γ_δ with increasing degrees of zoom. The parameters for the computation are $n = 1$, $\Omega = (-3, 3)$, $s = \frac{1}{2}$, $\delta = 1$ and $w_\delta \in C_c^\infty(-1, 1)$ is a bump function equal to 1 on $(-\frac{1}{2}, \frac{1}{2})$.

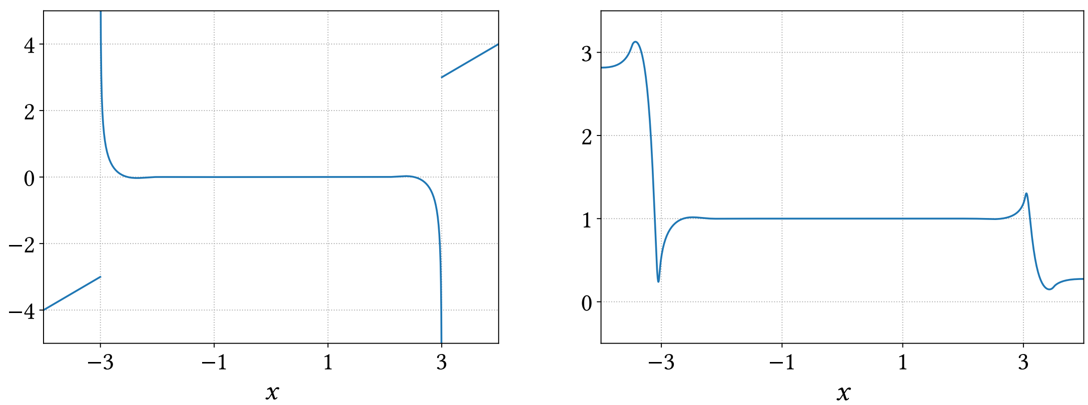

Figure 4.3: Left: A numerical approximation of the function $h_{c,g} \in N^{s,2,\delta}(\Omega)$ with $c = 0$ and $g(x) = x$ for $x \in \Gamma_\delta$. Right: A plot of $\mathcal{P}_\delta^s v|_{\Omega_\delta} \in \bigcap_{p\in[1,\infty]} N^{s,p,\delta}(\Omega)$ with $v(x) = 1 + 5\varphi(x + 4) - 2\varphi(x - 4)$ for a non-negative bump function $\varphi \in C_c^\infty(-1, 1)$. The parameters are the same as in Figure 4.2.

4.4.1 Connection between classical and nonlocal Sobolev spaces

As we know from Section 4.2.3, the translation operators $\mathcal{Q}_\delta^s$ and its inverse $\mathcal{P}_\delta^s$ provide a isomorphism between the spaces $H^{s,p,\delta}(\mathbb{R}^n)$ and $W^{1,p}(\mathbb{R}^n)$ with the properties $\nabla \circ \mathcal{Q}_\delta^s = D_\delta^s$ and

$D^s_\delta \circ \mathcal{P}^s_\delta = \nabla$. On a bounded open set $\Omega \subset \mathbb{R}^n$, it still holds for $u \in H^{s,p,\delta}(\Omega)$ that $\mathcal{Q}^s_\delta u$ lies in $W^{1,p}(\Omega)$ with $\nabla(\mathcal{Q}^s_\delta u) = D^s_\delta u$ (cf. Lemma 4.2.4), but $\mathcal{P}^s_\delta$ is not defined on $W^{1,p}(\Omega)$, which prevents an identification with $H^{s,p,\delta}(\Omega)$ in analogy to the setting on the whole space $\mathbb{R}^n$.

It turns out that one can resolve this issue and find a perfect translation mechanism also between the classical and nonlocal Sobolev spaces on bounded sets, by considering the spaces modulo the functions with zero (nonlocal) gradient. In this spirit, our next theorem gives a natural generalization of Lemma 4.2.4 and 4.2.5, cf. also [72, Section 2.4].

To state the result precisely, let us introduce the quotient spaces

$$\widetilde{H}^{s,p,\delta}(\Omega) := H^{s,p,\delta}(\Omega)/N^{s,p,\delta}(\Omega) \quad \text{and} \quad \widetilde{W}^{1,p}(\Omega) := W^{1,p}(\Omega)/\mathcal{C}(\Omega),$$

where $\mathcal{C}(\Omega) := \{v \in L^p(\Omega) : v \text{ is constant}\}$; for the equivalence classes in $\widetilde{H}^{s,p,\delta}(\Omega)$, we write $[u]^s_\delta = u + N^{s,p,\delta}(\Omega)$ with a representative in $u \in H^{s,p,\delta}(\Omega)$, and analogously, $[v] = v + \mathcal{C}(\Omega)$ with $v \in W^{1,p}(\Omega)$ for elements in $\widetilde{W}^{1,p}(\Omega)$. We endow these spaces with the norms given by

$$\big\|[u]^s_\delta\big\|_{\widetilde{H}^{s,p,\delta}(\Omega)} := \|D^s_\delta u\|_{L^p(\Omega;\mathbb{R}^n)} \quad \text{and} \quad \big\|[v]\big\|_{\widetilde{W}^{1,p}(\Omega)} := \|\nabla v\|_{L^p(\Omega;\mathbb{R}^n)}, \tag{4.41}$$

noting that $D^s_\delta u$ and ∇v are both independent of the chosen representative of $[u]^s_\delta$ and $[v]$, respectively. Moreover, let $\widetilde{D}^s_\delta[u]^s_\delta := D^s_\delta u$ and $\widetilde{\nabla}[v] := \nabla v$ for $u \in H^{s,p,\delta}(\Omega)$ and $v \in W^{1,p}(\Omega)$, respectively, where the choice of representative is irrelevant.

Theorem 4.4.1 (Isomorphism between $\widetilde{H}^{s,p,\delta}(\Omega)$ and $\widetilde{W}^{1,p}(\Omega)$). *Let $p \in [1,\infty]$. Then, the linear map*

$$\widetilde{\mathcal{Q}}^s_\delta : \widetilde{H}^{s,p,\delta}(\Omega) \to \widetilde{W}^{1,p}(\Omega), \quad [u]^s_\delta \mapsto [\mathcal{Q}^s_\delta u]$$

defines a isometric isomorphism, and it holds with $\widetilde{\mathcal{P}}^s_\delta := (\widetilde{\mathcal{Q}}^s_\delta)^{-1}$ that

$$\widetilde{\nabla} \circ \widetilde{\mathcal{Q}}^s_\delta = \widetilde{D}^s_\delta \quad \textit{and} \quad \widetilde{D}^s_\delta \circ \widetilde{\mathcal{P}}^s_\delta = \widetilde{\nabla}. \tag{4.42}$$

Proof. Note first that $\widetilde{\mathcal{Q}}^s_\delta$ is well-defined since $\mathcal{Q}^s_\delta h$ is constant for any $h \in N^{s,p,\delta}(\Omega)$. The first identity in (4.42) follows immediately from $\nabla \circ \mathcal{Q}^s_\delta = D^s_\delta$, and we can compute that

$$\|\widetilde{\mathcal{Q}}^s_\delta[u]^s_\delta\|_{\widetilde{W}^{1,p}(\Omega)} = \|\nabla(\mathcal{Q}^s_\delta u)\|_{L^p(\Omega;\mathbb{R}^n)} = \|D^s_\delta u\|_{L^p(\Omega;\mathbb{R}^n)} = \|[u]^s_\delta\|_{\widetilde{H}^{s,p,\delta}(\Omega)}$$

for all $u \in H^{s,p,\delta}(\Omega)$, which shows that $\widetilde{\mathcal{Q}}^s_\delta$ is an isometry. To prove the bijectivity, we claim that the inverse of $\widetilde{\mathcal{Q}}^s_\delta$ is given by

$$\widetilde{\mathcal{P}}^s_\delta[v] = \big[\mathcal{P}^s_\delta(\mathcal{E}v)|_{\Omega_\delta}\big]^s_\delta \quad \text{for } v \in W^{1,p}(\Omega),$$

where $\mathcal{E} : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ is any bounded linear extension operator. Indeed, it holds that

$$D^s_\delta(\widetilde{\mathcal{P}}^s_\delta[v]) = D^s_\delta(\mathcal{P}^s_\delta(\mathcal{E}v)|_{\Omega_\delta}) = \nabla(\mathcal{E}v)|_\Omega = \nabla v \quad \text{for } v \in W^{1,p}(\Omega),$$

from which we infer the second part of (4.42), as well as $\widetilde{\mathcal{P}}^s_\delta \circ \widetilde{\mathcal{Q}}^s_\delta = \mathrm{Id}$ and $\widetilde{\mathcal{Q}}^s_\delta \circ \widetilde{\mathcal{P}}^s_\delta = \mathrm{Id}$.

□

Remark 4.4.2. The boundedness of $\widetilde{\mathcal{Q}}^s_\delta$ and $\widetilde{\mathcal{P}}^s_\delta$ holds as well, if $\widetilde{H}^{s,p,\delta}(\Omega)$ is equipped with the associated quotient norm, i.e.,

$$|||[u]^s_\delta|||_{\widetilde{H}^{s,p,\delta}(\Omega)} := \inf_{h\in N^{s,p,\delta}(\Omega)} \|u-h\|_{L^p(\Omega_\delta)} + \|D^s_\delta u\|_{L^p(\Omega;\mathbb{R}^n)}$$

for $u \in H^{s,p,\delta}(\Omega)$. Indeed, for $\widetilde{\mathcal{Q}}^s_\delta$ this is clear, whereas for $\widetilde{\mathcal{P}}^s_\delta$ we can compute for $v \in W^{1,p}(\Omega)$ with $\int_\Omega v\,dx = 0$ that

$$\begin{aligned}\left\|\left|\widetilde{\mathcal{P}}^s_\delta[v]\right|\right\|_{\widetilde{H}^{s,p,\delta}(\Omega)} &\leq \|\mathcal{P}^s_\delta(\mathcal{E}v)\|_{L^p(\Omega_\delta)} + \|\nabla v\|_{L^p(\Omega;\mathbb{R}^n)} \\ &\leq C\|\mathcal{E}v\|_{W^{1,p}(\mathbb{R}^n)} + \|\nabla v\|_{L^p(\Omega;\mathbb{R}^n)} \\ &\leq C\|v\|_{W^{1,p}(\Omega)} + \|\nabla v\|_{L^p(\Omega;\mathbb{R}^n)} \leq C\|[v]\|_{\widetilde{W}^{1,p}(\Omega)},\end{aligned}$$

where the second inequality uses Lemma 4.2.5, and the last the classical Poincaré-Wirtinger inequality. Moreover, for $p \in (1,\infty)$, the operator norm of $\widetilde{\mathcal{P}}^s_\delta$ is independent of s by (4.20). We use this observation later in Corollary 4.4.7 to deduce a new nonlocal Poincaré-Wirtinger equation.

In the classical Sobolev setting, the norm $\|\cdot\|_{\widetilde{W}^{1,p}(\Omega)}$ in (4.41) is equivalent to the quotient norm on $\widetilde{W}^{1,p}(\Omega)$ by the standard Poincaré-Wirtinger inequality. △

If the characterization of $N^{s,p,\delta}(\Omega)$ in Theorem 4.3.8 holds, then any boundary values can be attained in the layer Γ_δ by elements in an equivalence class of $\widetilde{H}^{s,p,\delta}(\Omega)$. In other words, for each $u \in H^{s,p,\delta}(\Omega)$ and each $g \in L^p(\Gamma_\delta)$, there exists a representative of $[u]^s_\delta = u + N^{s,p,\delta}(\Omega)$ that coincides with g in Γ_δ. Based on this observation, we can state the following consequence of Theorem 4.4.1.

Corollary 4.4.3. *Let $p \in (1, \frac{2}{1-s})$ and Ω be a bounded $C^{1,1}$-domain. Then, for every $v \in W^{1,p}(\Omega)$ and $g \in L^p(\Gamma_\delta)$, there is a $u \in H^{s,p,\delta}(\Omega)$ such that*

$$\begin{cases} D^s_\delta u = \nabla v & \text{a.e. in } \Omega, \\ u = g & \text{a.e. in } \Gamma_\delta. \end{cases}$$

Proof. Let $v \in W^{1,p}(\Omega)$. Then, Theorem 4.4.1 implies that $\nabla v = \widetilde{\nabla}[v] = \widetilde{D}^s_\delta[u]^s_\delta = D^s_\delta u$ for some $u \in H^{s,p,\delta}(\Omega)$. By Theorem 4.3.8, we may assume that u coincides with g in the boundary layer Γ_δ, which yields the desired function. □

4.4.2 Extension modulo functions with zero nonlocal gradient.

While not every function in $H^{s,p,\delta}(\Omega)$ is the restriction of a function in $H^{s,p,\delta}(\mathbb{R}^n)$ (cf. Proposition 4.3.3), we can show nevertheless that extensions to the whole space $\mathbb{R}^n$ are possible up to function with zero nonlocal gradient. This technical tool has several applications within this paper. It has appeared already in the proof of Proposition 4.3.3, where it provided an efficient way for generating functions in $N^{s,p,\delta}(\Omega)$.

With $p \in [1,\infty]$, we define for a given bounded linear extension operator $\mathcal{E} : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$,

$$\mathcal{E}^s_\delta : H^{s,p,\delta}(\Omega) \to H^{s,p,\delta}(\mathbb{R}^n), \quad \mathcal{E}^s_\delta := \mathcal{P}^s_\delta \circ \mathcal{E} \circ \mathcal{Q}^s_\delta, \tag{4.43}$$

with the translation operators $\mathcal{Q}^s_\delta$ and $\mathcal{P}^s_\delta$ from Section 4.2.3. As the composition of bounded linear operators, $\mathcal{E}^s_\delta$ is bounded, even uniformly in s when $p \in (1,\infty)$, see (4.20). In view of (4.15), we infer for every $u \in H^{s,p,\delta}(\Omega)$ that $D^s_\delta \mathcal{E}^s_\delta u = D^s_\delta u$ on Ω, and thus,

$$u - \mathcal{E}^s_\delta u|_{\Omega_\delta} \in N^{s,p,\delta}(\Omega) \quad \text{for } u \in H^{s,p,\delta}(\Omega).$$

In this sense, $\mathcal{E}^s_\delta$ can be viewed as an extension operator on $H^{s,p,\delta}(\Omega)$ modulo functions in $N^{s,p,\delta}(\Omega)$.

Note further that $\mathcal{E}^s_\delta$, as a map from $H^{s,p,\delta}(\Omega)$ to $L^p(\Omega_\delta)$, is compact for $p \in (1,\infty)$ due to the compact embedding of $H^{s,p,\delta}(\mathbb{R}^n) = H^{s,p}(\mathbb{R}^n)$ into $L^p(\Omega_\delta)$, see Section 4.2.2. Thus, if $(u_j)_j$ is a weakly convergent sequence in $H^{s,p,\delta}(\Omega)$ with limit $u \in H^{s,p,\delta}(\Omega)$, then

$$\mathcal{E}^s_\delta u_j \to \mathcal{E}^s_\delta u \quad \text{in } L^p(\Omega_\delta). \tag{4.44}$$

4.4.3 A new nonlocal Poincaré inequality

Another application of Theorem 4.3.8 is that we can derive a new Poincaré inequality for the nonlocal gradient. As opposed to the Poincaré inequality in [30, Theorem 6.1], which requires functions to be zero in the double collar $\Gamma_{\pm\delta}$, the new one only imposes a condition in Γ_δ together with an average-value condition. Precisely, will work with functions in the linear subspace

$$\mathring{H}^{s,p,\delta}(\Omega) := \big\{u \in H^{s,p,\delta}(\Omega) : u = 0 \text{ a.e. in } \Gamma_\delta,\ \textstyle\int_\Omega u\,dx = 0\big\}.$$

Theorem 4.4.4 (Nonlocal Poincaré inequality). *Let $p \in (1, \frac{2}{1-s})$ and $\Omega \subset \mathbb{R}^n$ be a bounded $C^{1,1}$-domain. Then, there exists a constant $C > 0$ such that*

$$\|u\|_{L^p(\Omega_\delta)} \le C\|D^s_\delta u\|_{L^p(\Omega;\mathbb{R}^n)}$$

for all $u \in \mathring{H}^{s,p,\delta}(\Omega)$.

Proof. The proof stragegy follows a well-known contradiction argument, with Lemma 4.4.5 below as main technical ingredient. Suppose the statement is false, then there is a sequence $(u_j)_j \subset \mathring{H}^{s,p,\delta}(\Omega)$ with $\|u_j\|_{L^p(\Omega_\delta)} > j\|D^s_\delta u_j\|_{L^p(\Omega;\mathbb{R}^n)}$ for all j. By defining the sequence $(\tilde{u}_j)_j \subset \mathring{H}^{s,p,\delta}(\Omega)$ via

$$\tilde{u}_j := \frac{u_j}{\|u_j\|_{L^p(\Omega_\delta)}},$$

we obtain $\|\tilde{u}_j\|_{L^p(\Omega_\delta)} = 1$ and $\|D^s_\delta \tilde{u}_j\|_{L^p(\Omega;\mathbb{R}^n)} \le 1/j$ for each j. This allows us to conclude for a non-relabeled subsequence that

$$\tilde{u}_j \rightharpoonup h \quad \text{in } H^{s,p,\delta}(\Omega) \text{ as } j \to \infty,$$

with a limit function $h \in H^{s,p,\delta}(\Omega)$ that satisfies $D^s_\delta h = 0$, or in other words, $h \in N^{s,p,\delta}(\Omega)$. Due to the weak closedness of $\mathring{H}^{s,p,\delta}(\Omega)$, we also find that $u \in \mathring{H}^{s,p,\delta}(\Omega)$, which yields $h = 0$ by Proposition 4.3.9.

Finally, we infer from Lemma 4.4.5 that $\tilde{u}_j \to 0$ in $L^p(\Omega_\delta)$ as $j \to \infty$, which contradicts $\|\tilde{u}_j\|_{L^p(\Omega_\delta)} = 1$ for all j and thereby, proves the result. □

The previous proof used the compact embedding of $\mathring{H}^{s,p,\delta}(\Omega)$ into $L^p(\Omega_\delta)$, which is the subject of the following lemma. We point out that it builds substantially on the identification of $N^{s,p,\delta}(\Omega)$ from Theorem 4.3.8.

Lemma 4.4.5. *Let $p \in (1, \frac{2}{1-s})$ and suppose $\Omega \subset \mathbb{R}^n$ is a bounded $C^{1,1}$-domain. If $(u_j)_j \subset \mathring{H}^{s,p,\delta}(\Omega)$ is such that $u_j \rightharpoonup u$ in $H^{s,p,\delta}(\Omega)$ as $j \to \infty$ with some $u \in H^{s,p,\delta}(\Omega)$, then $u \in \mathring{H}^{s,p,\delta}(\Omega)$ and*

$$u_j \to u \quad \text{in } L^p(\Omega_\delta) \text{ as } j \to \infty. \tag{4.45}$$

Proof. The fact that $u \in \mathring{H}^{s,p,\delta}(\Omega)$ is clear, since $\mathring{H}^{s,p,\delta}(\Omega)$ is weakly closed in $L^p(\Omega_\delta)$. As for (4.45), we use the extension operator $\mathcal{E}^s_\delta$ from Section 4.4.2 to obtain

$$\mathcal{E}^s_\delta u_j \to \mathcal{E}^s_\delta u \quad \text{in } L^p(\Omega_\delta) \text{ as } j \to \infty$$

by (4.44). Therefore, with the sequence $(h_j)_j \subset N^{s,p,\delta}(\Omega)$ given by $h_j := u_j - \mathcal{E}^s_\delta u_j$ and $h := u - \mathcal{E}^s_\delta u$ it holds that

$$h_j \rightharpoonup h \text{ in } L^p(\Omega_\delta) \quad \text{and} \quad h_j \to h \text{ in } L^p(\Gamma_\delta) \text{ as } j \to \infty, \tag{4.46}$$

where the second convergence follows from $u_j = 0 = u$ a.e. on Γ_δ. If we consider the norm on $N^{s,p,\delta}(\Omega)$ from Remark 4.3.11, then (4.46) implies

$$|||h_j|||_{N^{s,p,\delta}(\Omega)} = \|h_j\|_{L^p(\Omega_j)} + \left|\int_\Omega Q_\delta^s * h_j\right| \to 0 \quad \text{as } j \to \infty.$$

Since $|||\cdot|||_{N^{s,p,\delta}(\Omega)}$ is equivalent to the norm induced on $N^{s,p,\delta}(\Omega)$ by $\|\cdot\|_{L^p(\Omega_\delta)}$, we obtain $h_j \to h$ in $L^p(\Omega_\delta)$, and thus,

$$u_j = \mathcal{E}_\delta^s u_j + h_j \to \mathcal{E}_\delta^s u + h = u \quad \text{in } L^p(\Omega_\delta) \text{ as } j \to \infty,$$

which concludes the proof. □

Remark 4.4.6. The contradiction argument in Theorem 4.4.4 works more generally for any weakly closed subset $X \subset H^{s,p,\delta}(\Omega)$ that is compactly contained in $L^p(\Omega_\delta)$ and satisfies $X \cap N^{s,p,\delta}(\Omega) = \{0\}$. For example, one could replace the condition $\int_\Omega u\,dx = 0$ by the condition $\int_\Omega Q_\delta^s * u\,dx = 0$ or remove the mean-value condition completely and assume $u = 0$ a.e. in $\Omega_\delta \setminus \overline{O}$ for any $O \Subset \Omega$ (cf. Remark 4.3.10). △

4.4.4 Nonlocal Poincaré-Wirtinger inequality

Here, we derive an inequality involving the nonlocal gradient in the spirit of the classical Poincaré-Wirtinger inequality, by subtracting suitable functions with zero nonlocal gradient. Moreover, we complement the inequality with a compactness result. This will be used later in Section 4.6 to prove the well-posedness and localization as $s \uparrow 1$ of nonlocal variational problems with Neumann-type boundary conditions.

Let $p \in (1,\infty)$, and consider the metric projection $\pi_\delta^s : L^p(\Omega_\delta) \to N^{s,p,\delta}(\Omega)$, which minimizes the distance to the functions with vanishing nonlocal gradient in the L^p-norm, i.e., for $u \in L^p(\Omega_\delta)$,

$$\|u - \pi_\delta^s(u)\|_{L^p(\Omega_\delta)} = \min_{h \in N^{s,p,\delta}(\Omega)} \|u - h\|_{L^p(\Omega_\delta)};$$

Note that the minimum exists, considering that $N^{s,p,\delta}(\Omega)$ is weakly closed in $H^{s,p,\delta}(\Omega)$, and also in $L^p(\Omega)$, since $\|\cdot\|_{H^{s,p,\delta}(\Omega)} = \|\cdot\|_{L^p(\Omega)}$ on $N^{s,p,\delta}(\Omega)$. In the case $p = 2$, π_δ^s corresponds to the (linear) orthogonal projection onto $N^{s,p,\delta}(\Omega)$. Even though π_δ^s need not be linear when $p \neq 2$, one does have that π_δ^s is 1-homogeneous and that

$$\pi_\delta^s(u + h) = \pi_\delta^s(u) + h \qquad \text{for all } h \in N^{s,p,\delta}(\Omega). \tag{4.47}$$

It is also well-known that π_δ^s is continuous, given that $L^p(\Omega_\delta)$ is uniformly convex, see e.g., [119].

We now formulate and prove the Poincaré-Wirtinger inequality with the help of the metric projection.

Lemma 4.4.7 (Nonlocal Poincaré-Wirtinger inequality). *Let $p \in (1,\infty)$. Then, there exists a constant $C = C(\Omega, p, \delta) > 0$ such that*

$$\|u - \pi_\delta^s(u)\|_{L^p(\Omega_\delta)} \leq C\|D_\delta^s u\|_{L^p(\Omega;\mathbb{R}^n)}$$

for all $u \in H^{s,p,\delta}(\Omega)$.

Proof. It follows from Theorem 4.4.1 (in the version where $\widetilde{H}^{s,p,\delta}(\Omega)$ is equipped with the quotient norm, see Remark 4.4.2) and Lemma 4.2.4 that

$$\begin{aligned}\|u - \pi_\delta^s(u)\|_{L^p(\Omega_\delta)} &\leq |||[u]_\delta^s|||_{\widetilde{H}^{s,p,\delta}(\Omega)} = \left|\left|\left|\widetilde{\mathcal{P}}_\delta^s \widetilde{\mathcal{Q}}_\delta^s [u]_\delta^s\right|\right|\right|_{\widetilde{H}^{s,p,\delta}(\Omega)} \leq C\|\widetilde{\mathcal{Q}}_\delta^s [u]_\delta^s\|_{\widetilde{W}^{1,p}(\Omega)} \\ &= C\|[\mathcal{Q}_\delta^s u]\|_{\widetilde{W}^{1,p}(\Omega)} = C\|\nabla(\mathcal{Q}_\delta^s u)\|_{L^p(\Omega;\mathbb{R}^n)} = C\|D_\delta^s u\|_{L^p(\Omega;\mathbb{R}^n)},\end{aligned}$$

with a constant $C > 0$ independent of s. □

Second, one obtains the following compactness result. It can be seen as the trace-free analogue to [28, Theorem 2.3] in the setting of complementary-value spaces.

Lemma 4.4.8 (Compactness in $H^{s,p,\delta}(\Omega)$). *Let $p \in (1,\infty)$, then any sequence $(u_j)_j \subset H^{s,p,\delta}(\Omega)$ converging weakly to u in $H^{s,p,\delta}(\Omega)$ satisfies*

$$u_j - \pi^s_\delta(u_j) \to u - \pi^s_\delta(u) \quad \text{in } L^p(\Omega_\delta) \text{ as } j \to \infty.$$

Proof. Using the extension operator modulo $N^{s,p,\delta}(\Omega)$ from (4.43), we define

$$h_j := \mathcal{E}^s_\delta u_j - u_j + \mathcal{E}^s_\delta u - u \in N^{s,p,\delta}(\Omega) \quad \text{for all } j.$$

Since $\mathcal{E}^s_\delta u_j \to \mathcal{E}^s_\delta u$ in $L^p(\Omega_\delta)$ according to (4.44), it follows that $u_j + h_j \to u$ in $L^p(\Omega_\delta)$, and hence,

$$\lim_{j\to\infty} \|u_j - u - \pi^s_\delta(u_j - u)\|_{L^p(\Omega_\delta)} \leq \lim_{j\to\infty} \|u_j - u + h_j\|_{L^p(\Omega_\delta)} = 0,$$

by definition of the metric projection. This shows that

$$u_j - \pi^s_\delta(u_j - u) \to u \text{ in } L^p(\Omega_\delta) \text{ as } j \to \infty.$$

In view of (4.47) and the continuity of π^s_δ, we then find that

$$u_j - \pi^s_\delta(u_j) = u_j - \pi^s_\delta(u_j - u) - \pi^s_\delta(u_j - \pi^s_\delta(u_j - u)) \to u - \pi^s_\delta(u) \quad \text{in } L^p(\Omega_\delta),$$

which concludes the proof. □

4.5 Nonlocal differential inclusion problems

In the present section we discuss results on the solvability of differential inclusion problems involving the nonlocal gradient. This means that for a given set $E \subset \mathbb{R}^{m\times n}$, we aim to find all $u \in H^{s,p,\delta}(\Omega;\mathbb{R}^m)$ that satisfy

$$D^s_\delta u \in E \quad \text{a.e. in } \Omega, \tag{4.48}$$

and optionally, also a boundary condition in the single layer Γ_δ or the double layer $\Gamma_{\pm\delta}$. Problems of the type (4.48) have not appeared in the literature before, although related results such as fractional Korn inequalities have been studied recently in various settings [32, 130, 190].

Throughout this section, let $p \in [1,\infty]$ and $\Omega \subset \mathbb{R}^n$ be a bounded $C^{1,1}$-domain. Additionally, whenever we work with Dirichlet conditions in the double layer $\Gamma_{\pm\delta}$, we also assume that $\Omega_{-\delta} \neq \emptyset$ and $|\partial\Omega_{-\delta}| = 0$. The set $N^{s,p,\delta}(\Omega)$ naturally plays a key role in the discussion of (4.48), considering that it can be interpreted as the solution to the most basic nonlocal inclusion, namely, with the choice $E = \{0\}$. On the one hand, for any solution to (4.48), adding a function from $N^{s,p,\delta}(\Omega)$ generates a new solution, that is, if $u \in H^{s,p,\delta}(\Omega)$ solves (4.48), then so does any other element in $[u]^s_\delta = u + N^{s,p,\delta}(\Omega)$, cf. Section 4.4.1. When $p \in (1, \frac{2}{1-s})$, any single-layer boundary condition $g \in L^p(\Gamma_\delta)$ can therefore be attained, by the characterization of $N^{s,p,\delta}(\Omega)$ in Theorem 4.3.8.

Our overall strategy in dealing with (4.48) is to relate them with classical differential inclusions, and to carry over the by now well-known results on their classical counterparts, that is, solving

$$\nabla v \in E \quad \text{a.e. in } \Omega. \tag{4.49}$$

for $v \in W^{1,p}(\Omega)$, also subject to boundary conditions. A rich literature on the latter has emerged over the last decades, including [76–78, 170, 200], see also [75, 169, 182] for an overview. While there is no unified theory available, the results fall roughly into two groups, relating to the complementary themes of rigidity and flexibilty. This division, which we will adopt here as well, is partly motivated by models in materials science, where differential inclusions appear naturally when studying microstructure formation, cf. [169, 182].

The connection between nonlocal and standard gradients established in Section 4.4.1 implies that (4.48) and (4.49) are equivalent when it comes to solvability. Indeed, due to Theorem 4.4.1 the map $\widetilde{\mathcal{Q}}_\delta^s$ gives a bijection between the solutions of (4.49) modulo constants and the solutions to (4.48) modulo functions in $N^{s,p,\delta}(\Omega)$. In the following, we take a look into selected aspects of flexibility and rigidity in the nonlocal setting, starting with the latter.

One calls the classical differential inclusion (4.49) rigid, if all its solutions v have constant gradient, meaning that, $v(x) = l_A(x) + c = Ax + c$ for $A \in E$ and $c \in \mathbb{R}^m$; recall the notation l_A with $A \in \mathbb{R}^{m\times n}$ for the linear function $l_A(x) = Ax$ with $x \in \mathbb{R}^n$. The nonlocal gradient of a linear function agrees with the classical gradient, since

$$D_\delta^s l_A = Q_\delta^s * \nabla l_A = Q_\delta^s * A = A, \tag{4.50}$$

where we have used $\|Q_\delta^s\|_{L^1(\mathbb{R}^n)} = 1$ (see also [32, Proposition 4.1]). Based on this observation, one obtains that rigidity carries over to the nonlocal setting in the following sense.

Corollary 4.5.1 (Nonlocal rigidity). *Let $E \subset \mathbb{R}^{m\times n}$ be such that the differential inclusion* (4.49) *is rigid. Then, all solutions $u \in H^{s,p,\delta}(\Omega)$ to the nonlocal inclusion*

$$D_\delta^s u \in E \quad \text{a.e. in } \Omega, \tag{4.51}$$

are of the form $u = l_A + h$ with $A \in E$ and $h \in N^{s,p,\delta}(\Omega)$. In particular, if $p \in (1, \frac{2}{1-s})$, then for any $g \in L^p(\Gamma_\delta)$, there is a solution u of (4.51) *with $u = g$ a.e. in Γ_δ.*

Proof. As $u \in [l_A]_\delta^s$ with $A \in E$ clearly solves (4.51) in view of (4.50), it remains to show that these are the only solutions. Indeed, by the assumption of rigidity, the solutions to (4.49) are exactly the functions that lie in $[l_A]$ for some $A \in E$, so that any u solving (4.51) needs to satisfy $\widetilde{\mathcal{Q}}_\delta^s [u]_\delta^s = [l_A]$. Since also $\widetilde{\mathcal{Q}}_\delta^s [l_A]_\delta^s = [l_A]$ and $\widetilde{\mathcal{Q}}_\delta^s$ is injective according to Theorem 4.4.1, we finally conclude that $u - l_A \in N^{s,p,\delta}(\Omega)$.

When the assumptions of Theorem 4.3.8 are satisfied, we may use Theorem 4.3.8 to find that any boundary condition is attained in $N^{s,p,\delta}(\Omega)$, which yields the second part of the statement. □

The preceding result characterizes all solutions in terms of the set $N^{s,p,\delta}(\Omega)$ and shows that there is no restriction on the boundary conditions that can be achieved in the single layer. If one prescribes boundary conditions in the double layer $\Gamma_{\pm\delta}$, instead, the set of solutions is considerably more restrictive. Our next statement addresses a nonlocal inclusion problem with linear boundary data l_A with $A \in \mathbb{R}^{m\times n}$, precisely,

$$\begin{cases} D_\delta^s u \in E & \text{a.e. in } \Omega_{-\delta}, \\ D_\delta^s u = A & \text{a.e. in } \Gamma_{-\delta}, \\ u = l_A & \text{a.e. in } \Gamma_{\pm\delta}, \end{cases} \tag{4.52}$$

for $u \in H^{s,p,\delta}(\Omega)$. Note that the reason for prescribing the nonlocal gradient in the collar $\Gamma_{-\delta}$ is that the condition $u = l_A$ a.e. in $\Gamma_{\pm\delta}$ automatically implies $D_\delta^s u = A$ near $\partial\Omega$ in light of (H2). The inclusion $D_\delta^s u \in E$ a.e. in Ω would therefore only be possible if $A \in E$, which renders the problem trivial. We now show a rigidity statement for (4.52).

Corollary 4.5.2 (Nonlocal rigidity with linear boundary conditions). *Let $E \subset \mathbb{R}^{m\times n}$ be such that the inclusion* (4.49) *is rigid and let $A \in \mathbb{R}^{m\times n}$. Then, the nonlocal inclusion problem* (4.52) *has a solution if and only if $A \in E$, which is then uniquely given by $u = l_A$.*

Proof. Let u be a solution of (4.52) and define $v := Q_\delta^s u \in W^{1,p}(\Omega)$, which then satisfies $\nabla v \in E$ a.e. in $\Omega_{-\delta}$ and $\nabla v = A$ a.e. in $\Gamma_{-\delta}$, cf. Lemma 4.2.4. Since (4.49) is rigid, there is an $A' \in E$ such that $\nabla v = A'$ a.e. in $\Omega_{-\delta}$. Hence, it holds that $v = l_{A'} + c$ a.e. in $\Omega_{-\delta}$ for some $c \in \mathbb{R}$. Moreover, for a.e. $x \in \Omega \setminus (\overline{\Omega_{-\delta}} + \mathrm{supp}(Q_\delta^s))$ (by (H2), this open set is non-empty), we obtain

$$v(x) = (Q_\delta^s * u)(x) = (Q_\delta^s * l_A)(x) = Ax.$$

Combining this with $\nabla v = A$ a.e. in $\Gamma_{-\delta}$, yields that $v = l_A$ a.e. in $\Gamma_{-\delta}$. We conclude that

$$\begin{cases} v = l_{A'} + c & \text{a.e. in } \Omega_{-\delta}, \\ v = l_A & \text{a.e. in } \Gamma_{-\delta}, \end{cases}$$

so that we must have $l_{A'} + c = l_A$ on $\partial\Omega_{-\delta}$ for v to be a Sobolev function. Unless $A = A'$ and $c = 0$, we find that the set where $l_{A'} + c = l_A$ is an affine subspace of dimension at most $n-1$, which cannot contain the boundary of the bounded open set $\Omega_{-\delta}$. Therefore, we must have $A = A'$ and $c = 0$, which yields, in particular, that $A \in E$ and $D_\delta^s u = \nabla v = A$ a.e. in Ω. We now infer from the nonlocal Poincaré inequality for double-layer boundary conditions (see [30, Theorem 6.1]) that $u = l_A$ is indeed the only solution. □

Next is a statement on flexibility for (4.52), which also allows for solutions with non-constant nonlocal gradients and reveals a relation between the attainable boundary conditions and the set E. In doing so, we restrict our attention to a weaker notion of solutions, though, calling a sequence $(u_j)_j \subset H^{s,\infty,\delta}(\Omega;\mathbb{R}^m)$ an approximate solution to (4.52), if

$$\begin{cases} \mathrm{dist}(D_\delta^s u_j, E) \to 0 & \text{in measure on } \Omega_{-\delta}, \\ D_\delta^s u_j \to A & \text{in measure on } \Gamma_{-\delta}, \\ u_j = l_A & \text{in } \Gamma_{\pm\delta}. \end{cases} \tag{4.53}$$

In the classical case, it is well-known that approximate solutions to (4.49) subject to linear boundary values l_A exist if and only if A lies in the quasiconvex hull of E defined by

$$E^{\mathrm{qc}} := \big\{B \in \mathbb{R}^{m\times n} \,:\, f(B) \le \sup_E f \text{ for all quasiconvex } f : \mathbb{R}^{m\times n} \to \mathbb{R}\big\},$$

see e.g., [169, Theorem 4.10],[75, Chapter 7]. For the approximate solutions as in (4.53), we can use the translation method to prove an analogous statement.

Proposition 4.5.3 (Approximate solutions to nonlocal differential inclusions). *Let $E \subset \mathbb{R}^{m\times n}$ be compact and $A \in \mathbb{R}^{m\times n}$. Then,* (4.52) *admits an approximate solution in the sense of* (4.53) *if and only if $A \in E^{\mathrm{qc}}$.*

Proof. First, suppose that $A \in E^{\mathrm{qc}}$, then by [169, Theorem 4.10], there is a bounded sequence $(v_j)_j \subset W_0^{1,\infty}(\Omega;\mathbb{R}^m)$ such that

$$\mathrm{dist}(A + \nabla v_j, E) \to 0 \quad \text{in measure on } \Omega. \tag{4.54}$$

We may assume without loss of generality that $v_j \to 0$ in $L^\infty(\Omega;\mathbb{R}^m)$ and hence, also $\nabla v_j \overset{*}{\rightharpoonup} 0$ in $L^\infty(\Omega;\mathbb{R}^{m\times n})$; otherwise, we glue together suitably scaled and translated copies of v_j for each j.

After identifying v_j with its extension to $\mathbb{R}^n$ by zero, we define the sequence $(\tilde{u}_j)_j$ by

$$\tilde{u}_j := \mathcal{P}^s_\delta v_j \in H^{s,\infty,\delta}(\mathbb{R}^n;\mathbb{R}^m) \quad \text{for } j \in \mathbb{N}.$$

Since $D^s_\delta \tilde{u}_j = \nabla v_j \overset{*}{\rightharpoonup} 0$ in $L^\infty(\mathbb{R}^n;\mathbb{R}^{m\times n})$, and the sequence $(v_j)_j$ is also bounded in $W^{1,p}(\mathbb{R}^n;\mathbb{R}^m)$, it follows along with the weak continuity of $\mathcal{P}^s_\delta$ that $\tilde{u}_j \rightharpoonup 0$ in $H^{s,p,\delta}(\mathbb{R}^n;\mathbb{R}^m) = H^{s,p}(\mathbb{R}^n;\mathbb{R}^m)$ as $j \to \infty$ for all $p \in (1,\infty)$. In addition, the compact embedding of $H^{s,p}(\mathbb{R}^n;\mathbb{R}^m)$ into $L^\infty(\Omega_\delta;\mathbb{R}^m)$ for $sp > n$ (see Section 4.2.2), yields

$$\tilde{u}_j \to 0 \quad \text{in } L^\infty(\Omega_\delta;\mathbb{R}^m) \quad \text{and} \quad \tilde{u}_j \overset{*}{\rightharpoonup} 0 \text{ in } H^{s,\infty,\delta}(\Omega;\mathbb{R}^m). \tag{4.55}$$

We now introduce a sequence of cut-off functions $(\chi_j)_j \subset C_c^\infty(\Omega_{-\delta};[0,1])$ such that

$$|\Omega_{-\delta} \setminus \{\chi_j = 1\}| \to 0 \quad \text{and} \quad \mathrm{Lip}(\chi_j)\|\tilde{u}_j\|_{L^\infty(\Omega_\delta;\mathbb{R}^m)} \to 0 \quad \text{as } j \to \infty, \tag{4.56}$$

where $\mathrm{Lip}(\chi_j)$ denotes the Lipschitz constant of χ_j, and we define $(u_j)_j$ via

$$u_j := \chi_j \tilde{u}_j \in H^{s,\infty,\delta}(\Omega;\mathbb{R}^m),$$

which guarantees

$$u_j = 0 \quad \text{in } \Gamma_{\pm\delta}. \tag{4.57}$$

Moreover, by the nonlocal Leibniz rule (see [72, Lemma 2]),

$$D^s_\delta u_j = \chi_j D^s_\delta \tilde{u}_j + K_{\chi_j}(\tilde{u}_j) = \chi_j \nabla v_j + K_{\chi_j}(\tilde{u}_j), \tag{4.58}$$

where $K_{\chi_j} : L^\infty(\Omega_\delta) \to L^\infty(\Omega;\mathbb{R}^n)$ are bounded linear operators that satisfy

$$\|K_{\chi_j}(\tilde{u}_j)\|_{L^\infty(\Omega;\mathbb{R}^{m\times n})} \leq C\mathrm{Lip}(\chi_j)\|\tilde{u}_j\|_{L^\infty(\Omega_\delta;\mathbb{R}^m)} \to 0 \text{ as } j \to \infty; \tag{4.59}$$

the last convergence follows from (4.56). Since $\mathrm{dist}(A + \chi_j \nabla v_j, E) \to 0$ in measure on $\Omega_{-\delta}$ due to (4.54) and the first convergence in (4.56), we conclude along with (4.58) and (4.59) that

$$\mathrm{dist}(A + D^s_\delta u_j, E) \to 0 \quad \text{in measure on } \Omega_{-\delta}.$$

Moreover, as [72, Lemma 3] yields convergence $D^s_\delta u_j \to 0$ in L^∞ in any compactly contained subset of the collar $\Gamma_{-\delta}$, we have that

$$A + D^s_\delta u_j \to A \quad \text{in measure on } \Gamma_{-\delta} \text{ as } j \to \infty.$$

Hence, we obtain the desired approximate solution to (4.52), after adding the linear function l_A to $(u_j)_j$.

Conversely, if $(u_j)_j \subset H^{s,\infty,\delta}(\Omega;\mathbb{R}^m)$ is a sequence satisfying (4.53), we set $v_j := \mathcal{Q}^s_\delta u_j$ for all $j \in \mathbb{N}$ to find that $(v_j)_j \subset W^{1,\infty}(\Omega;\mathbb{R}^m)$ is a sequence with $v_j = l_A$ on $\partial\Omega$ for all $j \in \mathbb{N}$ in the sense of traces, and

$$\begin{cases} \mathrm{dist}(\nabla v_j, E) = \mathrm{dist}(D^s_\delta u_j, E) \to 0 & \text{in measure on } \Omega_{-\delta} \text{ as } j \to \infty, \\ \nabla v_j = D^s_\delta u_j \to A & \text{in measure on } \Gamma_{-\delta} \text{ as } j \to \infty. \end{cases}$$

A small adaptation to the argument in [169, Theorem 4.10 (i)] now shows that $A \in E^{\mathrm{qc}}$, as desired. □

4.6 Well-posedness and localization of nonlocal Neumann-type problems

This section is concerned with the analysis of nonlocal differential equations with homogeneous Neumann-type boundary conditions. In fact, it even covers a more general setting with natural boundary conditions. Our main results are the well-posedness for these problems for any fixed fractional parameter $s \in (0,1)$ and a rigorous proof of localisation, i.e., the convergence to the classical analogues of these boundary-value problems as the fractional parameter s goes to 1.

We approach these problems from the variational perspective, where the objects of interest are the associated energy functionals: For $\Omega \subset \mathbb{R}^n$ a bounded Lipschitz domain and $p \in (1,\infty)$, consider $\mathcal{F}^s_\delta : H^{s,p,\delta}(\Omega;\mathbb{R}^m) \to \mathbb{R}_\infty$ given by

$$\mathcal{F}^s_\delta(u) = \int_\Omega f(x, D^s_\delta u)\,dx - \int_{\Omega_\delta} F \cdot u\,dx, \tag{4.60}$$

where $F \in L^{p'}(\Omega_\delta;\mathbb{R}^m)$ with p' the dual exponent of p and the Carathéodory function $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}_\infty$ are suitably given.

Due to the absence of any constraints in the space of admissible functions $H^{s,p,\delta}(\Omega)$, the minimization of $\mathcal{F}^s_\delta$ gives rise to natural boundary conditions when passing to the Euler-Lagrange equations. Nonlocal variational problems on complementary-value spaces, in contrast, lead to Dirichlet boundary-value problems, see e.g., [30, Section 8].

4.6.1 Existence theory for a class of nonlocal Neumann-type variational problems

In this section we prove the existence of minimizers of the functional in (4.60), on a suitable subspace of $H^{s,p,\delta}(\Omega)$ where the Poincaré-Wirtinger inequality from Section 4.4.4 can be applied. Precisely, recalling the metric projection $\pi^s_\delta : L^p(\Omega_\delta;\mathbb{R}^m) \to N^{s,p,\delta}(\Omega;\mathbb{R}^m)$ from Section 4.4.4 (extended to vector-valued functions), we introduce the sets

$$N^{s,p,\delta}(\Omega;\mathbb{R}^m)^\perp = \{u \in H^{s,p,\delta}(\Omega;\mathbb{R}^m) \,:\, \pi^s_\delta(u) = 0\}.$$

For $p = 2$, this corresponds to the orthogonal complement of $N^{s,2,\delta}(\Omega;\mathbb{R}^m)$ in L^2, whereas for the case $p \neq 2$, it need not be a linear subspace, given the nonlinearity of the metric projection.

We now present the main result of this section, which establishes the existence of minimizers for $\mathcal{F}^s_\delta$ on the subspaces $N^{s,p,\delta}(\Omega;\mathbb{R}^m)^\perp$.

Theorem 4.6.1 (Existence of minimizers for $\mathcal{F}^s_\delta$). *Let $p \in (1,\infty)$, $F \in L^{p'}(\Omega_\delta;\mathbb{R}^m)$ and $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$ be a Carathéodory integrand such that*

$$f(x,A) \geq c\left(|A|^p - 1\right) \quad \text{for a.e. } x \in \Omega \text{ and all } A \in \mathbb{R}^{m\times n}$$

with a constant $c > 0$. If $v \mapsto \int_\Omega f(x,\nabla v)\,dx$ is weakly lower semicontinuous on $W^{1,p}(\Omega;\mathbb{R}^m)$, then the functional $\mathcal{F}^s_\delta$ in (4.60), i.e.,

$$\mathcal{F}^s_\delta(u) = \int_\Omega f(x, D^s_\delta u)\,dx - \int_{\Omega_\delta} F \cdot u\,dx,$$

admits a minimizer over $N^{s,p,\delta}(\Omega;\mathbb{R}^m)^\perp$.

Proof. We apply the direct method in the calculus of variations. Note first that $N^{s,p,\delta}(\Omega;\mathbb{R}^m)^{\perp}$ is a weakly closed subset of $H^{s,p,\delta}(\Omega;\mathbb{R}^m)$ as a consequence of Lemma 4.4.8. The coercivity then follows from the lower bound on f along with the Poincaré-Wirtinger inequality from Corollary 4.4.7, which reduces to

$$\|u\|_{L^p(\Omega_\delta;\mathbb{R}^m)} \leq C\|D^s_\delta u\|_{L^p(\Omega;\mathbb{R}^{m\times n})},$$

for $u \in N^{s,p,\delta}(\Omega;\mathbb{R}^m)^{\perp}$. For the weak lower semicontinuity of $\mathcal{F}^s_\delta$, we observe that if $u_j \rightharpoonup u$ in $H^{s,p,\delta}(\Omega;\mathbb{R}^m)$, then $v_j := \mathcal{Q}^s_\delta u_j \rightharpoonup \mathcal{Q}^s_\delta u =: v$ in $W^{1,p}(\Omega;\mathbb{R}^m)$ with $\nabla v_j = D^s_\delta u_j$ for all j and $\nabla v = D^s_\delta u$, cf. Lemma 4.2.4. Hence,

$$\begin{aligned}\mathcal{F}^s_\delta(u) &= \int_\Omega f(x,\nabla v)\,dx - \int_{\Omega_\delta} F\cdot u\,dx\\ &\leq \liminf_{j\to\infty}\int_\Omega f(x,\nabla v_j)\,dx - \int_{\Omega_\delta} F\cdot u_j\,dx = \liminf_{j\to\infty}\mathcal{F}^s_\delta(u_j),\end{aligned}$$

showing that $\mathcal{F}^s_\delta$ is weakly lower semicontinuous on $H^{s,p,\delta}(\Omega;\mathbb{R}^m)$. In combination with the coercivity, this yields the desired existence of a minimizer of $\mathcal{F}^s_\delta$ in $N^{s,p,\delta}(\Omega;\mathbb{R}^m)^{\perp}$. □

Remark 4.6.2. a) For the sake of generality, the previous theorem assumes that the classical integral functional (with standard gradients) associated to $\mathcal{F}^s_\delta$ is weakly lower semicontinuous. Well-known sufficient conditions for this include polyconvexity of the integrand f in the second argument or quasiconvexity of the latter along with a suitable upper bound, see e.g., [75, Theorems 8.11 and 8.31].

b) Note that if $F \in L^{p'}(\Omega_\delta;\mathbb{R}^m)$ satisfies the compatibility condition

$$\int_{\Omega_\delta} F\cdot h\,dx = 0 \quad \text{for all } h \in N^{s,p,\delta}(\Omega;\mathbb{R}^m), \tag{4.61}$$

then $\mathcal{F}^s_\delta$ is invariant under translations in $N^{s,p,\delta}(\Omega;\mathbb{R}^m)$. As a consequence of Theorem 4.6.1, $\mathcal{F}^s_\delta$ then admits minimizers over the whole space $H^{s,p,\delta}(\Omega;\mathbb{R}^m)$. △

As a consequence of Theorem 4.6.2, and specifically Remark 4.6.2, one can infer, by passing to Euler-Lagrange equations, the existence of weak solutions for a class of nonlocal differential equations with natural boundary conditions. Namely, suppose that F satisfies (4.61) and let f be continuously differentiable in its second argument and $C > 0$ such that

$$|f(x,A)| \leq C(1+|A|^p) \quad\text{and}\quad |D_Af(x,A)| \leq C(1+|A|^{p-1}) \quad \text{for all } (x,A)\in\Omega_\delta\times\mathbb{R}^{m\times n}, \tag{4.62}$$

with D_Af the differential of f with respect to its second argument. Then, using a standard argument, see [75, Theorem 3.37], we find that the minimizers $u \in H^{s,p,\delta}(\Omega;\mathbb{R}^m)$ of $\mathcal{F}^s_\delta$ solve the weak Euler-Lagrange equation

$$\int_\Omega D_Af(x,D^s_\delta u)\cdot D^s_\delta v\,dx = \int_{\Omega_\delta} F\cdot v\,dx \quad \text{for all } v\in H^{s,p,\delta}(\Omega;\mathbb{R}^m). \tag{4.63}$$

We note that the compatibility condition in (4.61) is also necessary for (4.63) to hold, since the left hand-side is zero for $v \in N^{s,p,\delta}(\Omega;\mathbb{R}^m)$. Moreover, by the definition of the weak nonlocal divergence via nonlocal integration by parts, the equation (4.63) corresponds to the weak formulation of

$$-\operatorname{div}^s_\delta\big(\mathbb{1}_\Omega D_Af(\cdot,D^s_\delta u)\big) = F \quad \text{in } \Omega_\delta. \tag{4.64}$$

Within the region $\Omega_{-\delta}$, this equation reduces to the nonlocal Euler-Lagrange equation from [30, Theorem 8.2], while in the double boundary layer $\Gamma_{\pm\delta}$, the equation takes into account the geometry of the boundary $\partial\Omega$. More precisely, one obtains

$$\begin{cases} -\operatorname{div}_\delta^s(D_A f(x, D_\delta^s u)) = F & \text{in } \Omega_{-\delta}, \\ \mathcal{N}_\delta^s(D_A f(x, D_\delta^s u)) = F & \text{in } \Gamma_{\pm\delta}, \end{cases} \tag{4.65}$$

where $\mathcal{N}_\delta^s := -\operatorname{div}_\delta^s(\mathbb{1}_\Omega \,\cdot)$ coincides with the nonlocal boundary operator, recently introduced in [26, Definition 3.1] to prove a concise nonlocal integration by parts formula.

Now, if u solves (4.64) or (4.65) (weakly), the nonlocal divergence imposes that $\mathbb{1}_\Omega D_A f(\cdot, D_\delta^s u)$ must be regular enough across $\partial\Omega$. As $s \uparrow 1$, we expect to recover the natural boundary conditions $D_A f(\cdot, \nabla u) \cdot \nu = 0$ on $\partial\Omega$ with ν an outer normal to $\partial\Omega$. This intuition is made rigorous in the next section.

4.6.2 Localization for $s \uparrow 1$

We now turn to studying the limiting behavior of the nonlocal variational problem from Theorem 4.6.1, and the closely related nonlocal Neumann-type problems, as the fractional parameter s tends to 1. Our main result in this section (see Theorem 4.6.4) rigorously confirms the expectation that these problems localize, that is, they converge to their classical counterparts with usual gradients.

To start, let us collect in the next lemma some preparatory tools revolving around the asymptotic behavior of the sets $N^{s,p,\delta}(\Omega)$ and $N^{s,p,\delta}(\Omega)^\perp$ as s tends to 1. To capture the limit objects, we introduce $H^{1,p,\delta}(\Omega) := \{u \in L^p(\Omega_\delta) \,:\, u|_\Omega \in W^{1,p}(\Omega)\}$ and

$$N^{1,p,\delta}(\Omega) = \{u \in H^{1,p,\delta}(\Omega) \,:\, \nabla u = 0 \text{ in } \Omega\} = \{u \in L^p(\Omega_\delta) : u|_\Omega \text{ is constant}\}. \tag{4.66}$$

along with its corresponding metric projection $\pi_\delta^1 : L^p(\Omega_\delta) \to N^{1,p,\delta}(\Omega)$, and we also set

$$N^{1,p,\delta}(\Omega)^\perp := \{u \in H^{1,p,\delta}(\Omega; \mathbb{R}^m) \,:\, \pi_\delta^1(u) = 0\}.$$

Given the definition in (4.66), the projection $\pi_\delta^1(u)$ agrees with u in Γ_δ and is constant on Ω. Considering that $\arg\min_{c\in\mathbb{R}} \|u - c\|_{L^p(\Omega)} = 0$ is equivalent to $\int_\Omega |u|^{p-1} \operatorname{sign}(u)\, dx = 0$ for any $u \in L^p(\Omega_\delta)$, one can represent $N^{1,p,\delta}(\Omega)^\perp$ as

$$N^{1,p,\delta}(\Omega)^\perp = \left\{ u \in L^p(\Omega_\delta) : u|_\Omega \in W^{1,p}(\Omega),\ u = 0 \text{ a.e. in } \Gamma_\delta,\ \int_\Omega |u|^{p-1} \operatorname{sign}(u)\, dx = 0 \right\}. \tag{4.67}$$

When $p = 2$, the nonlinear integral condition in (4.67) reduces simply to the requirement of zero mean value.

Lemma 4.6.3. *Let $p \in (1, \infty)$ and let $(s_j)_j \subset (0, 1)$ be a sequence converging to 1. Then, these statements hold:*

(i) For all $v \in W^{1,p}(\mathbb{R}^n)$ it holds that $\mathcal{P}_\delta^{s_j} v \to v$ in $L^p(\Omega_\delta)$ as $j \to \infty$.

(ii) If $(u_j)_j \subset L^p(\Omega_\delta)$ converges to $u \in L^p(\Omega_\delta)$, then $\pi_\delta^{s_j}(u_j) \to \pi_\delta^1(u)$ as $j \to \infty$.

(iii) Let $(u_j)_j \subset L^p(\Omega_\delta)$ with $u_j \in N^{s_j,p,\delta}(\Omega)^\perp$ for all j. If $\sup_j \|D_\delta^{s_j} u_j\|_{L^p(\Omega;\mathbb{R}^n)} < \infty$, then there is a $u \in N^{1,p,\delta}(\Omega)^\perp$ such that (up to a non-relabeled subsequence)

$$u_j \to u \text{ in } L^p(\Omega_\delta) \quad \text{and} \quad D_\delta^{s_j} u_j \rightharpoonup \nabla u \text{ in } L^p(\Omega; \mathbb{R}^n) \text{ as } j \to \infty.$$

Proof. *Part* (i). Let $v \in W^{1,p}(\mathbb{R}^n)$. In light of (4.11) and (4.20), we find for $0 < \bar{s} \leq \inf_j s_j$ that

$$\sup_j \|\mathcal{P}_\delta^{s_j} v\|_{H^{\bar{s},p}(\mathbb{R}^n)} \leq \sup_j \|\mathcal{P}_\delta^{s_j} v\|_{H^{s_j,p}(\mathbb{R}^n)} < \infty.$$

Due to the compact embedding of $H^{\bar{s},p}(\mathbb{R}^n)$ into $L^p(\Omega_{2\delta})$ (see Section 4.2.2), there is a subsequence (not relabeled) such that $\mathcal{P}_\delta^{s_j} v \to w$ in $L^p(\Omega_{2\delta})$ for some $w \in L^p(\Omega_{2\delta})$. To identify w, consider an arbitrary test function $\varphi \in C_c^\infty(\Omega_\delta)$. As shown in [72, Eq. (3.4)], it holds that $\mathcal{Q}_\delta^{s_j}\varphi = Q_\delta^{s_j} * \varphi \to \varphi$ uniformly as $j \to \infty$. Together with Fubini's theorem, this implies

$$\begin{aligned}\int_{\Omega_\delta} w\varphi \, dx &= \lim_{j\to\infty} \int_{\Omega_{2\delta}} (\mathcal{P}_\delta^{s_j} v)\,(Q_\delta^{s_j} * \varphi)\, dx = \lim_{j\to\infty} \int_{\Omega_\delta} [Q_\delta^{s_j} * (\mathcal{P}_\delta^{s_j} v)]\,\varphi\, dx \\ &= \lim_{j\to\infty} \int_{\Omega_\delta} (\mathcal{Q}_\delta^{s_j}\mathcal{P}_\delta^{s_j} v)\,\varphi\, dx = \int_{\Omega_\delta} v\,\varphi\, dx,\end{aligned}$$

from which we infer $w = v$ on Ω_δ.

Part (ii). Since $0 \in N^{s,p,\delta}(\Omega)$ for all $s \in (0,1]$, we deduce from the definition of the metric projection that

$$\|\pi_\delta^{s_j}(u_j)\|_{L^p(\Omega_\delta)} \leq 2\|u_j\|_{L^p(\Omega_\delta)} \quad \text{for all } j.$$

As $(u_j)_j$ is bounded in $L^p(\Omega_\delta)$, so is $(\pi_\delta^{s_j}(u_j))_j$, and there exists a (non-relabeled) subsequence and a $w \in L^p(\Omega_\delta)$ with $\pi_\delta^{s_j}(u_j) \rightharpoonup w$ in $L^p(\Omega_\delta)$ as $j \to \infty$. For any test function $\psi \in C_c^\infty(\Omega;\mathbb{R}^n)$, one then obtains

$$\int_\Omega w \operatorname{div}\psi\, dx = \lim_{j\to\infty} \int_{\Omega_\delta} \pi_\delta^{s_j}(u_j)\, \operatorname{div}_\delta^{s_j}\psi\, dx = 0, \tag{4.68}$$

where the first inequality uses $\operatorname{div}_\delta^{s_j}\psi \to \operatorname{div}\psi$ uniformly on Ω_δ (see [72, Lemma 7]), and the last equality follows from integration by parts and the fact that $\pi_\delta^{s_j}(u_j) \in N^{s_j,p,\delta}(\Omega)$ has zero gradient $D_\delta^{s_j}$ for each j. By (4.68), the limit function w is constant on Ω, and hence, $w \in N^{1,p,\delta}(\Omega)$, cf. (4.66). It remains to show that $w = \pi_\delta^1(u)$ and that $\pi_\delta^{s_j}(u_j)$ converges even strongly.

To this aim, we first construct an auxiliary sequence $(h_j)_j \subset L^p(\Omega_\delta)$ with the properties that

$$h_j \in N^{s_j,p,\delta}(\Omega) \text{ for all } j \qquad \text{and} \qquad h_j \to \pi_\delta^1(u) \text{ in } L^p(\Omega_\delta) \text{ as } j \to \infty. \tag{4.69}$$

Since $\pi_\delta^1(u)$ is constant on Ω, one can find a sequence $(\varphi_k)_k \subset C_c^\infty(\Omega_\delta)$ that approximates $\pi_\delta^1(u)$ strongly in $L^p(\Omega_\delta)$ and satisfies that φ_k is constant on Ω for every k. Then, $\mathcal{P}_\delta^{s_j}\varphi_k \in N^{s_j,p,\delta}(\Omega)$ because of

$$D_\delta^{s_j}(\mathcal{P}_\delta^{s_j}\varphi_k) = \nabla(\mathcal{Q}_\delta^{s_j}\mathcal{P}_\delta^{s_j}\varphi_k) = \nabla\varphi_k = 0 \quad \text{on } \Omega,$$

and, along with part (i),

$$\lim_{k\to\infty}\lim_{j\to\infty} \|\mathcal{P}_\delta^{s_j}\varphi_k - \pi_\delta^1(u)\|_{L^p(\Omega_\delta)} = \lim_{k\to\infty} \|\varphi_k - \pi_\delta^1(u)\|_{L^p(\Omega_\delta)} = 0.$$

By extracting a suitable diagonal sequence, we obtain a sequence as in (4.69).

Now, with $(h_j)_j$ and the convergences $u_j \to u$ and $\pi_\delta^{s_j}(u_j) \rightharpoonup w$ in $L^p(\Omega_\delta)$ at hand, it follows that

$$\begin{aligned}\|u - \pi_\delta^1(u)\|_{L^p(\Omega_\delta)} \leq \|u - w\|_{L^p(\Omega_\delta)} &\leq \liminf_{j\to\infty} \|u_j - \pi_\delta^{s_j}(u_j)\|_{L^p(\Omega_\delta)} \\ &\leq \limsup_{j\to\infty} \|u_j - h_j\|_{L^p(\Omega_\delta)} = \|u - \pi_\delta^1(u)\|_{L^p(\Omega_\delta)}.\end{aligned}$$

As the inequalities in the previous lines turn to equalities, we infer $\pi_\delta^{s_j}(u_j) \to w = \pi_\delta^1(u)$ in $L^p(\Omega_\delta)$, which finishes the proof of (ii).

Part (iii). By Corollary 4.4.7, the sequence $(u_j)_j$ is bounded in $L^p(\Omega_\delta)$. Using that the extension operator $\mathcal{E}_\delta^s$ (see Section 4.4.2) is uniformly bounded with respect to s gives

$$\sup_j \|\mathcal{E}_\delta^{s_j} u_j\|_{H^{\bar{s},p}(\mathbb{R}^n)} \leq \sup_j \|\mathcal{E}_\delta^{s_j} u_j\|_{H^{s_j,p}(\mathbb{R}^n)} < \infty,$$

with $\bar{s} \in (0, \inf_j s_j]$. By the compact embedding of $H^{\bar{s},p}(\mathbb{R}^n)$ into $L^p(\Omega_\delta)$, we can extract a subsequence (not relabeled) and find a $w \in L^p(\Omega_\delta)$ such that $\mathcal{E}_\delta^{s_j} u_j \to w$ in $L^p(\Omega_\delta)$. A distributional argument in analogy to [72, Lemma 9] allows us to deduce that $w|_\Omega \in W^{1,p}(\Omega)$, or equivalently, $w \in H^{1,p,\delta}(\Omega)$, and

$$D_\delta^{s_j} u_j = D_\delta^{s_j} \mathcal{E}_\delta^{s_j} u_j \rightharpoonup \nabla w \qquad \text{in } L^p(\Omega; \mathbb{R}^n) \text{ as } j \to \infty. \tag{4.70}$$

Part (ii) shows on the other hand that $\pi_\delta^{s_j}(\mathcal{E}_\delta^{s_j} u_j) \to \pi_\delta^1(w)$ in $L^p(\Omega_\delta)$ as $j \to \infty$. Hence,

$$u_j = \mathcal{E}_\delta^{s_j} u_j + (u_j - \mathcal{E}_\delta^{s_j} u_j) = \mathcal{E}_\delta^{s_j} u_j - \pi_\delta^{s_j}(\mathcal{E}_\delta^{s_j} u_j) \to w - \pi_\delta^1(w) \quad \text{in } L^p(\Omega) \text{ as } j \to \infty; \tag{4.71}$$

note that the second equality is a consequence of $u_j - \mathcal{E}_\delta^{s_j} u_j \in N^{s_j,p,\delta}(\Omega)$, equation (4.47), and $u_j \in N^{s_j,p,\delta}(\Omega)^\perp$, which imply $\pi_\delta^{s_j}(\mathcal{E}_\delta^{s_j} u_j) - \mathcal{E}_\delta^{s_j} u_j + u_j = \pi_\delta^{s_j}(u_j) = 0$.

Finally, the statement follows from (4.70) and (4.71) with $u := w - \pi_\delta^1(w) \in N^{1,p,\delta}(\Omega)^\perp$, and the observation that $\pi_\delta^1(w) \in N^{1,p,\delta}(\Omega)$ is constant in Ω. □

We can now state and prove our localisation result in terms of variational convergence for $s \uparrow 1$. Using the framework of Γ-convergence (see e.g., [49,80]) guarantees the convergence of minimizers as a particular consequence.

Theorem 4.6.4 (Γ-convergence to classical variational integral). *Let $p \in (1, \infty)$, $F \in L^{p'}(\Omega_\delta; \mathbb{R}^m)$ and $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}_\infty$ be a Carathéodory integrand such that*

$$f(x, A) \geq c\left(|A|^p - 1\right) \quad \textit{for a.e. } x \in \Omega \textit{ and all } A \in \mathbb{R}^{m\times n} \tag{4.72}$$

with a constant $c > 0$. If $v \mapsto \int_\Omega f(x, \nabla v)\,dx$ is weakly lower semicontinuous on $W^{1,p}(\Omega; \mathbb{R}^m)$, then the family of functionals $(\mathcal{F}_\delta^s)_s$ with $\mathcal{F}_\delta^s : L^p(\Omega_\delta; \mathbb{R}^m) \to \mathbb{R}_\infty$ defined by

$$\mathcal{F}_\delta^s(u) = \begin{cases} \displaystyle\int_\Omega f(x, D_\delta^s u)\,dx - \int_{\Omega_\delta} F \cdot u\,dx & \textit{for } u \in N^{s,p,\delta}(\Omega; \mathbb{R}^m)^\perp, \\ \infty & \textit{else,} \end{cases}$$

Γ-converge with respect to $L^p(\Omega_\delta; \mathbb{R}^m)$-convergence as $s \to 1$ to $\mathcal{F}_\delta^1 : L^p(\Omega_\delta; \mathbb{R}^m) \to \mathbb{R}_\infty$ given by

$$\mathcal{F}_\delta^1(u) = \begin{cases} \displaystyle\int_\Omega f(x, \nabla u)\,dx - \int_\Omega F \cdot u\,dx & \textit{for } u \in N^{1,p,\delta}(\Omega; \mathbb{R}^m)^\perp, \\ \infty & \textit{else,} \end{cases}$$

with $N^{1,p,\delta}(\Omega; \mathbb{R}^m)^\perp$ as in (4.67). In addition, the family $(\mathcal{F}_\delta^s)_s$ is equi-coercive in $L^p(\Omega_\delta; \mathbb{R}^m)$.

Proof. Let $(s_j)_j$ be a sequence in $(0, 1)$ that converge to 1 as $j \to \infty$.

Step 1: Equi-coercivity. Let $(u_j)_j \subset L^p(\Omega_\delta)$ with $\sup_j \mathcal{F}_\delta^{s_j}(u_j) < \infty$, in particular, $u_j \in N^{s_j,p,\delta}(\Omega)^\perp$ for each j. The lower bound (4.72) together with the nonlocal Poincaré inequality in Corollary 4.4.7

with a constant independent of s shows that $(D_\delta^{s_j} u_j)_j$ is bounded in $L^p(\Omega;\mathbb{R}^{m\times n})$. Hence, the compactness result in Lemma 4.6.3 (iii) is applicable and immediately yields a subsequence of $(u_j)_j$ that converges strongly in $L^p(\Omega;\mathbb{R}^m)$ to a function in $N^{1,p,\delta}(\Omega)^\perp$.

Step 2: Liminf-inequality. Let $(s_j)_j \subset (0,1)$ and $(u_j)_j \subset L^p(\Omega_\delta)$ be sequences such that $s_j \to 1$, $u_j \to u$ in $L^p(\Omega_\delta;\mathbb{R}^m)$ as $j \to \infty$ and assume without loss of generality that $\sup_j \mathcal{F}_\delta^{s_j}(u_j) < \infty$. Then, according to Lemma 4.6.3 (iii) (cf. also Step 1), $u \in N^{1,p,\delta}(\Omega)^\perp$ with $D_\delta^{s_j} u_j \rightharpoonup \nabla u$ in $L^p(\Omega;\mathbb{R}^{m\times n})$ as $j \to \infty$. The desired liminf-inequality

$$\mathcal{F}_\delta^1(u) \le \liminf_{j\to\infty} \mathcal{F}_\delta^{s_j}(u_j)$$

is straightforward, if we exploit the weak lower semicontinuity of $v \mapsto \int_\Omega f(x,\nabla v)\,dx$ as in the proof of Theorem 4.6.1, but now with $\mathcal{Q}_\delta^{s_j}$ varying with j.

Step 3: Recovery sequence. Let $u \in N^{1,p,\delta}(\Omega)^\perp$ with $\mathcal{F}_\delta^1(u) < \infty$ and take $v \in W^{1,p}(\mathbb{R}^n;\mathbb{R}^m)$ with $v = u$ on Ω. We define a sequence $(u_j)_j \subset L^p(\Omega_\delta)$ by setting

$$u_j := \mathcal{P}_\delta^{s_j} v - \pi_\delta^{s_j}(\mathcal{P}_\delta^{s_j} v) \in N^{s_j,p,\delta}(\Omega;\mathbb{R}^m)^\perp.$$

By construction, it holds in view of (4.15) that, for every j,

$$D_\delta^{s_j} u_j = D_\delta^{s_j}(\mathcal{P}_\delta^{s_j} v) = \nabla v = \nabla u \quad \text{on } \Omega, \tag{4.73}$$

and Lemma 4.6.3 (i) and (ii) imply

$$u_j \to v - \pi_\delta^1(v) = u \quad \text{in } L^p(\Omega_\delta;\mathbb{R}^m) \text{ as } j \to \infty.$$

Observe that the identification of the limit function results from the fact that both u and $v - \pi_\delta^1(v)$ lie in $N^{1,p,\delta}(\Omega)^\perp$ and they have the same gradient in Ω.

Altogether, we have shown that $u_j \to u$ in $L^p(\Omega_\delta;\mathbb{R}^m)$ and

$$\mathcal{F}_\delta^{s_j}(u_j) = \int_\Omega f(x, D_\delta^{s_j} u_j)\,dx - \int_{\Omega_\delta} F\cdot u_j\,dx = \int_\Omega f(x,\nabla u)\,dx - \int_{\Omega_\delta} F\cdot u_j\,dx \longrightarrow \mathcal{F}_\delta^1(u)$$

as $j \to \infty$, which proves the stated Γ-convergence. □

Remark 4.6.5. We point out that the statement of Theorem 4.6.4 does not require any growth bound on f from above. This is of particular relevance in settings with polyconvex integrands, which - motivated by applications in elasticity theory - are often chosen to be extended-valued. In terms of the proof, the waiver of any upper bound on f is possible by the specific construction of the recovery sequence, whose nonlocal gradients are independent of j, see (4.73). △

Finally, we address what the previously shown convergence of the variational problems implies for the relation between local and nonlocal differential equations subject to natural and Neumann-type boundary conditions.

Indeed, if the classical compatibility condition $\int_\Omega F\,dx = 0$ holds, then any minimizer $u \in L^p(\Omega_\delta;\mathbb{R}^m)$ of $\mathcal{F}_\delta^1$, when restricted to Ω, also minimizes the functional

$$v \mapsto \int_\Omega f(x,\nabla v)\,dx - \int_\Omega F\cdot v\,dx$$

over the full space $W^{1,p}(\Omega;\mathbb{R}^m)$. In particular, if f is continuously differentiable in its second argument with f and $D_A f$ satisfying (4.62), then the minimizer u weakly satisfies the Euler-Lagrange system with natural boundary conditions

$$\begin{cases} -\operatorname{div}(D_A f(\cdot, \nabla u)) = F & \text{in } \Omega, \\ D_A f(\cdot, \nabla u) \cdot \nu = 0 & \text{on } \partial\Omega, \end{cases} \tag{4.74}$$

where ν is an outward pointing unit normal to $\partial\Omega$. Therefore, Theorem 4.6.4 implies that the minimizers of $\mathcal{F}^s_\delta$ converge up to subsequence in $L^p(\Omega;\mathbb{R}^m)$ to a weak solution of (4.74) as $s \uparrow 1$.

Chapter 5

Nonlocal gradients: Fundamental theorem of calculus, Poincaré inequalities and embeddings

This chapter corresponds to the preprint

[36] J. C. Bellido, C. Mora-Corral and H. Schönberger. Nonlocal gradients: Fundamental theorem of calculus, Poincaré inequalities and embeddings. Preprint arXiv:2402.16487, 2024.

5.1 Introduction

Many phenomena in nature, science and engineering are modeled with differential equations or local variational principles. Locality in this context means that the behaviour of an object depends only on its immediate neighborhood. However, there are situations where long-range interactions have to be taken into account. This gives rise to nonlocal models involving integro-differential equations or integral operators. The study of nonlocal models has proliferated in the last decades, as they provide effective ways to bridge between different length scales and lead to refined predictions. Areas that have benefited from nonlocal modeling include materials science, diffusion processes, imaging and machine learning.

An initial motivation for this work is *peridynamics*, which is a new nonlocal approach to solid mechanics [196] that has experienced huge progress and has led to a substantial literature; see, e.g., the books [44, 91, 116, 154], or the more mathematically-oriented articles [33, 47, 159]. Even though bond-based peridynamics, based on double-integral energies, is among the most widespread nonlocal models in mechanics, it was demonstrated in [27, 160] that it is incompatible with the classical model of nonlinear elasticity in the nonlocal-to-local limit.

To remedy this shortcoming of bond-based peridynamics, the authors of [28] adopted a model similar to the classical one ([23]) that involves, instead of the classical gradient, the *Riesz fractional gradient* D^s for $s \in (0,1)$, which is defined for smooth functions $u : \mathbb{R}^n \to \mathbb{R}$ as

$$D^s u(x) = c_{n,s} \int_{\mathbb{R}^n} \frac{u(x) - u(y)}{|x - y|^{n+s}} \frac{x - y}{|x - y|} \, dy. \tag{5.1}$$

The Riesz fractional gradient is a paradigmatic example of a nonlocal gradient. It was Shieh & Spector [193,194] who brought it to the attention of the PDE and calculus of variations communities. They introduced the function spaces associated to the Riesz gradient, with the key observation that they are equivalent to the *Bessel potential spaces* $H^{s,p}(\mathbb{R}^n)$. Moreover, a series of useful inequalities and embeddings hold. These fundamental insights laid the basis for an analysis of the equations and minimization problems related to the fractional gradient, which has led to the understanding of fractional function spaces [66, 179], the existence of solutions in variational problems [28, 140], and the relationship with local models [29].

A drawback of the Riesz fractional gradient for certain applications is the fact that it requires integration over the whole space for its computation. To be able to work on bounded domains, as is desirable, for instance, for realistic materials modeling, the previous approach was modified in [31, 72] by incorporating a horizon parameter. This was implemented by truncating the Riesz fractional gradient. To be more explicit, the nonlocal gradient for a fractional parameter $s \in (0,1)$, a horizon $\delta > 0$ and an appropriate non-negative, smooth, radial cut-off function w_δ supported in the ball $B_\delta(0)$ is defined as

$$D_\delta^s u(x) = \int_{B_\delta(x)} \frac{u(x)-u(y)}{|x-y|} \frac{x-y}{|x-y|} \frac{w_\delta(x-y)}{|x-y|^{n+s-1}} \, dy \quad \text{for } x \in \Omega. \tag{5.2}$$

Note that u needs to be defined in the larger region $\Omega_\delta = \{x \in \mathbb{R}^n : \operatorname{dist}(x,\Omega) < \delta\}$. As in the case of Sobolev spaces (and in the fractional case above), there are two ways to define the space $H^{s,p,\delta}(\mathbb{R}^n)$ associated with nonlocal gradients: as a completion of the smooth compactly supported functions under the norm $\|u\|_{L^p(\mathbb{R}^n)} + \|D_\delta^s u\|_{L^p(\mathbb{R}^n)}$, and through a distributional definition, which is based on a suitable integration by parts formula.

Analyses of the variational problems based on $D_\delta^s u$ can be found in [30, 31, 72, 141]. Two techniques established in those papers are the *translation method* [31, 72] and the *nonlocal fundamental theorem of calculus* [31]. The translation property is based on the observation that every nonlocal gradient is a classical gradient. Precisely,

$$D_\delta^s u = \nabla(Q_\delta^s * u), \tag{5.3}$$

where Q_δ^s is an integrable kernel supported in the ball $B_\delta(0)$. The nonlocal fundamental theorem of calculus refers to the representation formula

$$u = V_\delta^s * D_\delta^s u, \tag{5.4}$$

where V_δ^s is a locally integrable function implicitly given via Fourier transform. The identity (5.4) can be used to prove various embeddings and inequalities related to the nonlocal gradient D_δ^s and the function spaces $H^{s,p,\delta}(\mathbb{R}^n)$. Both (5.3) and (5.4) are modifications of analogous results in the fractional setting [66, 140, 179, 193].

Expressions (5.1) and (5.2) lead naturally to the central objects in this paper, which are general nonlocal gradients of the form

$$D_\rho u(x) = \int_{\mathbb{R}^n} \frac{u(x)-u(y)}{|x-y|} \frac{x-y}{|x-y|} \rho(x-y) \, dy$$

for some kernel ρ, typically with a singularity at the origin. Despite the fact that the axiomatic properties (invariance by translation and rotation, homogeneity and continuity [208]) of the Riesz gradient are desirable in many contexts, in some situations the use of other kernels ρ presents advantages. Perhaps the most relevant kernels for applications are those with compact support, since they allow for modeling phenomena in bounded domains. We denote by $H^{\rho,p}(\mathbb{R}^n)$ the set of $L^p(\mathbb{R}^n)$ functions with an $L^p(\mathbb{R}^n)$ distributional nonlocal gradient, while, given an open $\Omega \subset \mathbb{R}^n$, the set $H_0^{\rho,p}(\Omega)$ comprises those $H^{\rho,p}(\mathbb{R}^n)$ functions vanishing in the complement of Ω.

References on general nonlocal gradients are [92, 93, 102], where vector calculus for nonlocal gradients is addressed, and [161], on localization properties of nonlocal gradients. More precisely, this article benefits from the works [148] for non-radial kernels, [31, 72] for the operator D_δ^s, and

[129] for kernels supported in a half-ball. In fact, while this article was being written, we became aware of the preprint [16], which studies general nonlocal operators similar to ours, but focuses on different aspects: nonlocal-to-local estimates, localization, quasiconvexity and fine properties.

In this work we examine radial kernels, which model isotropic interactions, and its aim is to ascertain the assumptions on ρ that are necessary to develop a satisfactory theory for nonlocal gradients and its associated function spaces. In particular, we establish a set of hypotheses on ρ for which the main structural properties of the Riesz gradient carry over to general gradients. This lays the basis for the study of PDEs and variational problems based on nonlocal gradients.

We give here an overview of our main results. The first key property that we establish is an analogue of the translation method in (5.3), that is, we identify a locally integrable function Q_ρ such that

$$D_\rho u = \nabla(Q_\rho * u) \quad \text{for } u \in C_c^\infty(\mathbb{R}^n), \tag{5.5}$$

see Proposition 5.2.6. Beyond the simple representation that (5.5) provides, we use this formula to gain information about the operator D_ρ from the Fourier perspective by studying the Fourier transform of Q_ρ. This enables us to prove Poincaré inequalities and compact embeddings for the spaces $H_0^{\rho,p}(\Omega)$ with $\Omega \subset \mathbb{R}^n$ open and bounded. Precisely, beyond technical assumptions on ρ, we show that if

$$\liminf_{x \to 0} |x|^{n-1}\rho(x) > 0, \tag{5.6}$$

then there is a $C > 0$ such that

$$\|u\|_{L^p(\Omega)} \leq C\|D_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \quad \text{for all } u \in H_0^{\rho,p}(\Omega),$$

whereas $H_0^{\rho,p}(\Omega)$ is compactly embedded into $L^p(\mathbb{R}^n)$ if

$$\lim_{x \to 0} |x|^{n-1}\rho(x) = \infty, \tag{5.7}$$

see Theorem 5.4.5 for the case $p = 2$ and Theorem 5.4.11 for the general case $p \in (1,\infty)$; the setting with $p \in (1,\infty)$ requires an additional smoothness assumption on ρ in order to apply the Mihlin-Hörmander multiplier theorem. Remarkably, we show in Proposition 5.7.5 that the conditions (5.6) and (5.7) are essentially optimal when $p = 2$, which indicates that a singularity in the kernel ρ is unavoidable.

Under additional assumptions on ρ, we can push the Fourier analysis of D_ρ further and prove a nonlocal fundamental theorem of calculus as in (5.4),

$$u = V_\rho * D_\rho u \quad \text{for all } u \in H_0^{\rho,p}(\Omega), \tag{5.8}$$

with a locally integrable function V_ρ whose singularity at the origin is related to that of ρ, see Theorem 5.5.2 and Corollary 5.5.3. As an application, we use (5.8) to show embeddings of $H_0^{\rho,p}(\Omega)$ into Orlicz spaces and spaces with prescribed modulus of continuity, where the behavior of the Orlicz function and modulus of continuity are linked to the singularity of ρ at the origin, cf. Theorem 5.6.2 and Theorem 5.6.5. This provides a refinement and generalization of the fractional Sobolev and Morrey inequalities for D^s and D^s_δ [31, 193] that is not restricted to the scale of Lebesgue or Hölder spaces.

As is apparent from the previous paragraphs, not all results require the same assumptions on ρ. In fact, during the development of the theory, we will increasingly impose more conditions on the kernel. We will see several examples of ρ along the paper, but we anticipate a few of them for which the general theory holds:

$$\frac{\chi(x)}{|x|^{n+s-1}}, \qquad \frac{\chi(x)(-\log|x|)}{|x|^{n+s-1}}, \qquad \frac{\chi(x)}{|x|^{n+s-1}(-\log|x|)}, \qquad \frac{\chi(x)}{|x|^{n+s(|x|)-1}},$$

with $\chi \in C_c^\infty(\mathbb{R}^n)$ a non-negative, non-zero, radial function with some weak property of being decreasing (and supported in the unit ball in the second and third examples), $0 < s < 1$ and, in the last example, $s : (0, \infty) \to (0, 1)$ a smooth function. All those kernels have compact support; we will also see that this can be assumed without loss of generality, since only the behaviour of ρ near zero is relevant for the properties of the function spaces $H^{\rho,p}(\mathbb{R}^n)$ and $H_0^{\rho,p}(\Omega)$, see Proposition 5.3.10.

We finish this introduction with the comment that the community has not reached a consensus on the name or the notation for the spaces related to nonlocal gradients. In the introduction of [59], there is an interesting discussion on the origin of the spaces $H^{s,p}(\mathbb{R}^n)$, which are commonly called *Bessel potential spaces*, but for which the authors propose the name *Lions-Calderón spaces*. Of course, the spaces $H^{\rho,p}(\mathbb{R}^n)$ generalize Bessel potential spaces, which, in turn, are nonlocal versions of Sobolev spaces. Actually, in the case of Bessel potential spaces, those are obtained through complex interpolation between Sobolev $W^{1,p}$ and Lebesgue L^p spaces. Hence, an appropriate name for $H^{s,p}(\mathbb{R}^n)$ could be *fractional Sobolev spaces*, but this name is commonly reserved for the *Gagliardo* (or *Sobolev-Slobodeckij*) spaces $W^{s,p}(\mathbb{R}^n)$. In contrast to $H^{s,p}(\mathbb{R}^n)$, Gagliardo spaces are obtained by real interpolation between Sobolev and Lebesgue spaces. From our point of view, $H^{s,p}(\mathbb{R}^n)$ deserves the name of fractional Sobolev space in more right than $W^{s,p}(\mathbb{R}^n)$, as for the former there is a fractional differential object (the Riesz fractional gradient), whereas for the latter there is just a fractional seminorm. Although *ρ-Bessel space* could be a sensible name for $H^{\rho,p}(\mathbb{R}^n)$, we propose to call it *ρ-nonlocal Sobolev space*, since there is no immediate connection with the Bessel potentials.

The outline of this article is as follows. In Section 5.2, we establish the basic assumptions on ρ and properties of D_ρ including the translation method of (5.5). An analysis of the associated spaces $H^{\rho,p}(\mathbb{R}^n)$ and $H_0^{\rho,p}(\Omega)$ is performed in Section 5.3, providing a distributional definition of nonlocal gradients, a Leibniz rule and density results. In addition, Proposition 5.3.10 establishes a simple sufficient condition for the equality of spaces associated to different kernels, and for carrying over Poincaré inequality from one gradient to the other. Section 5.4 is devoted to the proof of the Poincaré inequalities and compact embeddings under the assumptions (5.6) and (5.7). The nonlocal fundamental theorem of calculus as in (5.8) is proven in Section 5.5, and the subsequent embeddings into Orlicz spaces and spaces of continuous function are presented in Section 5.6. Finally, in Section 5.7, we establish conditions for inclusions and equality of spaces associated to different kernels, and show that the conditions in (5.6) and (5.7) are almost optimal in order to have Poincaré inequalities and compact embeddings in L^2, respectively.

Notation

We fix the dimension $n \in \mathbb{N}$ of the space and an open set $\Omega \subset \mathbb{R}^n$. The vectors of the canonical basis of $\mathbb{R}^n$ are e_j, $j = 1, \dots, n$. The characteristic function of $A \subset \mathbb{R}^n$ is denoted by $\mathbb{1}_A$. The complement of a subset A in $\mathbb{R}^n$ is denoted by A^c, its closure by $\bar{A}$ and its boundary by ∂A. We write $B_r(x)$ for the open ball centred at $x \in \mathbb{R}^n$ of radius $r > 0$. We also set $B_r = B_r(0)$, $\mathbb{S}^{n-1} = \partial B_1$ and $\mathbb{S}_+^{n-1} = \{z \in \mathbb{S}^{n-1} : z_1 > 0\}$. The surface area in integrals is indicated by $\mathcal{H}^{n-1}$, while we set $\sigma_{n-1} = \mathcal{H}^{n-1}(\mathbb{S}^{n-1})$.

We will use an exponent $p \in [1, \infty]$ of integrability; its conjugate exponent $p' = \frac{p}{p-1}$. The notation for Lebesgue L^p and Sobolev $W^{1,p}$ spaces is standard. So is the notation for functions that are of class C^k, for k an integer or infinity. Their version of compact support are C_c^k. The support of a function is indicated by supp. We will indicate the domain and target of the functions, as in $L^p(\Omega, \mathbb{R}^n)$. The target is omitted if it is $\mathbb{R}$. We will use the abbreviation *a.e.* for *almost everywhere* or *almost every*. For $\alpha \in \mathbb{N}^n$, we give the standard meaning to the partial derivative ∂^α and the size $|\alpha|$; see [122, Sect. 2.2]. The operation of convolution is denoted by $*$.

The convention for the Fourier transform of a function $f \in L^1(\mathbb{R}^n)$ is

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i x\cdot\xi}\, dx.$$

This definition is extended by continuity and duality to other function and distribution spaces. The Schwartz space is denoted by $\mathcal{S}$ and the space of tempered distributions by $\mathcal{S}'$. The variable in the Fourier space is generically taken to be ξ. The inverse Fourier transform is denoted by $f^\vee$.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *radial* if there exists $\bar{f} : [0,\infty) \to \mathbb{R}$ (the *radial representation* of f) such that $f(x) = \bar{f}(|x|)$ for every $x \in \mathbb{R}^n$. A radial function $f : \mathbb{R}^n \to \mathbb{R}$ is *radially decreasing* if its radial representation is decreasing. A function $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ is *vector radial* if there exists $\bar{\varphi} : [0,\infty) \to \mathbb{R}$ such that $\varphi(x) = \bar{\varphi}(|x|)x$ for every $x \in \mathbb{R}^n$.

The words *increasing* and *decreasing* are meant in their wide (not strict) sense. In contrast, we use *positive* and *negative* with their strict meaning.

From Section 5.5, we will use the notion of *almost increasing* and *almost decreasing*. A function $f : I \to \mathbb{R}^n$ is almost increasing in the interval I if there exists a $C > 0$ such that for any $x_1, x_2 \in I$ with $x_1 \le x_2$, we have $f(x_1) \le Cf(x_2)$. An analogous definition is given for almost decreasing. We will denote by $C, C_k \ldots$ generic positive constants, which may vary from line to line.

For convenience of the reader, we collect here the assumptions made on the radial kernel $\rho : \mathbb{R}^n \to [0,\infty]$ that appear throughout the paper:

(H1) The function $f_\rho : (0,\infty) \to \mathbb{R},\ t \mapsto t^{n-2}\overline{\rho}(t)$ is decreasing, and there is a $0 < \mu < 1$ such that $\mu f_\rho(t/2) \ge f_\rho(t)$ for $t \in (0,\varepsilon)$;

(H2) the function f_ρ is smooth in $(0,\infty)$, and for $t \in (0,\varepsilon)$,

$$-C\frac{d}{dt}f_\rho(t) \ge \frac{f_\rho(t)}{t} \quad \text{and} \quad \left|\frac{d^k}{dt^k}f_\rho(t)\right| \le C_k\frac{f_\rho(t)}{t^k} \quad \text{for } k \in \mathbb{N};$$

(H3) the function $g_\rho : (0,\infty) \to \mathbb{R},\ t \mapsto t^{n+\sigma-1}\overline{\rho}(t)$ is almost decreasing on $(0,\varepsilon)$ for some $\sigma \in (0,1)$;

(H4) the function $h_\rho : (0,\infty) \to \mathbb{R},\ t \mapsto t^{n+\gamma-1}\overline{\rho}(t)$ is almost increasing on $(0,\varepsilon)$ for some $\gamma \in (0,1)$.

5.2 First properties of $\mathcal{G}_\rho$

In this section we show some principal properties of the ρ-derivative $\mathcal{G}_\rho u$ and its Fourier transform for $u \in C_c^\infty(\mathbb{R}^n)$.

We always make the following basic assumptions on ρ:

$$\text{(H0)} \quad \begin{cases} \rho : \mathbb{R}^n \to [0,\infty] \text{ is radial with } \rho(x) \in \mathbb{R} \text{ for all } x \in \mathbb{R}^n \setminus \{0\}, \\ \rho \in L^1_{\mathrm{loc}}(\mathbb{R}^n) \text{ with } \displaystyle\int_{\mathbb{R}^n} \min\{1, |x|^{-1}\}\rho(x)\, dx < \infty, \\ \displaystyle\inf_{B_\varepsilon} \rho > 0 \text{ for some } \varepsilon > 0. \end{cases}$$

Similar assumptions to (H0) have appeared in the literature; e.g., [129]. Note that we consider ρ as a function everywhere defined in $\mathbb{R}^n$, and not as an equivalence class of functions defined a.e.

Clearly, (H0) implies that

$$\int_{B_r} \rho(x)\, dx + \int_{B_r^c} \frac{\rho(x)}{|x|}\, dx < \infty, \qquad r > 0, \tag{5.9}$$

which, in terms of the radial representation $\overline{\rho}$, can be written as

$$\int_0^r \overline{\rho}(t)t^{n-1}\,dt + \int_r^\infty \overline{\rho}(t)t^{n-2}\,dt < \infty, \qquad r > 0. \tag{5.10}$$

In fact, under the assumption (H0), we have the equivalence

$$\begin{aligned} \rho \in L^1(\mathbb{R}^n) &\iff \int_r^\infty \overline{\rho}(t)t^{n-1}\,dt < \infty \text{ for some } r > 0 \\ &\iff \int_r^\infty \overline{\rho}(t)t^{n-1}\,dt < \infty \text{ for all } r > 0. \end{aligned} \tag{5.11}$$

Example 5.2.1. Classes of kernels ρ satisfying (H0) are:

(a) Given $s \in (0,1)$,
$$\rho(x) = \frac{1}{|x|^{n+s-1}}. \tag{5.12}$$

(b) Given a continuous $s : [0,\infty) \to (0,1)$ with $\inf_{[0,\infty)} s > 0$,
$$\rho(x) = \frac{1}{|x|^{n+s(|x|)-1}}.$$

(c) Given $0 \leq \alpha < n$ and $\beta > n-1$,
$$\rho(x) = \frac{\mathbb{1}_{B_1}(x)}{|x|^\alpha} + \frac{\mathbb{1}_{B_1^c}(x)}{|x|^\beta}.$$

(d) If ρ satisfies (H0) and $\chi \in L^\infty(\mathbb{R}^n)$ is radial with $\chi \geq 0$ and $\inf_{B_\varepsilon} \chi > 0$ then $\chi\rho$ satisfies (H0).

(e) If ρ_1, ρ_2 satisfy (H0) and $\alpha_1, \alpha_2 > 0$ then $\alpha_1\rho_1 + \alpha_2\rho_2$ satisfies (H0).

(f) If ρ_1, ρ_2 satisfy (H0) then any measurable radial ρ with $\rho_1 \leq \rho \leq \rho_2$ satisfies (H0).

Definition 5.2.2. *For $u \in C_c^\infty(\mathbb{R}^n)$, we define the nonlocal gradient of u as*

$$\mathcal{G}_\rho u(x) = \int_{\mathbb{R}^n} \frac{u(x)-u(y)}{|x-y|}\frac{x-y}{|x-y|}\rho(x-y)\,dy, \qquad x \in \mathbb{R}^n. \tag{5.13}$$

The following result shows that the nonlocal gradient defines an integrable and bounded function, which can be deduced from the more general statement in [92, Prop. 1], but we provide the details for the reader's convenience.

Lemma 5.2.3. *Let $u \in C_c^\infty(\mathbb{R}^n)$. Then $\mathcal{G}_\rho u \in L^1(\mathbb{R}^n,\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n,\mathbb{R}^n)$ and the integral* (5.13) *is absolutely convergent for each $x \in \mathbb{R}^n$.*

Proof. Let $L > 0$ be a Lipschitz constant of u, then we can bound

$$\begin{aligned} \int_{\mathbb{R}^n} \left| \frac{u(x)-u(y)}{|x-y|}\frac{x-y}{|x-y|}\rho(x-y) \right| dy &\leq \max\{L, 2\|u\|_\infty\} \int_{\mathbb{R}^n} \min\{1, |x-y|^{-1}\}\rho(x-y)\,dy \\ &= \max\{L, 2\|u\|_\infty\} \int_{\mathbb{R}^n} \min\{1, |z|^{-1}\}\rho(z)\,dz. \end{aligned}$$

Thus, $\mathcal{G}_\rho u$ is bounded thanks to (H0).

Next, let $K = \operatorname{supp} u$, $\delta > 0$, and define $K_\delta := K + B_\delta$. In order to prove that $\mathcal{G}_\rho u \in L^1(\mathbb{R}^n, \mathbb{R}^n)$, it is enough to show that $\mathcal{G}_\rho u \in L^1(K_\delta^c, \mathbb{R}^n)$. For $x \notin K_\delta$ we have

$$|\mathcal{G}_\rho u(x)|\, dx \le \int_K \frac{|u(y)|}{|x-y|}\rho(x-y)\, dy,$$

so

$$\int_{K_\delta^c} |\mathcal{G}_\rho u(x)|\, dx \le \int_K |u(y)| \int_{K_\delta^c} \frac{1}{|x-y|}\rho(x-y)\, dx\, dy \le \|u\|_1 \int_{B_\delta^c} \frac{\rho(z)}{|z|}\, dz,$$

which is finite thanks to (5.9). □

Define $\lambda_\rho : \mathbb{R}^n \to \mathbb{C}^n$ as

$$\lambda_\rho(\xi) := \int_{\mathbb{R}^n} \frac{\rho(x)x}{|x|^2}(e^{2\pi i \xi\cdot x} - 1)\, dx.$$

Similarly as in [129, Lemma 2.13] (see also [148, Lemma 1.2]), the following result shows that λ_ρ is the Fourier multiplier associated to the operator $\mathcal{G}_\rho$.

Lemma 5.2.4. *If $u \in C_c^\infty(\mathbb{R}^n)$ then*

$$\widehat{\mathcal{G}_\rho u} = \lambda_\rho \widehat{u}. \tag{5.14}$$

Moreover,

$$\lambda_\rho(\xi) = \frac{i\xi}{|\xi|} \int_{\mathbb{R}^n} \frac{\rho(x)x_1}{|x|^2} \sin(2\pi|\xi|x_1)\, dx, \qquad \xi \in \mathbb{R}^n \setminus \{0\}$$

and $|\lambda_\rho(\xi)| < \infty$ for all $\xi \in \mathbb{R}^n$.

Proof. Thanks to Lemma 5.2.3, $\mathcal{G}_\rho u \in L^1(\mathbb{R}^n, \mathbb{R}^n)$ and the following calculation is valid for each $\xi \in \mathbb{R}^n$:

$$\begin{aligned}
\widehat{\mathcal{G}_\rho u}(\xi) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{u(x)-u(y)}{|x-y|}\frac{x-y}{|x-y|}\rho(x-y)e^{-2\pi i x\cdot\xi}\, dy\, dx \\
&= -\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{u(y-z)-u(y)}{|z|}\frac{z}{|z|}\rho(z)e^{-2\pi i(y-z)\cdot\xi}\, dy\, dz \\
&= -\int_{\mathbb{R}^n}\left[\int_{\mathbb{R}^n} u(y-z)e^{-2\pi i(y-z)\cdot\xi}\, dy - e^{2\pi i z\cdot\xi}\int_{\mathbb{R}^n} u(y)e^{-2\pi i y\cdot\xi}\, dy\right]\frac{z}{|z|^2}\rho(z)\, dz \\
&= -\int_{\mathbb{R}^n}\left[\widehat{u}(\xi) - e^{2\pi i z\cdot\xi}\widehat{u}(\xi)\right]\frac{z}{|z|^2}\rho(z)\, dz \\
&= \widehat{u}(\xi)\int_{\mathbb{R}^n}\left(e^{2\pi i z\cdot\xi} - 1\right)\frac{z}{|z|^2}\rho(z)\, dz,
\end{aligned}$$

so (5.14) is proved. The alternative expression for λ_ρ is obtained as follows. The integral

$$\int_{\mathbb{R}^n} \frac{\rho(x)x}{|x|^2}(\cos(2\pi\xi\cdot x) - 1)\, dx$$

is zero since it is absolutely convergent with an odd integrand. Therefore,

$$\lambda_\rho(\xi) = i\int_{\mathbb{R}^n} \frac{\rho(x)x}{|x|^2}\sin(2\pi\xi\cdot x)\, dx.$$

Now, as ρ is radial, if $R \in \mathrm{SO}(n)$ is a rotation then

$$\lambda_\rho(R\xi) = i\int_{\mathbb{R}^n} \frac{\rho(x)x}{|x|^2}\sin(2\pi R\xi\cdot x)\, dx = i\int_{\mathbb{R}^n} \frac{\rho(x)x}{|x|^2}\sin(2\pi\xi\cdot R^T x)\, dx = R\lambda_\rho(\xi).$$

Now, given $\xi \in \mathbb{R}^n$, choose $R \in \mathrm{SO}(n)$ such that $R\xi = |\xi|e_1$. Then

$$\begin{aligned}\lambda_\rho(\xi) = R^T\lambda_\rho(|\xi|e_1) &= iR^T\int_{\mathbb{R}^n}\frac{\rho(x)x}{|x|^2}\sin(2\pi|\xi|x_1)\,dx = iR^T\int_{\mathbb{R}^n}\frac{\rho(x)x_1e_1}{|x|^2}\sin(2\pi|\xi|x_1)\,dx\\ &= \frac{i\xi}{|\xi|}\int_{\mathbb{R}^n}\frac{\rho(x)x_1}{|x|^2}\sin(2\pi|\xi|x_1)\,dx,\end{aligned}$$

as desired.

Finally, using the bound $|\sin t| \leq \min\{1,|t|\}$ ($t \in \mathbb{R}$), we find that

$$|\lambda_\rho(\xi)| \leq \int_{\mathbb{R}^n}\frac{\rho(x)}{|x|}|\sin(2\pi|\xi|x_1)|\,dx \leq \int_{\mathbb{R}^n}\rho(x)\min\{|x|^{-1},2\pi|\xi|\}\,dx < \infty,$$

in light of (5.9). □

Now we employ a strategy described in [31, Prop. 4.3] (itself based on [179, Lemma 15.9]) consisting of studying a potential of $x \mapsto -\frac{\overline{\rho}(|x|)}{|x|}\frac{x}{|x|}$. For this, we define the function

$$Q_\rho(x) := \int_{|x|}^\infty \frac{\overline{\rho}(t)}{t}\,dt, \qquad x \in \mathbb{R}^n\setminus\{0\},$$

which is well defined and finite due to (5.10). We present some of its immediate properties.

Lemma 5.2.5. *The following statements hold:*

(i) For each $0 < a < b$ *we have*

$$Q_\rho \in W^{1,1}(B_b\setminus B_a), \quad Q_\rho \in L^1(B_b) \quad \text{and} \quad \nabla Q_\rho \in L^1(B_a^c,\mathbb{R}^n),$$

and for a.e. $x \in \mathbb{R}^n\setminus\{0\}$,

$$\nabla Q_\rho(x) = -\frac{\overline{\rho}(|x|)}{|x|}\frac{x}{|x|}.$$

(ii) For every $M > 0$ *there is a* $C = C(n,M) > 0$ *such that*

$$|Q_\rho(x)| \leq \frac{C}{|x|^{n-1}} \quad \text{for } |x| \geq M.$$

(iii) If $\rho \in L^1(\mathbb{R}^n)$ *then* $Q_\rho \in L^1(\mathbb{R}^n)$. *Moreover, when* ρ *has compact support then* Q_ρ *lies in* $L^1(\mathbb{R}^n)$ *and also has compact support.*

Proof. Part (i). Let $0 < a < 1$. Then

$$\int_a^1\frac{\overline{\rho}(t)}{t}\,dt + \int_1^\infty\frac{\overline{\rho}(t)}{t}\,dt \leq \frac{1}{a^n}\int_a^1\overline{\rho}(t)t^{n-1}\,dt + \int_1^\infty\overline{\rho}(t)t^{n-2}\,dt < \infty, \tag{5.15}$$

in view of (5.10). Consequently, the radial representation $\overline{Q}_\rho$ of Q_ρ is locally Sobolev in $(0,\infty)$ (see, e.g., [149, Thms. 3.29 and 7.16]). An argument similar to [24, Lemma 4.1] shows that Q_ρ is locally Sobolev in $\mathbb{R}^n\setminus\{0\}$ and its distributional derivative coincides with its classical derivative a.e.

A straightforward calculation based on the coarea formula and Fubini's theorem shows that

$$\int_{B_b}Q_\rho(x)\,dx = \frac{\sigma_{n-1}}{n}\left[\int_0^b\overline{\rho}(t)t^{n-1}\,dt + b^n\int_b^\infty\frac{\overline{\rho}(t)}{t}\,dt\right] < \infty,$$

thanks to (5.10) and (5.15).

As $\overline{Q}_\rho$ is absolutely continuous,

$$\overline{Q}_\rho'(t) = -\frac{\overline{\rho}(t)}{t}, \qquad \text{for a.e. } t > 0,$$

so

$$\nabla Q_\rho(x) = \overline{Q}_\rho'(|x|)\frac{x}{|x|}, \qquad \text{for a.e. } x \in \mathbb{R}^n \setminus \{0\}.$$

With this we clearly have $\nabla Q_\rho \in L^1(B_a^c, \mathbb{R}^n)$, due to (5.9).

Part (ii). For any $t \geq M > 0$ we have

$$\overline{Q}_\rho(t) = \int_t^\infty \frac{\overline{\rho}(r)}{r}\, dr \leq \frac{1}{t^{n-1}} \int_t^\infty \overline{\rho}(r) r^{n-2}\, dr \leq \frac{1}{t^{n-1}} \int_M^\infty \overline{\rho}(r) r^{n-2}\, dr$$

and (ii) is concluded thanks to (5.10).

Part (iii). Assume $\rho \in L^1(\mathbb{R}^n)$. In order to show that $Q_\rho \in L^1(\mathbb{R}^n)$ it is enough to check that $Q_\rho \in L^1(B_b^c)$, due to (i). A straightforward calculation shows that

$$\int_{B_b^c} Q_\rho(x)\, dx = \sigma_{n-1} \int_b^\infty \frac{\overline{\rho}(r)}{r} \frac{r^n - b^n}{n}\, dr \leq \sigma_{n-1} \int_b^\infty \overline{\rho}(r) r^{n-1}\, dr < \infty,$$

in view of (5.11). If, in addition, ρ has compact support then so does Q_ρ, and, hence, $Q_\rho \in L^1(\mathbb{R}^n)$. □

As a consequence of part (iii) of Lemma 5.2.5, if $\rho \in L^1(\mathbb{R}^n)$ then $\widehat{Q}_\rho$ is a continuous function, which is analytic if ρ has compact support. In fact, in the development of the theory, we will often assume that $\operatorname{supp} \rho = \overline{B_\delta}$ for some $\delta > 0$. In this case, for an open set $\Omega \subset \mathbb{R}^n$ we define $\Omega_\delta = \Omega + B_\delta$, and note that $\mathcal{G}_\rho \varphi$ is supported in Ω_δ for $\varphi \in C_c^\infty(\Omega)$.

We now show that the nonlocal gradient can be written as the convolution of Q_ρ with the classical gradient, and derive a formula for $\widehat{Q}_\rho$.

Proposition 5.2.6. *The following two statements hold:*

(i) For $u \in C_c^\infty(\mathbb{R}^n)$, we have that $\mathcal{G}_\rho u \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ and

$$\mathcal{G}_\rho u = Q_\rho * \nabla u = \nabla (Q_\rho * u).$$

(ii) If, in addition, $\rho \in L^1(\mathbb{R}^n)$, then

$$\widehat{\mathcal{G}_\rho u}(\xi) = 2\pi i \xi \widehat{Q}_\rho(\xi) \widehat{u}(\xi) \quad \text{and} \quad \lambda_\rho(\xi) = 2\pi i \xi \widehat{Q}_\rho(\xi), \qquad \xi \in \mathbb{R}^n$$

and

$$\widehat{Q}_\rho(\xi) = \frac{1}{2\pi|\xi|} \int_{\mathbb{R}^n} \frac{\rho(x) x_1}{|x|^2} \sin(2\pi |\xi| x_1)\, dx, \qquad \xi \in \mathbb{R}^n \setminus \{0\}. \tag{5.16}$$

Proof. Part (i). We consider $x, e \in \mathbb{R}^n$ with $|e| = 1$ and the vector field

$$\beta : \mathbb{R}^n \setminus \{x\} \to \mathbb{R}^n, \qquad \beta(y) = (u(x) - u(y))\, Q_\rho(x - y)\, e.$$

For any $0 < a < b$, by Lemma 5.2.5, $\beta \in W^{1,1}(B_b(x) \setminus B_a(x), \mathbb{R}^n)$. By the divergence theorem (e.g., [149, Th. 18.1])

$$\begin{aligned}\int_{B_b(x)\setminus B_a(x)} \operatorname{div} \beta(y)\, dy = \int_{\partial B_b(x)} &\beta(y) \cdot \nu_{B_b(x)}(y)\, d\mathcal{H}^{n-1}(y) \\ &- \int_{\partial B_a(x)} \beta(y) \cdot \nu_{B_a(x)}(y)\, d\mathcal{H}^{n-1}(y),\end{aligned} \tag{5.17}$$

where $\nu_{B_r(x)}$ is the exterior normal to $B_r(x)$, for $r = a, b$. In fact,

$$\operatorname{div}\beta(y) = -Q_\rho(x-y)\,\nabla u(y)\cdot e - (u(x)-u(y))\nabla Q_\rho(x-y)\cdot e, \qquad \text{a.e. } y \in \mathbb{R}^n.$$

It turns out that both terms of the right-hand side of the formula above are in $L^1(\mathbb{R}^n)$. Indeed, by Lemma 5.2.5 and the fact that u has compact support, we have that the map $y \mapsto Q_\rho(x-y)\,\nabla u(y)$ is in $L^1(\mathbb{R}^n,\mathbb{R}^n)$. Analogously, the map $y \mapsto (u(x)-u(y))\nabla Q_\rho(x-y)$ is in $L^1(B_a(x)^c,\mathbb{R}^n)$, and, as u is Lipschitz, we have

$$\begin{aligned}\int_{B_a(x)} |(u(x)-u(y))\nabla Q_\rho(x-y)|\,dy &\leq \|\nabla u\|_\infty \int_{B_a(x)} |x-y||\nabla Q_\rho(x-y)|\,dy\\ &= \|\nabla u\|_\infty \sigma_{n-1}\int_0^a t^{n-1}\overline{\rho}(t)\,dt < \infty,\end{aligned}$$

in view of Lemma 5.2.5 and (5.10). In particular,

$$\begin{aligned}\lim_{\substack{a\searrow 0\\ b\to\infty}} \int_{B_b(x)\setminus B_a(x)} \operatorname{div}\beta(y)\,dy = &-\int_{\mathbb{R}^n} Q_\rho(x-y)\,\nabla u(y)\cdot e\,dy\\ &-\int_{\mathbb{R}^n}(u(x)-u(y))\nabla Q_\rho(x-y)\cdot e\,dy.\end{aligned} \tag{5.18}$$

By Lemma 5.2.5, $Q_\rho \in L^1(B_r)$ for all $r > 0$, so

$$\int_0^r \overline{Q}_\rho(t)t^{n-1}\,dt < \infty,$$

which implies that $\liminf_{a\downarrow 0} a^n\overline{Q}_\rho(a) = 0$. Let $\{a_j\}_{j\in\mathbb{N}}$ be a sequence of positive numbers tending to zero such that $\lim_{j\to\infty} a_j^n\overline{Q}_\rho(a_j) = 0$. As u is Lipschitz,

$$\left|\int_{\partial B_{a_j}(x)} \beta\cdot\nu_{B_{a_j}(x)}\,d\mathcal{H}^{n-1}\right| \leq \int_{\partial B_{a_j}(x)} |\beta|\,d\mathcal{H}^{n-1} \leq \|\nabla u\|_\infty\,\sigma_{n-1}a_j^n\overline{Q}_\rho(a_j) \to 0 \quad \text{as } j\to\infty,$$

and, as u has compact support, if b is big enough,

$$\int_{\partial B_b(x)} \beta\cdot\nu_{B_b(x)}\,d\mathcal{H}^{n-1} = u(x)\int_{\partial B_b(x)} Q_\rho(x-y)\,e\cdot\nu_{B_b(x)}\,d\mathcal{H}^{n-1}(y) = 0$$

by symmetry. Together with (5.17) and (5.18), this yields

$$\int_{\mathbb{R}^n} Q_\rho(x-y)\,\nabla u(y)\cdot e\,dy = -\int_{\mathbb{R}^n}(u(x)-u(y))\nabla Q_\rho(x-y)\cdot e\,dy.$$

As this is true for every $e \in \mathbb{R}^n$ with $|e| = 1$, we conclude that

$$\int_{\mathbb{R}^n} Q_\rho(x-y)\,\nabla u(y)\,dy = -\int_{\mathbb{R}^n}(u(x)-u(y))\nabla Q_\rho(x-y)\,dy.$$

In light of Lemma 5.2.5 (i), this equality shows that $Q_\rho * \nabla u(x) = \mathcal{G}_\rho u(x)$. Naturally, we also have $Q_\rho * \nabla u = \nabla(Q_\rho * u)$, since $u \in C_c^\infty(\mathbb{R}^n)$ and $Q_\rho \in L^1_{\text{loc}}(\mathbb{R}^n)$. In particular, $\mathcal{G}_\rho u \in C^\infty(\mathbb{R}^n,\mathbb{R}^n)$.

Part (ii). If $\rho \in L^1(\mathbb{R}^n)$ we have $Q_\rho \in L^1(\mathbb{R}^n)$ thanks to Lemma 5.2.5. Consequently, taking Fourier transforms in the expression $\mathcal{G}_\rho u = Q_\rho * \nabla u$, we conclude that $\widehat{\mathcal{G}_\rho u}(\xi) = 2\pi i\xi\widehat{Q}_\rho(\xi)\widehat{u}(\xi)$ for all $\xi \in \mathbb{R}^n$. Comparing this expression with that of (5.14), we obtain the correspondence $\lambda_\rho(\xi) = 2\pi i\xi\widehat{Q}_\rho(\xi)$ and, hence, the equality

$$\widehat{Q}_\rho(\xi) = \frac{-i\xi}{2\pi|\xi|^2} \cdot \lambda_\rho(\xi) = \frac{1}{2\pi|\xi|} \int_{\mathbb{R}^n} \frac{\rho(x)x_1}{|x|^2} \sin(2\pi|\xi|x_1)\, dx, \qquad \xi \in \mathbb{R}^n \setminus \{0\}$$

holds thanks to Lemma 5.2.4. □

Remark 5.2.7. If $Q_\rho \in L^1(\mathbb{R}^n)$ and $\xi \in \mathbb{R}^n \setminus \{0\}$, we can use the formula for the Fourier transform of a radial function, see [122, Appendix B.5], to find the following alternative expression

$$\begin{aligned} \widehat{Q}_\rho(\xi) &= \frac{2\pi}{|\xi|^{\frac{n-2}{2}}} \int_0^\infty \int_r^\infty \frac{\overline{\rho}(t)}{t}\, dt\, J_{\frac{n}{2}-1}(2\pi|\xi|r) r^{\frac{n}{2}}\, dr \\ &= \frac{2\pi}{|\xi|^{\frac{n-2}{2}}} \int_0^\infty \frac{\overline{\rho}(t)}{t} \int_0^t J_{\frac{n}{2}-1}(2\pi|\xi|r) r^{\frac{n}{2}}\, dr\, dt = \frac{1}{|\xi|^{\frac{n}{2}}} \int_0^\infty \overline{\rho}(t) t^{\frac{n}{2}-1} J_{\frac{n}{2}}(2\pi|\xi|t)\, dt, \end{aligned} \tag{5.19}$$

with J_ν for $\nu > 0$ the Bessel function of the first kind. In the last line we used

$$\int_0^t J_{\frac{n}{2}-1}(2\pi|\xi|r) r^{\frac{n}{2}}\, dr = \frac{t^{\frac{n}{2}}}{2\pi|\xi|} J_{\frac{n}{2}}(2\pi t|\xi|),$$

which follows from the identity in [122, Appendix B.3]. The integral in (5.19) also appears in [16] through different methods. △

Part (ii) of Proposition 5.2.6 formally shows that, for the nonlocal gradient $\mathcal{G}_\rho$, the fundamental theorem of calculus in Fourier space looks like

$$\widehat{u}(\xi) = \frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)} \cdot \widehat{\mathcal{G}_\rho u}(\xi), \qquad \xi \in \mathbb{R}^n \setminus \{0\}, \tag{5.20}$$

which motivates the further study of the Fourier transform of Q_ρ. This will be carried out in Section 5.4.

5.3 Function spaces

In this section we establish the definition and first properties of the spaces $H^{\rho,p}(\mathbb{R}^n)$ and $H_0^{\rho,p}(\Omega)$, including density results. Then, we show a sufficient condition for the equivalence of spaces associated with different kernels.

5.3.1 Definition and first properties

The aim of this section is to introduce the space of L^p-functions whose nonlocal gradient is an L^p-function in analogy to Sobolev spaces, and derive some of its general properties. We start by defining the nonlocal divergence.

Definition 5.3.1. *For a vector field $\Phi \in C_c^\infty(\mathbb{R}^n, \mathbb{R}^n)$, we define its nonlocal divergence as*

$$\operatorname{div}_\rho \Phi(x) = \int_{\mathbb{R}^n} \frac{\Phi(x) - \Phi(y)}{|x-y|} \cdot \frac{x-y}{|x-y|} \rho(x-y)\, dy, \qquad x \in \mathbb{R}^n.$$

The following proposition states that the nonlocal divergence is the dual operator of the nonlocal gradient $\mathcal{G}_\rho$ in the sense of integration by parts. Different proofs of analogous results have appeared in the literature, for instance [28, 66, 161, 208]. We do not include the proof here and refer the interested reader to any of those references.

Proposition 5.3.2. *Let $u \in L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$ be such that*

$$\int_K \int_{\mathbb{R}^n} \frac{|u(x) - u(y)|}{|x-y|} \rho(x-y)\, dx\, dy < \infty \tag{5.21}$$

for any compact set $K \subset \mathbb{R}^n$. Then, for any $\Phi \in C_c^\infty(\mathbb{R}^n, \mathbb{R}^n)$, the integration by parts formula

$$\int_{\mathbb{R}^n} \mathcal{G}_\rho u(x) \cdot \Phi(x)\, dx = -\int_{\mathbb{R}^n} u(x) \operatorname{div}_\rho \Phi(x)\, dx$$

holds.

Notice that the hypothesis (5.21) guarantees that $\mathcal{G}_\rho u(x)$ exists as a Lebesgue integral for a.e. $x \in \mathbb{R}^n$, and $\mathcal{G}_\rho u \in L^1_{\text{loc}}(\mathbb{R}^n, \mathbb{R}^n)$; the assumption $u \in L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$ combined with Lemma 5.2.3 ensures that $u \operatorname{div}_\rho \Phi$ is integrable. The integration by parts formula leads naturally to the definition of the distributional nonlocal gradient.

Definition 5.3.3. *Given $u \in L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$, we define its distributional, or weak, nonlocal gradient $D_\rho u$ as the distribution*

$$\langle D_\rho u, \Phi \rangle = -\int_{\mathbb{R}^n} u(x) \operatorname{div}_\rho \Phi(x)\, dx, \qquad \Phi \in C_c^\infty(\mathbb{R}^n, \mathbb{R}^n).$$

Checking that $D_\rho u$ is a distribution is elementary, given Lemma 5.2.3. Thanks to Proposition 5.3.2, if $u \in L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$ satisfies (5.21), its nonlocal gradient and its distributional nonlocal gradient coincide: $\mathcal{G}_\rho u = D_\rho u$.

Definition 5.3.4. *Let $p \in [1, \infty]$. We define the ρ-nonlocal Sobolev space $H^{\rho,p}(\mathbb{R}^n)$ as the set of functions $u \in L^p(\mathbb{R}^n)$ such that $D_\rho u \in L^p(\mathbb{R}^n, \mathbb{R}^n)$, equipped with the norm*

$$\|u\|_{H^{\rho,p}(\mathbb{R}^n)} = \begin{cases} \left(\|u\|_{L^p(\mathbb{R}^n)} + \|D_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)}\right)^{\frac{1}{p}} & \text{for } p \in [1,\infty), \\ \max\{\|u\|_{L^\infty(\mathbb{R}^n)}, \|D_\rho u\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)}\} & \text{for } p = \infty. \end{cases}$$

For an open set $\Omega \subset \mathbb{R}^n$, we define the closed subspace

$$H_0^{\rho,p}(\Omega) := \{u \in H^{\rho,p}(\mathbb{R}^n) \,:\, u = 0 \text{ a.e. in } \Omega^c\}.$$

We also denote $H^\rho(\mathbb{R}^n) = H^{\rho,2}(\mathbb{R}^n)$ and $H_0^\rho(\Omega) = H_0^{\rho,2}(\Omega)$.

Standard arguments show that $H^{\rho,p}(\mathbb{R}^n)$ is a Banach space, which is separable for $p \in [1, \infty)$, reflexive for $p \in (1, \infty)$ and Hilbert for $p = 2$ (see, e.g., [52, Prop. 8.1], [158, Th. 2.1] or [31, Prop. 3.4]). We note that the choice $\rho(x) = \frac{c_{n,s}}{|x|^{n+s-1}}$ for $s \in (0, 1)$ and $p \in (1, \infty)$ gives rise to the usual Bessel-potential space $H^{s,p}(\mathbb{R}^n)$, see e.g., [193, Th. 1.7] and [54, Th. A.1]. Moreover, we deduce immediately from the definition that $H_0^{\rho,p}(\mathbb{R}^n) = H^{\rho,p}(\mathbb{R}^n)$.

The following result shows the embedding of classical Sobolev spaces into $H^{\rho,p}(\mathbb{R}^n)$.

Proposition 5.3.5. *Let $p \in [1, \infty]$. Assume ρ satisfies (H0). Then, $W^{1,p}(\mathbb{R}^n) \subset H^{\rho,p}(\mathbb{R}^n)$ with continuous embedding, i.e., there exists a constant $C = C(n, p, \rho) > 0$ such that for any $u \in W^{1,p}(\mathbb{R}^n)$,*

$$\|u\|_{H^{\rho,p}(\mathbb{R}^n)} \le C\|u\|_{W^{1,p}(\mathbb{R}^n)}.$$

Moreover, $D_\rho u = \mathcal{G}_\rho u$.

Proof. Let $u \in W^{1,p}(\mathbb{R}^n)$. In the case $p = \infty$, we notice that for any $x \in \mathbb{R}^n$,

$$
\begin{aligned}
|\mathcal{G}_\rho u(x)| &\leq \int_{B_1} \frac{|u(x+h)-u(x)|}{|h|}\rho(h)\,dh + \int_{B_1^c} \frac{|u(x+h)-u(x)|}{|h|}\rho(h)\,dh \\
&\leq \|\nabla u\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)} \int_{B_1} \rho(h)\,dh + 2\|u\|_{L^\infty(\mathbb{R}^n)} \int_{B_1^c} \frac{\rho(h)}{|h|}\,dh \leq C\|u\|_{W^{1,\infty}(\mathbb{R}^n)},
\end{aligned}
$$

where we have used (H0). Thus, $\|\mathcal{G}_\rho u\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)} \leq C\|u\|_{W^{1,\infty}(\mathbb{R}^n)}$.

For $p \in [1,\infty)$ we follow the proof of [29, Prop. 2.7], which we include here for the reader's convenience. Clearly,

$$
\|\mathcal{G}_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \leq \left\| \int_{B_1} \frac{|u(\cdot+h)-u(\cdot)|}{|h|}\rho(h)\,dh \right\|_{L^p(\mathbb{R}^n)} + \left\| \int_{B_1^c} \frac{|u(\cdot+h)-u(\cdot)|}{|h|}\rho(h)\,dh \right\|_{L^p(\mathbb{R}^n)}.
$$

Applying Minkowski's integral inequality to the first term on the right-hand side of this inequality yields

$$
\begin{aligned}
\left\| \int_{B_1} \frac{|u(\cdot+h)-u(\cdot)|}{|h|}\rho(h)\,dh \right\|_{L^p(\mathbb{R}^n)} &\leq \int_{B_1} \frac{\rho(h)}{|h|} \left(\int_{\mathbb{R}^n} |u(x+h)-u(x)|^p\,dx \right)^{\frac{1}{p}} dh \\
&\leq C\|\nabla u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \int_{B_1} \rho(h)\,dh \leq C\|\nabla u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)},
\end{aligned}
$$

where we have used (H0) and [52, Prop. 9.3]. For the second term, applying Fubini's theorem and (H0), we find

$$
\begin{aligned}
\left\| \int_{B_1^c} \frac{|u(\cdot+h)-u(\cdot)|}{|h|}\rho(h)\,dh \right\|_{L^p(\mathbb{R}^n)} &\leq \left\| \int_{B_1^c} |u(\cdot+h)| \frac{\rho(h)}{|h|}\,dh \right\|_{L^p(\mathbb{R}^n)} + \left\| \int_{B_1^c} |u(\cdot)| \frac{\rho(h)}{|h|}\,dh \right\|_{L^p(\mathbb{R}^n)} \\
&\leq 2\left(\int_{B_1^c} \frac{\rho(h)}{|h|}\,dh \right) \|u\|_{L^p(\mathbb{R}^n)} \leq C\|u\|_{L^p(\mathbb{R}^n)}.
\end{aligned}
$$

Consequently, $\|\mathcal{G}_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(\mathbb{R}^n)}$.

Notice that the previous arguments show that, in both cases $p = \infty$ and $p < \infty$, any $u \in W^{1,p}(\mathbb{R}^n)$ satisfies the hypothesis of Proposition 5.3.2, and, hence, $\mathcal{G}_\rho u = D_\rho u$. □

The next result shows that the gradient operator commutes with convolution, see also [129, Lemma 3.7].

Lemma 5.3.6. *Let $p \in [1,\infty]$ and let ρ satisfy (H0). Let $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $u \in H^{\rho,p}(\mathbb{R}^n)$. Then, $\varphi * u \in C^\infty(\mathbb{R}^n)$ and $D_\rho(\varphi * u) = \varphi * D_\rho u$.*

Proof. Let $\Phi \in C_c^\infty(\mathbb{R}^n,\mathbb{R}^n)$ and $x \in \mathbb{R}^n$. A straightforward calculation shows that

$$
\begin{aligned}
\varphi * \operatorname{div}_\rho \Phi(x) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \varphi(x-y) \frac{\Phi(y)-\Phi(z)}{|y-z|} \cdot \frac{y-z}{|y-z|}\rho(y-z)\,dz\,dy \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \varphi(z') \frac{\Phi(x-z')-\Phi(y'-z')}{|x-y'|} \cdot \frac{x-y'}{|x-y'|}\rho(x-y')\,dz'\,dy' \\
&= \operatorname{div}_\rho(\varphi * \Phi)(x),
\end{aligned}
$$

after applying the changes of variables $y' = x - y + z$ and $z' = x - y$, and having in mind that all integrals involved are absolutely convergent since $\Phi \in C_c^\infty(\mathbb{R}^n,\mathbb{R}^n)$ and ρ satisfies (H0).

Let $\tilde{\varphi}$ be the reflection of φ, i.e., $\tilde{\varphi}(x) = \varphi(-x)$. Now, for any $u \in H^{\rho,p}(\mathbb{R}^n)$, by standard properties of convolution, $\varphi * u \in C^\infty(\mathbb{R}^n)$, and, by the previous computation, the definition of the distributional nonlocal gradient and Fubini's theorem (in particular, [52, Prop. 4.16]), we have

$$\begin{aligned}\int_{\mathbb{R}^n} D_\rho(\varphi * u)(x) \cdot \Phi(x)\,dx &= -\int_{\mathbb{R}^n} \varphi * u(x) \operatorname{div}_\rho \Phi(x)\,dx \\ &= -\int_{\mathbb{R}^n} u(x)(\tilde{\varphi} * \operatorname{div}_\rho \Phi)(x)\,dx \\ &= -\int_{\mathbb{R}^n} u(x) \operatorname{div}_\rho(\tilde{\varphi} * \Phi)(x)\,dx \\ &= \int_{\mathbb{R}^n} D_\rho u(x) \cdot (\tilde{\varphi} * \Phi)(x)\,dx \\ &= \int_{\mathbb{R}^n} (\varphi * D_\rho u)(x) \cdot \Phi(x)\,dx,\end{aligned}$$

and having in mind that Φ is an arbitrary test function, this concludes the proof. □

As a consequence of Lemma 5.3.6 and the standard method of approximation by convolution (see, e.g., [149, Th. C.16]), we find the following result.

Proposition 5.3.7. *Let ρ satisfy (H0).*

(i) For $p \in [1,\infty)$, it holds that $C^\infty(\mathbb{R}^n) \cap H^{\rho,p}(\mathbb{R}^n)$ is dense in $H^{\rho,p}(\mathbb{R}^n)$.

(ii) For all $u \in H^{\rho,\infty}(\mathbb{R}^n)$, there exists a sequence $\{u_j\}_{j\in\mathbb{N}}$ in $C^\infty(\mathbb{R}^n) \cap H^{\rho,\infty}(\mathbb{R}^n)$ such that

$$\begin{aligned}&u_j \to u \text{ a.e.},\quad \mathcal{G}_\rho u_j \to D_\rho u \text{ a.e.},\\ &\|u_j\|_{L^\infty(\mathbb{R}^n)} \to \|u\|_{L^\infty(\mathbb{R}^n)} \quad\text{and}\quad \|\mathcal{G}_\rho u_j\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)} \to \|D_\rho u\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)} \qquad \text{as } j \to \infty.\end{aligned}$$

The next lemma is a Leibniz rule in this nonlocal context, which is of interest in own right, and also needed in the proof of Theorem 5.3.9.

Lemma 5.3.8. *Let $p \in [1,\infty]$ and let ρ satisfy (H0). Let $g \in H^{\rho,p}(\mathbb{R}^n)$ and $f \in C_c^\infty(\mathbb{R}^n)$. Then, $fg \in H^{\rho,p}(\mathbb{R}^n)$ and*

$$D_\rho(fg) = fD_\rho g + K_f(g), \tag{5.22}$$

with

$$K_f(g)(x) = \int_{\mathbb{R}^n} \frac{f(x)-f(y)}{|x-y|} g(y) \frac{x-y}{|x-y|} \rho(x-y)\,dy, \quad \text{a.e. } x \in \mathbb{R}^n.$$

Moreover, there exists $C = C(\rho) > 0$ such that

$$\|K_f(g)\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \le C\|f\|_{W^{1,\infty}(\mathbb{R}^n)}\|g\|_{L^p(\mathbb{R}^n)} \tag{5.23}$$

and, if in addition $\rho \in L^1(\mathbb{R}^n)$,

$$\|K_f(g)\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \le C\|\nabla f\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)}\|g\|_{L^p(\mathbb{R}^n)}. \tag{5.24}$$

Proof. Clearly, $fg \in L^p(\mathbb{R}^n)$. Let $\Phi \in C_c^\infty(\mathbb{R}^n,\mathbb{R}^n)$. It is immediate to check that for all $x \in \mathbb{R}^n$,

$$\operatorname{div}_\rho(f\Phi)(x) = f(x) \operatorname{div}_\rho \Phi(x) + \int_{\mathbb{R}^n} \frac{f(x)-f(y)}{|x-y|} \frac{x-y}{|x-y|} \rho(x-y) \cdot \Phi(y)\,dy.$$

Thus,

$$
\begin{aligned}
-\int_{\mathbb{R}^n} & f(x)g(x)\,\mathrm{div}_\rho\,\Phi(x)\,dx \\
&= -\int_{\mathbb{R}^n} g(x)\,\mathrm{div}_\rho(f\Phi)(x)\,dx + \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} g(x)\frac{f(x)-f(y)}{|x-y|}\frac{x-y}{|x-y|}\rho(x-y)\cdot\Phi(y)\,dy\,dx \\
&= \int_{\mathbb{R}^n} f(x)D_\rho g(x)\cdot\Phi(x)\,dx + \int_{\mathbb{R}^n} K_f(g)(x)\cdot\Phi(x)\,dx,
\end{aligned}
$$

which shows (5.22). The bound (5.23) follows from Young's convolution inequality by using (H0) and the estimate

$$
\left|\frac{f(x)-f(y)}{|x-y|}\frac{x-y}{|x-y|}\rho(x-y)\right| \le C\|f\|_{W^{1,\infty}(\mathbb{R}^n)}\min\{1,|x-y|^{-1}\}\rho(x-y),
$$

while (5.24) uses

$$
\left|\frac{f(x)-f(y)}{|x-y|}\frac{x-y}{|x-y|}\rho(x-y)\right| \le C\|\nabla f\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)}\rho(x-y).
$$

Therefore, $D_\rho(fg) \in L^p(\mathbb{R}^n)$ and $fg \in H^{\rho,p}(\mathbb{R}^n)$. □

The following result explores the density of C_c^∞ functions, whose proof utilizes well-known mollification and cut-off arguments, see also [72, Th. 1] and [129, Th. 3.3].

Theorem 5.3.9. *Let ρ satisfy (H0).*

(i) Let $1 \le p < \infty$. Then, $C_c^\infty(\mathbb{R}^n)$ is dense in $H^{\rho,p}(\mathbb{R}^n)$.

(ii) For all $u \in H^{\rho,\infty}(\mathbb{R}^n)$, there exists a sequence $\{u_j\}_{j\in\mathbb{N}}$ in $C_c^\infty(\mathbb{R}^n)$ such that

$$
\begin{aligned}
&u_j \to u \text{ a.e.}, \quad \mathcal{G}_\rho u_j \to D_\rho u \text{ a.e.}, \\
&\|u_j\|_{L^\infty(\mathbb{R}^n)} \to \|u\|_{L^\infty(\mathbb{R}^n)} \quad \text{and} \quad \|\mathcal{G}_\rho u_j\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)} \to \|D_\rho u\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)} \qquad \text{as } j\to\infty.
\end{aligned}
$$

(iii) Let $1 \le p < \infty$ and let $\Omega \subset \mathbb{R}^n$ be an open and bounded set with a Lipschitz boundary. Then, $C_c^\infty(\Omega)$ is dense in $H_0^{\rho,p}(\Omega)$.

Proof. Due to Proposition 5.3.10 and Example 5.3.11 (a) below, we assume without loss of generality that $\rho \in L^1(\mathbb{R}^n)$. Parts (i) and (ii) can be proved as in [72, Th. 1]. Indeed, consider $\chi \in C_c^\infty(\mathbb{R}^n)$ with $0 \le \chi \le 1$ everywhere, $\chi = 1$ in B_1, and define $\chi_k := \chi(\frac{\cdot}{k})$ for $k \in \mathbb{N}$. Thanks to Proposition 5.3.7, it is enough to construct an approximating sequence for $u \in C^\infty(\mathbb{R}^n) \cap H^{\rho,p}(\mathbb{R}^n)$. Then, the sequence $\{\chi_k u\}_{k\in\mathbb{N}}$ is in $C_c^\infty(\mathbb{R}^n)$. Moreover, for $p \in [1,\infty)$ we have $\chi_k u \to u$ in $L^p(\mathbb{R}^n)$ as $k \to \infty$ and, by Lemma 5.3.8,

$$
\|D_\rho u - D_\rho(\chi_k u)\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \le \|(1-\chi_k)D_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} + C\frac{\|\nabla\chi\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)}}{k}\|u\|_{L^p(\mathbb{R}^n)} \to 0 \text{ as } k\to\infty,
$$

which yields (i). For $p = \infty$, we have $|\chi_k u| \le |u|$ and $\chi_k u \to u$ a.e. In addition, by Lemma 5.3.8, $D_\rho(\chi_k u) = \chi_k D_\rho u + K_{\chi_k}(u)$. Because $|\chi_k D_\rho u| \le |D_\rho u|$ on $\mathbb{R}^n$, $\chi_k D_\rho u \to D_\rho u$ a.e., and

$$
\|K_{\chi_k}(u)\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)} \le \frac{C}{k}\|\nabla\chi\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)}\|u\|_{L^\infty(\mathbb{R}^n)} \to 0 \quad \text{as } k\to\infty,
$$

we find that (ii) holds.

Part (iii). We first show that for each $u \in H_0^{\rho,p}(\Omega)$ and $\varepsilon > 0$, there exists a $\tilde{u} \in H^{\rho,p}(\mathbb{R}^n)$ with

$$\operatorname{supp} \tilde{u} \subset \Omega \quad \text{and} \quad \|u - \tilde{u}\|_{H^{\rho,p}(\mathbb{R}^n)} \leq \frac{\varepsilon}{2}. \tag{5.25}$$

To this aim, we may use the fact that the boundary of Ω is Lipschitz, to find a partition of unity $\chi_0, \ldots, \chi_N \in C_c^\infty(\mathbb{R}^n)$ subject to Ω and vectors $\zeta_1, \ldots, \zeta_N \in \mathbb{R}^n$ such that

$$\sum_{i=0}^N \chi_i = 1 \text{ on } \Omega, \quad \chi_0 \in C_c^\infty(\Omega), \tag{5.26}$$

and

$$(\operatorname{supp} \chi_i \cap \overline{\Omega}) + \lambda\zeta_i \subset \Omega \quad \text{for all } i = 1, \ldots, N \text{ and } \lambda > 0 \text{ small enough.} \tag{5.27}$$

For such λ, we define the function

$$\tilde{u} := \chi_0 u + \sum_{i=1}^N \tau_{\lambda\zeta_i}(\chi_i u),$$

where $\tau_\zeta(v) := v(\cdot - \zeta)$ for $v : \mathbb{R}^n \to \mathbb{R}$ denotes translation by the vector $\zeta \in \mathbb{R}^n$. In view of the Leibniz rule from Lemma 5.3.8 and the translation invariance of D_ρ, we deduce that $\tilde{u} \in H^{\rho,p}(\mathbb{R}^n)$. Moreover, due to (5.27), we find that $\operatorname{supp} \tilde{u} \subset \Omega$, which guarantees the first condition in (5.25).

For the norm estimate, we may use the L^p-continuity of the translation operator and the translation invariance of D_ρ, to find a $\lambda = \lambda_\varepsilon$ with

$$\|u - \tilde{u}\|_{H^{\rho,p}(\mathbb{R}^n)}^p = \left\|\sum_{i=1}^N \chi_i u - \tau_{\lambda_\varepsilon\zeta_i}(\chi_i u)\right\|_{L^p(\mathbb{R}^n)}^p + \left\|\sum_{i=1}^N D_\rho(\chi_i u) - \tau_{\lambda_\varepsilon\zeta_i} D_\rho(\chi_i u)\right\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)}^p \leq \left(\frac{\varepsilon}{2}\right)^p,$$

where we have used $u = \sum_{i=0}^N \chi_i u$ in light of the first part of (5.26) and the fact that $u \in H_0^{\rho,p}(\Omega)$. This proves (5.25). By mollifying $\tilde{u}$ suitably, we can find a $\varphi \in C_c^\infty(\Omega)$ such that $\|\tilde{u} - \varphi\|_{H^{\rho,p}(\mathbb{R}^n)} \leq \varepsilon/2$, which yields

$$\|u - \varphi\|_{H^{\rho,p}(\mathbb{R}^n)} \leq \|u - \tilde{u}\|_{H^{\rho,p}(\mathbb{R}^n)} + \|\tilde{u} - \varphi\|_{H^{\rho,p}(\mathbb{R}^n)} \leq \varepsilon,$$

and finishes the proof. □

5.3.2 Equivalence of spaces with different kernels

Here, we provide a sufficient condition so that two kernels give rise to the same space. This condition describes that the two kernels behave similarly at the origin. Moreover, one can carry over Poincaré inequalities from one gradient to the other.

Proposition 5.3.10. *Let $\Omega \subset \mathbb{R}^n$ be open. Let ρ_1, ρ_2 satisfy (H0) and assume that $(\rho_1 - \rho_2)/|\cdot| \in L^1(\mathbb{R}^n)$. Then, the following two statements hold:*

(i) Let $p \in [1, \infty]$. The identity $H_0^{\rho_1,p}(\Omega) = H_0^{\rho_2,p}(\Omega)$ holds with equivalent norms.

(ii) Let $p \in (1, \infty)$ and assume that Ω is bounded. If there is a $C_1 > 0$ such that for all $u \in H_0^{\rho_1,p}(\Omega)$,

$$\|u\|_{L^p(\Omega)} \leq C_1 \|D_{\rho_1} u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)}, \tag{5.28}$$

then there is a $C_2 > 0$ such that for all $u \in H_0^{\rho_2,p}(\Omega)$,

$$\|u\|_{L^p(\Omega)} \leq C_2 \|D_{\rho_2} u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)}.$$

Proof. Part (i). Set

$$F(x) = \frac{x}{|x|}\frac{\rho_2(x) - \rho_1(x)}{|x|}.$$

Then, we find that $F \in L^1(\mathbb{R}^n, \mathbb{R}^n)$ and, for all $\Phi \in C_c^\infty(\mathbb{R}^n, \mathbb{R}^n)$,

$$\operatorname{div}_{\rho_1} \Phi = \operatorname{div}_{\rho_2} \Phi + F * \Phi.$$

By Young's inequality, the operator $u \mapsto F * u$ is bounded on $L^p(\mathbb{R}^n)$. Hence, for $u \in H_0^{\rho_2,p}(\Omega)$ and $\Phi \in C_c^\infty(\mathbb{R}^n, \mathbb{R}^n)$ it holds that

$$\int_{\mathbb{R}^n} (D_{\rho_2} u + F * u) \cdot \Phi \, dx = -\int_{\mathbb{R}^n} u \, (\operatorname{div}_{\rho_2} \Phi + F * \Phi) \, dx = -\int_{\mathbb{R}^n} u \operatorname{div}_{\rho_1} \Phi \, dx,$$

where the first equality uses Fubini's theorem and the definition of the weak nonlocal gradient D_{ρ_2}. Therefore, $u \in H_0^{\rho_1,p}(\Omega)$,

$$D_{\rho_1} u = D_{\rho_2} u + F * u \tag{5.29}$$

and

$$\|D_{\rho_1} u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \leq \|D_{\rho_2} u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} + \|F\|_{L^1(\mathbb{R}^n,\mathbb{R}^n)} \|u\|_{L^p(\mathbb{R}^n)}.$$

The reverse inclusion and inequality are proved analogously.

Part (ii). We first prove that if $u \in C_c^\infty(\mathbb{R}^n)$ satisfies $\mathcal{G}_{\rho_2} u = 0$, then $u = 0$. By Lemma 5.2.4,

$$0 = \widehat{\mathcal{G}_{\rho_2} u}(\xi) = \lambda_{\rho_2}(\xi)\widehat{u}(\xi) \quad \text{for all } \xi \in \mathbb{R}^n.$$

If u were not identically zero, then $\widehat{u}$ would be a non-zero analytic function, and hence, non-zero in a set of full measure. As such, we deduce that $\lambda_{\rho_2} = 0$ a.e. A further application of Lemma 5.2.4 to any $\varphi \in C_c^\infty(\mathbb{R}^n)$ implies that $\widehat{\mathcal{G}_{\rho_2}\varphi} = 0$ a.e., hence $\mathcal{G}_{\rho_2}\varphi = 0$ a.e., and, in fact, everywhere thanks to Proposition 5.2.6 (i). On the other hand, choosing a φ such that $\varphi(y) = -\varphi(-y)$ for all $y \in \mathbb{R}^n$, $\varphi \geq 0$ in $\{y_1 > 0\}$ and φ is not identically zero on $B_\varepsilon \cap \{y_1 > 0\}$ yields

$$\mathcal{G}_{\rho_2}\varphi(0) = \int_{\{y_1>0\}} \frac{2\varphi(y)}{|y|} \frac{y}{|y|} \rho_2(y) \, dy > 0,$$

cf. (H0). This contradiction concludes that u must be zero, which proves the claim.

Now we show that if $u \in H_0^{\rho_2,p}(\Omega)$ satisfies $D_{\rho_2} u = 0$, then $u = 0$. Let $\varphi \in C_c^\infty(\mathbb{R}^n)$. By Lemma 5.3.6 and the fact that Ω is bounded, $\varphi * u \in C_c^\infty(\mathbb{R}^n)$ and

$$\mathcal{G}_{\rho_2}(\varphi * u) = \varphi * D_{\rho_2} u = 0.$$

By the claim above, $\varphi * u = 0$. As this is true for all $\varphi \in C_c^\infty(\mathbb{R}^n)$, we conclude, by taking φ to be a family of mollifiers, that $u = 0$.

Now, to prove the statement, we argue by contradiction. Suppose $(u_j)_j \subset H_0^{\rho_2,p}(\Omega)$ is a sequence satisfying

$$1 = \|u_j\|_{L^p(\Omega)} > j\|D_{\rho_2} u_j\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \quad \text{for all } j \in \mathbb{N}.$$

As $H_0^{\rho_2,p}(\Omega)$ is reflexive, there exists $u \in H_0^{\rho_2,p}(\Omega)$ such that, up to subsequence, $u_j \rightharpoonup u$ in $H_0^{\rho_2,p}(\Omega)$. As $D_{\rho_2} u_j \to 0$ in $L^p(\mathbb{R}^n, \mathbb{R}^n)$, we obtain that in fact $D_{\rho_2} u = 0$, which implies $u = 0$ thanks to the result of the previous paragraph.

Since for all large $R > 0$, it holds that $\operatorname{supp} u_j \subset B_R$ for all $j \in \mathbb{N}$, we find that

$$\|F * u_j\|_{L^p(B_{2R}^c,\mathbb{R}^n)} = \|(\mathbb{1}_{B_R^c} F) * u_j\|_{L^p(B_{2R}^c,\mathbb{R}^n)} \leq \|F\|_{L^1(B_R^c,\mathbb{R}^n)} \to 0,$$

as $R \to \infty$. Together with the compactness of convolution operators on bounded sets due to the Fréchet-Kolmogorov criterion (cf. [52, Cor. 4.28]), we infer that $F * u_j \to 0$ in $L^p(\mathbb{R}^n)$, and hence, by (5.29)

$$D_{\rho_1} u_j = D_{\rho_2} u_j + F * u_j \to 0 \quad \text{in } L^p(\mathbb{R}^n, \mathbb{R}^n).$$

In view of (5.28), we deduce that $u_j \to 0$ in $L^p(\Omega)$ which contradicts $\|u_j\|_{L^p(\Omega)} = 1$. □

We note that in part (i), we can take $\Omega = \mathbb{R}^n$, which yields the correspondence $H^{\rho_1,p}(\mathbb{R}^n) = H^{\rho_2,p}(\mathbb{R}^n)$ with equivalent norms. Additionally, having in mind (5.10), we can see that the condition $(\rho_1 - \rho_2)/|\cdot| \in L^1(\mathbb{R}^n)$ can be equivalently written as

$$\int_0^r (\overline{\rho_1}(t) - \overline{\rho_2}(t))\, t^{n-2}\, dt < \infty, \quad \text{for some } r > 0.$$

Example 5.3.11. We present two applications of Proposition 5.3.10.

(a) Let ρ satisfy (H0) and let $\chi \in L^\infty(\mathbb{R}^n)$ be radial with $\chi \geq 0$ and $\chi|_{B_r} = 1$ for some $r > 0$. Then $\chi\rho$ satisfies (H0) and $H_0^{\rho,p}(\Omega) = H_0^{\chi\rho,p}(\Omega)$ for all $p \in [1, \infty]$ and any open set $\Omega \subset \mathbb{R}^n$. The assumption $\chi|_{B_r} = 1$ can be weakened to $\inf_{B_r} \chi > 0$ and

$$\int_{B_r} \frac{\chi(x) - \chi(0)}{|x|} \rho(x)\, dx < \infty.$$

Moreover, if $p \in (1, \infty)$ and Ω is bounded, any Poincaré inequality for D_ρ implies an analogous one for $D_{\chi\rho}$. We conclude that we can associate to every kernel satisfying (H0) a kernel satisfying (H0) with compact support and giving rise to the same space. In the development of the theory we will often require that the kernel has compact support. Thanks to this observation, this can be assumed without loss of generality.

(b) Let $s \in (0, 1)$. Let $\chi \in L^\infty(\mathbb{R}^n)$ be radial with $\chi \geq 0$ and $(\chi(0) - \chi)/|\cdot|^{n+s} \in L^1(\mathbb{R}^n)$; the latter condition happens, for example, if χ is γ-Hölder continuous at 0 with $\gamma > s$. Then the kernel

$$\rho(x) = \frac{\chi(x)}{|x|^{n+s-1}}$$

can be compared with the kernel (5.12) of the Riesz s-fractional gradient. Thus, $H_0^{s,p}(\Omega) = H_0^{\rho,p}(\Omega)$ for any open $\Omega \subset \mathbb{R}^n$ and $p \in [1, \infty]$. If, in addition, Ω is bounded and $p \in (1, \infty)$, we have that $\|u\|_{L^p(\Omega)} \leq C\|D_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)}$ for all $u \in H_0^{\rho,p}(\Omega)$, as a consequence of the corresponding inequality for the Bessel-potential space $H^{s,p}$ (see [193, Th. 1.8]). This constitutes an alternative proof, as well as a generalization of the Poincaré inequality for truncated fractional gradients of [31, Th. 6.2].

5.4 Poincaré inequalities and compact embeddings

In this section, we derive conditions on the kernel ρ such that the ρ-nonlocal gradient satisfies a Poincaré inequality, and such that the spaces $H_0^{\rho,p}(\Omega)$ are compactly embedded into L^p. The argument is based on inverting the nonlocal gradient via (5.20), and showing with Fourier techniques that this is a bounded or compact operation. The results in the case $p = 2$ are tackled first and rely on Parseval's identity. Subsequently, the general case $p \in (1, \infty)$ is considered, which requires an additional assumption in order to apply the Mihlin-Hörmander multiplier theorem.

5.4.1 Positivity of $\widehat{Q}_\rho$

We show in this section that $\widehat{Q}_\rho > 0$, as a first step to make sense of the expression (5.20). We also relate the decay of $\widehat{Q}_\rho$ at infinity with the behavior of ρ around the origin. For this, we need the following assumption:

(H1) The function $f_\rho : (0,\infty) \to \mathbb{R},\ t \mapsto t^{n-2}\overline{\rho}(t)$ is decreasing, and there is a $0 < \mu < 1$ such that $\mu f_\rho(t/2) \geq f_\rho(t)$ for $t \in (0,\varepsilon)$.

Remark 5.4.1. We note that the second part (the doubling property) of (H1) is satisfied if there is a $\nu > 0$ such that $t \mapsto t^\nu f_\rho(t)$ is decreasing on $(0,\varepsilon)$. The constant μ is then given by $2^{-\nu}$.

Additionally, if f_ρ is differentiable, then a simple calculation with the product rule shows that $t \mapsto t^\nu f_\rho(t)$ is decreasing on $(0,\varepsilon)$ if and only if

$$-\frac{1}{\nu}\frac{d}{dt} f_\rho(t) \geq \frac{f_\rho(t)}{t} \quad \text{for all } t \in (0,\varepsilon).$$

The latter condition will appear again in (H2), and hence, implies the doubling property of (H1). △

Example 5.4.2. Classes of kernels ρ satisfying (H0)–(H1) are:

(a) ρ of Example 5.2.1 (a).

(b) Given $n-2 < \alpha < n$ and $\beta > n-1$,

$$\rho(x) = \frac{\mathbb{1}_{B_1}(x)}{|x|^\alpha} + \frac{\mathbb{1}_{B_1^c}(x)}{|x|^\beta}.$$

(c) Given $n-2 < \alpha < n$,

$$\rho(x) = \mathbb{1}_{B_1}(x)\frac{-\log|x|}{|x|^\alpha},$$

(d) Given $n-2 < \alpha < n$ and $r > 0$ sufficiently small,

$$\rho(x) = \frac{\mathbb{1}_{B_r(x)}}{|x|^\alpha(-\log|x|)},$$

(e) If ρ_1, ρ_2 satisfy (H0)–(H1) and $\alpha_1, \alpha_2 > 0$ then $\alpha_1\rho_1 + \alpha_2\rho_2$ satisfies (H0)–(H1).

We may now state the following result, whose proof takes inspiration from [31, Lemma 5.3].

Lemma 5.4.3. *Suppose that ρ satisfies (H0), (H1) and $\rho \in L^1(\mathbb{R}^n)$. Then, $\widehat{Q}_\rho$ is positive and there is a $C > 0$ such that*

$$\widehat{Q}_\rho(\xi) \geq C\frac{\overline{\rho}(1/|\xi|)}{|\xi|^n} \quad \text{and} \quad \widehat{Q}_\rho(\xi) \geq \frac{C}{|\xi|^2} \qquad \text{for all } \xi \in B^c_{2/\varepsilon}.$$

Proof. Note that since $\widehat{Q}_\rho$ is continuous (by Lemma 5.2.5 (iii)) and $\widehat{Q}_\rho(0) = \|Q_\rho\|_{L^1(\mathbb{R}^n)} > 0$, we have that $\widehat{Q}_\rho$ is positive around the origin. Next, for $\xi \neq 0$ we obtain from (5.16) and the coarea formula that

$$\begin{aligned}\widehat{Q}_\rho(\xi) &= \frac{1}{\pi|\xi|}\int_{\mathbb{S}^{n-1}_+} z_1 \int_0^\infty r^{n-2}\overline{\rho}(r)\sin(2\pi|\xi|rz_1)\,dr\,d\mathcal{H}^{n-1}(z) \\ &= \frac{1}{\pi|\xi|}\int_{\mathbb{S}^{n-1}_+} z_1 \int_0^\infty f_\rho(r)\sin(2\pi|\xi|rz_1)\,dr\,d\mathcal{H}^{n-1}(z).\end{aligned} \tag{5.30}$$

Let $\theta > 0$, which will play the role of $|\xi| z_1$. We have

$$\int_0^\infty f_\rho(r) \sin(2\pi\theta r)\, dr = \sum_{k=0}^\infty \int_{\frac{k}{\theta}}^{\frac{k+1}{\theta}} f_\rho(r) \sin(2\pi\theta r)\, dr \tag{5.31}$$

and for each $k \in \mathbb{N}$,

$$\int_{\frac{k}{\theta}}^{\frac{k+1}{\theta}} f_\rho(r) \sin(2\pi\theta r)\, dr = \int_{\frac{k}{\theta}}^{\frac{k+\frac{1}{2}}{\theta}} \left(f_\rho(r) - f_\rho\left(r + \frac{1}{2\theta}\right)\right) \sin(2\pi\theta r)\, dr \geq 0, \tag{5.32}$$

since f_ρ is decreasing. Moreover,

$$\int_0^{\frac{1}{2\theta}} \left(f_\rho(r) - f_\rho\left(r + \frac{1}{2\theta}\right)\right) \sin(2\pi\theta r)\, dr > 0,$$

since otherwise f_ρ would be constant near zero, contradicting the doubling property in (H1). This shows that $\widehat{Q}_\rho(\xi) > 0$.

Now, for $\theta > 1/\varepsilon$ and $0 < r < \frac{1}{2\theta}$, we have

$$f_\rho(r) - f_\rho\left(r + \frac{1}{2\theta}\right) \geq f_\rho(r) - f_\rho(2r) \geq (1-\mu)\, f_\rho(r) \geq (1-\mu)\, f_\rho\left(\frac{1}{2\theta}\right),$$

where we have used that f_ρ is decreasing, as well as the doubling condition in (H1). Therefore,

$$\int_0^{\frac{1}{2\theta}} \left(f_\rho(r) - f_\rho\left(r + \frac{1}{2\theta}\right)\right) \sin(2\pi\theta r)\, dr \geq (1-\mu)\, f_\rho\left(\frac{1}{2\theta}\right) \int_0^{\frac{1}{2\theta}} \sin(2\pi\theta r)\, dr = \frac{1-\mu}{\pi\theta} f_\rho\left(\frac{1}{2\theta}\right).$$

The above inequality together with (5.32) and (5.30) show that for all $|\xi| > 2/\varepsilon$,

$$\begin{aligned} \widehat{Q}_\rho(\xi) &\geq \frac{1}{2\pi|\xi|} \int_{\mathbb{S}^{n-1}\cap\{z_1 > 1/2\}} \int_0^\infty f_\rho(r) \sin(2\pi|\xi| r z_1)\, dr\, d\mathcal{H}^{n-1}(z) \\ &\geq \frac{1-\mu}{2\pi^2|\xi|^2} \int_{\mathbb{S}^{n-1}\cap\{z_1 > 1/2\}} \frac{1}{z_1} f_\rho\left(\frac{1}{2|\xi| z_1}\right) d\mathcal{H}^{n-1}(z) \geq \frac{C}{|\xi|^2} f_\rho\left(\frac{1}{|\xi|}\right), \end{aligned}$$

where in the last inequality we have used that f_ρ is decreasing. The constant C depends on μ and n. The proof is concluded thanks to the definition of f_ρ, as well as the fact that $f_\rho\left(\frac{1}{|\xi|}\right) \geq f_\rho\left(\frac{\varepsilon}{2}\right)$. □

The following proposition constitutes a further step to prove formula (5.20).

Proposition 5.4.4. *Suppose that ρ satisfies (H0), (H1) and $\rho \in L^1(\mathbb{R}^n)$. Then W_ρ given by*

$$\langle W_\rho, \eta\rangle = \lim_{r\downarrow 0} \int_{B_r^c} \frac{-i\xi}{2\pi|\xi|^2 \widehat{Q}_\rho(\xi)} \eta(\xi)\, d\xi, \qquad \text{for } \eta \in \mathcal{S}(\mathbb{R}^n), \tag{5.33}$$

defines a tempered distribution. Moreover, if ρ has compact support, it holds that

$$W_\rho \cdot \widehat{\mathcal{G}_\rho \varphi} = \widehat{\varphi} \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^n). \tag{5.34}$$

Proof. The integrand of (5.33) is well defined since $\widehat{Q}_\rho$ is positive (Lemma 5.4.3). Let us see that the integral is absolutely convergent for each $r > 0$, and for this we can assume that $r < 2/\varepsilon$. For $|\xi| \geq 2/\varepsilon$ we have that

$$\left|\frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)}\eta(\xi)\right| \leq \frac{|\eta(\xi)|}{2\pi|\xi|\widehat{Q}_\rho(\xi)} \leq \frac{|\eta(\xi)||\xi|}{2\pi C},$$

thanks to Lemma 5.4.3, and the right-hand side is integrable in $B^c_{2/\varepsilon}$ since $\eta \in \mathcal{S}(\mathbb{R}^n)$. For $r < |\xi| < 2/\varepsilon$ we have that

$$\left|\frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)}\eta(\xi)\right| \leq \frac{\|\eta\|_\infty}{2\pi|\xi|\widehat{Q}_\rho(\xi)},$$

and the right-hand side is integrable in $B_{2/\varepsilon} \setminus B_r$ since $\widehat{Q}_\rho$ is positive (Lemma 5.4.3) and continuous (Lemma 5.2.5 (iii)).

Now, by symmetry (as $\widehat{Q}_\rho$ is radial; cf. (5.16)),

$$\int_{B_{2/\varepsilon}\setminus B^c_r} \frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)}\eta(\xi)\, d\xi = \int_{B_{2/\varepsilon}\setminus B^c_r} \frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)}\,(\eta(\xi) - \eta(0))\, d\xi,$$

with

$$\left|\frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)}\,(\eta(\xi) - \eta(0))\right| \leq \frac{\|\nabla\eta\|_\infty}{2\pi\widehat{Q}_\rho(\xi)},$$

and the right-hand side is integrable in $B_{2/\varepsilon}$ since $\widehat{Q}_\rho$ is positive and continuous. Therefore, the limit of (5.33) exists and defines a tempered distribution.

For the final part, we note that if ρ has compact support, then $\mathcal{G}_\rho\varphi \in C^\infty_c(\mathbb{R}^n, \mathbb{R}^n)$ for all $\varphi \in C^\infty_c(\mathbb{R}^n)$ by Proposition 5.2.6 (i) and hence, $\widehat{\mathcal{G}_\rho\varphi} \in \mathcal{S}(\mathbb{R}^n, \mathbb{C}^n)$. With this and Proposition 5.2.6 (ii), we obtain that for all $\eta \in \mathcal{S}(\mathbb{R}^n)$,

$$\left\langle W_\rho \cdot \widehat{\mathcal{G}_\rho\varphi}, \eta\right\rangle = \left\langle W_\rho, \widehat{\mathcal{G}_\rho\varphi}\,\eta\right\rangle = \lim_{r\downarrow 0} \int_{B^c_r} \frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)} \cdot 2\pi i\xi\widehat{Q}_\rho(\xi)\widehat{\varphi}(\xi)\eta(\xi)\, d\xi = \int_{\mathbb{R}^n} \widehat{\varphi}(\xi)\eta(\xi)\, d\xi.$$

This proves that $W_\rho \cdot \widehat{\mathcal{G}_\rho\varphi} = \widehat{\varphi}$, as desired. □

Note that, when $n > 1$, the distribution W_ρ actually agrees with the locally integrable function

$$W_\rho(\xi) = \frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)}.$$

5.4.2 Poincaré inequality and Compactness in L^2

The bounds in Lemma 5.4.3 allow us to swiftly prove a Poincaré inequality and compactness result in the L^2-setting, by prescribing that ρ is of compact support and satisfies certain bounds.

Recall from Section 5.2 that if $\operatorname{supp}\rho = \overline{B_\delta}$ for some $\delta > 0$, given an open set $\Omega \subset \mathbb{R}^n$ and $\varphi \in C^\infty_c(\Omega)$, the function $\mathcal{G}_\rho\varphi$ is supported in $\Omega_\delta = \Omega + B_\delta$.

Theorem 5.4.5. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded and suppose that ρ satisfies (H0), (H1) and has compact support. Then, the following two statements hold:*

(i) If $\liminf_{t\downarrow 0} t^{n-1}\overline{\rho}(t) > 0$, then there is a $C = C(\Omega, n, \rho) > 0$ such that

$$\|u\|_{L^2(\Omega)} \leq C\|D_\rho u\|_{L^2(\mathbb{R}^n,\mathbb{R}^n)} \quad \textit{for all } u \in H^\rho_0(\Omega).$$

(ii) If $\lim_{t\downarrow 0} t^{n-1}\overline{\rho}(t) = \infty$*, then* $H_0^\rho(\Omega)$ *is compactly embedded into* $L^2(\mathbb{R}^n)$*.*

Proof. Note first that as ρ has compact support, $\rho \in L^1(\mathbb{R}^n)$, and, hence, Proposition 5.4.4 can be applied.

Part (i). Let $W_\rho \in \mathcal{S}'(\mathbb{R}^n, \mathbb{C}^n)$ be as in (5.33) and let $\chi \in C_c^\infty(\mathbb{R}^n)$ be radial with $\chi = 1$ on B_1. Then, we set

$$L = \left(\chi W_\rho\right)^\vee \in C^\infty(\mathbb{R}^n, \mathbb{R}^n) \quad \text{and} \quad M = (1-\chi)W_\rho.$$

We specify the above definitions. Naturally, $\chi W_\rho \in \mathcal{S}'(\mathbb{R}^n, \mathbb{C}^n)$ is the distribution defined as

$$\langle \chi W_\rho, \eta\rangle = \lim_{r\downarrow 0}\int_{B_r^c} \chi(\xi)\frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)}\eta(\xi)\,d\xi, \qquad \eta \in \mathcal{S}(\mathbb{R}^n);$$

cf. (5.33). As χ has compact support, so does χW_ρ, and, hence, by the Paley–Wiener theorem, L is analytic. Likewise, $M \in \mathcal{S}'(\mathbb{R}^n, \mathbb{C}^n)$ is the distribution defined as

$$\langle M, \eta\rangle = \int_{\mathbb{R}^n} (1-\chi(\xi))\,\frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)}\eta(\xi)\,d\xi, \qquad \eta \in \mathcal{S}(\mathbb{R}^n),$$

so M can be identified with the function

$$M(\xi) = (1-\chi(\xi))\,\frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)}, \qquad \xi \in \mathbb{R}^n. \tag{5.35}$$

Moreover, M is smooth (by Lemma 5.2.5 (iii), as Q_ρ has compact support) and bounded by Lemma 5.4.3 and the fact

$$\limsup_{t\uparrow\infty} \frac{t^{n-1}}{\overline{\rho}(1/t)} < \infty, \tag{5.36}$$

which is a consequence of the assumption in (i). Therefore, we may define the bounded operator

$$T_M : L^2(\mathbb{R}^n, \mathbb{R}^n) \to L^2(\mathbb{R}^n), \quad U \mapsto \left(M\cdot\widehat{U}\right)^\vee. \tag{5.37}$$

On the other hand, using Proposition 5.4.4 we have that for all $\varphi \in C_c^\infty(\mathbb{R}^n)$,

$$\widehat{\varphi} = W_\rho \cdot \widehat{\mathcal{G}_\rho\varphi} = (\chi W_\rho)\cdot\widehat{\mathcal{G}_\rho\varphi} + M\cdot\widehat{\mathcal{G}_\rho\varphi},$$

whence taking the inverse Fourier transform, we obtain

$$\varphi = L * \mathcal{G}_\rho\varphi + T_M(\mathcal{G}_\rho\varphi).$$

By Theorem 5.3.9 (i), this identity can be extended to

$$u = L * D_\rho u + T_M(D_\rho u) \ \text{ for all } u \in H_0^\rho(\Omega). \tag{5.38}$$

Therefore, part (i) follows with

$$C = \|L\|_{L^1(\Omega-\Omega_\delta,\mathbb{R}^n)} + \|M\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^n)}.$$

Part (ii). We note that the operator $U \mapsto L * U$ is compact from $L^2(\Omega_\delta, \mathbb{R}^n)$ to $L^2(\Omega)$, since L is locally bounded (see, e.g., [70, Prop. 4.7]); the precise definition of $L * U$ is

$$L * U(x) = \int_{\Omega_\delta} L(x-y)\cdot U(y)\,dy, \qquad x \in \Omega.$$

Moreover, as a consequence of the assumption in (ii) we have

$$\lim_{t\uparrow\infty} \frac{t^{n-1}}{\overline{\rho}(1/t)} = 0,$$

which implies that

$$|M(\xi)| \to 0 \text{ as } |\xi| \to \infty$$

thanks to Lemma 5.4.3. Therefore, for any $U \in L^2(\mathbb{R}^n, \mathbb{R}^n)$ with $\|U\|^2_{L^2(\mathbb{R}^n,\mathbb{R}^n)} \le 1$, it holds that

$$\lim_{R\to\infty} \int_{B_R^c} |\widehat{T_M U}|^2 \, d\xi \le \lim_{R\to\infty} \sup_{|\xi|\ge R} |M(\xi)|^2 = 0.$$

By a version of the Fréchet-Kolmogorov criterion in Fourier space, cf. [175, Th. 3], we deduce that $T_M : L^2(\mathbb{R}^n, \mathbb{R}^n) \to L^2(\Omega)$ is compact.

The compact embedding of $H_0^\rho(\Omega)$ into $L^2(\mathbb{R}^n)$ (or, equivalently, $L^2(\Omega)$) is concluded thanks to identity (5.38) and the fact that both operators $L * \cdot$ and T_M are compact from $L^2(\Omega_\delta, \mathbb{R}^n)$ to $L^2(\Omega)$, where in the case of T_M it is understood that $L^2(\Omega_\delta, \mathbb{R}^n)$ is extended as zero to $L^2(\mathbb{R}^n, \mathbb{R}^n)$. □

Example 5.4.6. Classes of kernels of compact support satisfying (H0)–(H1), and $\lim_{t\downarrow 0} t^{n-1}\overline{\rho}(t) = \infty$ are:

(a) Given $0 < s < 1$,

$$\rho(x) = \frac{\mathbb{1}_{B_1}(x)}{|x|^{n+s-1}}.$$

(b) Given $0 \le s < 1$,

$$\rho(x) = \mathbb{1}_{B_1}(x) \frac{-\log|x|}{|x|^{n+s-1}}.$$

(c) Given $0 < s < 1$ and $r > 0$ sufficiently small,

$$\rho(x) = \frac{\mathbb{1}_{B_r(x)}}{|x|^{n+s-1}(-\log|x|)}.$$

Example 5.4.7. A kernel of compact support satisfying (H0)–(H1), and $\lim_{t\downarrow 0} t^{n-1}\overline{\rho}(t) > 0$ is

$$\rho(x) = \frac{\mathbb{1}_{B_1}(x)}{|x|^{n-1}}.$$

We will see in Proposition 5.7.5, that under some additional assumptions on ρ, the growth conditions at the origin of Theorem 5.4.5 are sharp in order to obtain the validity of a Poincaré inequality or a compact embedding, respectively.

5.4.3 Poincaré inequality and Compactness in L^p

We derive here an analogue of Theorem 5.4.5 in the L^p setting, by applying the Mihlin-Hörmander theorem to show that the function M in the proof of Theorem 5.4.5 is an L^p multiplier. This requires us to study also the decay of the derivatives of $\widehat{Q}_\rho$, and we impose the following assumption for that:

(H2) The function f_ρ is smooth in $(0, \infty)$, and for $t \in (0, \varepsilon)$,

$$-C\frac{d}{dt} f_\rho(t) \ge \frac{f_\rho(t)}{t} \quad \text{and} \quad \left|\frac{d^k}{dt^k} f_\rho(t)\right| \le C_k \frac{f_\rho(t)}{t^k} \quad \text{for } k \in \mathbb{N}.$$

The first condition is equivalent to the property that $t \mapsto t^\nu f_\rho(t)$ is decreasing on $(0,\varepsilon)$ for some $\nu > 0$, cf. Remark 5.4.1, while the second imposes that f_ρ does not oscillate too much. A consequence of (H2) that will be repeatedly used is

$$\left|\frac{d^l}{dt^l}(t^k f_\rho(t))\right| \leq C_k t^{k-l} f_\rho(t), \qquad 0 \leq l \leq k, \quad t \in (0,\varepsilon), \tag{5.39}$$

as well as

$$\left|\frac{d^{k+1}}{dt^{k+1}}(t^k f_\rho(t))\right| \leq -C_k \frac{d}{dt} f_\rho(t), \qquad k \in \mathbb{N}, \quad t \in (0,\varepsilon). \tag{5.40}$$

The following consequence is also useful.

Lemma 5.4.8. *Let ρ satisfy (H0) and (H2). Then,* $\lim_{t\downarrow 0} t^2 f_\rho(t) = 0$.

Proof. As a consequence of (5.10), we have $\liminf_{t\downarrow 0} t^2 f_\rho(t) = 0$. Let $0 < t < \varepsilon$. By the fundamental theorem of calculus,

$$\varepsilon^2 f_\rho(\varepsilon) - t^2 f_\rho(t) = \int_t^\varepsilon \frac{d}{dr}(r^2 f_\rho(r))\, dr. \tag{5.41}$$

Now, by (5.39) and (5.10)

$$\int_0^\varepsilon \left|\frac{d}{dr}(r^2 f_\rho(r))\right| dr \leq C_2 \int_0^\varepsilon r f_\rho(r)\, dr < \infty.$$

Therefore, the limit when $t \downarrow 0$ of the right-hand side of (5.41) exists, and, consequently, so does the limit of the left-hand side, which proves the result. □

Example 5.4.9. Classes of kernels ρ satisfying (H0)–(H2) are:

(a) ρ of Example 5.2.1 (a).

(b) Given $0 \leq s < 1$ and a non-negative radial function $\chi \in C_c^\infty(\mathbb{R}^n)$ with $\chi(0) > 0$ and $\chi(x)/|x|^{1+s}$ radially decreasing,

$$\rho(x) = \frac{\chi(x)}{|x|^{n+s-1}}.$$

(c) Given $0 \leq s < 1$ and a non-negative, radial function $\chi \in C_c^\infty(B_1)$ with $\chi(0) > 0$ and $\chi(x)(-\log|x|)/|x|^{1+s}$ radially decreasing,

$$\rho(x) = \frac{\chi(x)(-\log|x|)}{|x|^{n+s-1}}.$$

Indeed, (H0) and (H1) are simple to verify, whereas (H2) follows since the derivatives of log behave similarly as a power function.

(d) Given $0 \leq s < 1$ and a non-negative, radial function $\chi \in C_c^\infty(B_1)$ with $\chi(0) > 0$ and $\chi(x)/(-|x|^{1+s}\log|x|)$ radially decreasing,

$$\rho(x) = \frac{\chi(x)}{|x|^{n+s-1}(-\log|x|)}.$$

The verification of (H0)–(H2) is similar to that of the previous example.

(e) Given a smooth $s : [0,\infty) \to (0,1)$ and a non-negative, radial function $\chi \in C_c^\infty(\mathbb{R}^n)$ with $\chi(0) > 0$ and $\chi(x)/|x|^{1+s(|x|)}$ radially decreasing,

$$\rho(x) = \frac{\chi(x)}{|x|^{n+s(|x|)-1}}.$$

Again, (H0) and (H1) follow readily, whereas for (H2) we first note that

$$\frac{d}{dt} f_\rho(t) = \overline{\chi}'(t)\frac{1}{t^{1+s(t)}} + \overline{\chi}(t)\left(\frac{-(1+s(t))}{t^{2+s(t)}} + \frac{-\log(t)s'(t)}{t^{1+s(t)}}\right),$$

which satisfies $|\frac{d}{dt} f_\rho(t)| \leq C f_\rho(t)/t$ since $t/\log t$ is locally bounded. The other derivatives can be bounded in a similar way, so (H2) holds.

(f) If ρ_1, ρ_2 satisfy (H0)–(H2) and $\alpha_1, \alpha_2 > 0$ then $\alpha_1\rho_1 + \alpha_2\rho_2$ satisfies (H0)–(H2).

Lemma 5.4.10. *Let ρ have compact support and satisfy (H0)–(H2). Then, for every $\alpha \in \mathbb{N}^n$,*

$$|\partial^\alpha \widehat{Q}_\rho(\xi)| \leq C_\alpha \left(|\xi|^{-|\alpha|}\widehat{Q}_\rho(\xi) + |\xi|^{-|\alpha|-1}\right), \qquad |\xi| \geq 1.$$

Proof. Step 1: Integral bounds. We show that for all $\theta > 1/\varepsilon$ and $k \in \mathbb{N}$,

$$\left|\int_0^\infty r^k f_\rho(r) a_k(2\pi\theta r)\,dr\right| \leq \frac{C_k}{\theta^k}\int_0^\infty f_\rho(r)\sin(2\pi\theta r)\,dr + \frac{C_k}{\theta^k}, \tag{5.42}$$

where $a_k = \cos$ when k is odd and $a_k = \sin$ when k is even.

Using integration by parts k times we obtain

$$\int_0^\infty r^k f_\rho(r) a_k(2\pi\theta r)\,dr = \frac{D_k}{\theta^k}\int_0^\infty \frac{d^k}{dr^k}(r^k f_\rho(r))\sin(2\pi\theta r)\,dr, \tag{5.43}$$

where the constants D_k may be negative. Indeed, equality (5.43) follows directly by integration by parts except possibly for the boundary terms. They turn out to be zero since ρ has compact support, $\lim_{t\downarrow 0}\frac{d^l}{dt^l}(t^k f_\rho(t)) = 0$ when $l \leq k-2$ and $\lim_{t\downarrow 0}\frac{d^l}{dt^l}(t^k f_\rho(t))\sin(2\pi\theta t) = 0$ when $l = k-1$ by (5.39) and Lemma 5.4.8; this also shows that the intermediate integrals leading to (5.43) in the induction process are finite. The final integral on the right-hand side of (5.43) is also finite by (5.39), (5.10) and the compact support of ρ. Setting $N = \lfloor\theta\varepsilon\rfloor \geq \theta\varepsilon/2$ (where $\lfloor\cdot\rfloor$ denotes the integer part) and $b_k(r) = \frac{d^k}{dr^k}(r^k f_\rho(r))$, we estimate the right-hand side of (5.43) as follows:

$$\left|\int_0^\infty \frac{d^k}{dr^k}(r^k f_\rho(r))\sin(2\pi\theta r)\,dr\right| \leq \left|\int_0^{\frac{N}{\theta}} b_k(r)\sin(2\pi\theta r)\,dr\right| + \int_{\varepsilon/2}^\infty \left|\frac{d^k}{dr^k}(r^k f_\rho(r))\right|dr, \tag{5.44}$$

since $N \geq \theta\varepsilon/2$. On the one hand,

$$\int_{\varepsilon/2}^\infty \left|\frac{d^k}{dr^k}(r^k f_\rho(r))\right|dr \leq C_k\int_{\varepsilon/2}^\infty f_\rho(r)\,dr \leq C_k,$$

by (5.39) and (5.10). On the other hand, as in (5.31) and (5.32),

$$\int_0^{\frac{N}{\theta}} b_k(r)\sin(2\pi\theta r)\,dr = \sum_{j=0}^{N-1}\int_{\frac{j}{\theta}}^{\frac{j+\frac{1}{2}}{\theta}}\left(b_k(r) - b_k\left(r + \frac{1}{2\theta}\right)\right)\sin(2\pi\theta r)\,dr. \tag{5.45}$$

By the fundamental theorem of calculus,

$$\begin{aligned}\left|b_k(r) - b_k\left(r + \frac{1}{2\theta}\right)\right| &\leq \int_r^{r+\frac{1}{2\theta}} \left|\frac{d}{dt} b_k(t)\right| dt \\ &\leq C_k \int_r^{r+\frac{1}{2\theta}} -\frac{d}{dt} f_\rho(t)\, dt = C_k \left(f_\rho(r) - f_\rho\left(r + \frac{1}{2\theta}\right)\right),\end{aligned}$$

where we have used (5.40). Thus,

$$\sum_{j=0}^{N-1} \int_{\frac{j}{\theta}}^{\frac{j+\frac{1}{2}}{\theta}} \left|b_k(r) - b_k\left(r + \frac{1}{2\theta}\right)\right| \sin(2\pi\theta r)\, dr \leq C_k \int_0^{\frac{N}{\theta}} f_\rho(r) \sin(2\pi\theta r)\, dr,$$

reasoning as in (5.45). Putting together the formulas following (5.44), we obtain

$$\left|\int_0^\infty \frac{d^k}{dr^k}(r^k f_\rho(r)) \sin(2\pi\theta r)\, dr\right| \leq C_k \int_0^{\frac{N}{\theta}} f_\rho(r) \sin(2\pi\theta r)\, dr + C_k.$$

Together with (5.43), the bounds (5.42) follow.

Step 2: Conclusion. Recalling formula (5.16), we obtain, from Leibniz' rule and interchanging the derivative with the integral, that

$$|\partial^\alpha \widehat{Q}(\xi)| \leq C_\alpha \sum_{k=0}^{|\alpha|} \frac{1}{|\xi|^{|\alpha|+1-k}} \left|\int_{\mathbb{R}^n} \frac{\rho(x) x_1^{k+1}}{|x|^2} a_k(2\pi|\xi|x_1)\, dx\right|, \qquad \xi \in \mathbb{R}^n \setminus \{0\}. \tag{5.46}$$

When $k = 0$, the term in the sum is equal to $2\pi|\xi|^{-|\alpha|}\widehat{Q}_\rho(\xi)$, cf. (5.16), which already has the correct form. For $k \geq 1$, we can compute as in (5.30) that

$$\int_{\mathbb{R}^n} \frac{\rho(x) x_1^{k+1}}{|x|^2} a_k(2\pi|\xi|x_1)\, dx = 2 \int_{\mathbb{S}_+^{n-1}} z_1^{k+1} \int_0^\infty r^k f_\rho(r) a_k(2\pi|\xi|z_1 r)\, dr\, d\mathcal{H}^{n-1}(z).$$

We split the integral in $\mathbb{S}_+^{n-1}$ into $\{z_1|\xi| > 1/\varepsilon\}$ and $\{z_1|\xi| \leq 1/\varepsilon\}$. In the first subset we have, thanks to (5.42) with $\theta = |\xi|z_1$,

$$\begin{aligned}&\left|\int_{\mathbb{S}_+^{n-1}\cap\{z_1|\xi|>1/\varepsilon\}} z_1^{k+1} \int_0^\infty r^k f_\rho(r) a_k(2\pi|\xi|z_1 r)\, dr\, d\mathcal{H}^{n-1}(z)\right| \\ &\leq C_k \left|\int_{\mathbb{S}_+^{n-1}\cap\{z_1|\xi|>1/\varepsilon\}} \frac{z_1^{k+1}}{(|\xi|z_1)^k} \left(\int_0^\infty f_\rho(r) \sin(2\pi|\xi|z_1 r)\, dr + 1\right) d\mathcal{H}^{n-1}(z)\right| \\ &\leq \frac{C_k}{|\xi|^k} \left|\int_{\mathbb{S}_+^{n-1}} z_1 \int_0^\infty f_\rho(r) \sin(2\pi|\xi|z_1 r)\, dr\, d\mathcal{H}^{n-1}(z)\right| + \frac{C_k}{|\xi|^k} = \frac{C_k}{|\xi|^{k-1}} \widehat{Q}_\rho(\xi) + \frac{C_k}{|\xi|^k},\end{aligned}$$

where the last equality is due to (5.30). In the second subset we have, for $|\xi| \geq 1$ and $k \geq 1$,

$$\begin{aligned}&\left|\int_{\mathbb{S}_+^{n-1}\cap\{z_1|\xi|\leq 1/\varepsilon\}} z_1^{k+1} \int_0^\infty r^k f_\rho(r) a_k(2\pi|\xi|z_1 r)\, dr\, d\mathcal{H}^{n-1}(z)\right| \\ &\leq \left|\int_{\mathbb{S}_+^{n-1}\cap\{z_1|\xi|\leq 1/\varepsilon\}} \frac{1}{(|\xi|\varepsilon)^{k+1}} \int_0^\infty r^k f_\rho(r)\, dr\, d\mathcal{H}^{n-1}(z)\right| \leq \frac{C_k}{|\xi|^{k+1}} \leq \frac{C_k}{|\xi|^k},\end{aligned}$$

where we have used (5.10) and the compact support of ρ. All in all, we have shown

$$\left|\int_{\mathbb{R}^n} \frac{\rho(x) x_1^{k+1}}{|x|^2} a_k(2\pi|\xi|x_1)\, dx\right| \leq \frac{C_k}{|\xi|^{k-1}} \widehat{Q}_\rho(\xi) + \frac{C_k}{|\xi|^k}, \qquad |\xi| \geq 1, \quad k \geq 0,$$

which yields the result thanks to (5.46). □

With these bounds we can apply the Mihlin-Hörmander theorem to prove the analogue of Theorem 5.4.5 in the L^p-setting.

Theorem 5.4.11. *Let $p \in (1, \infty)$. Let $\Omega \subset \mathbb{R}^n$ be open and bounded and suppose that ρ satisfies (H0)–(H2) and has compact support. Then, the following two statements hold:*

(i) If $\liminf_{t\downarrow 0} t^{n-1}\overline{\rho}(t) > 0$, then there is a $C = C(\Omega, n, \rho) > 0$ such that

$$\|u\|_{L^p(\Omega)} \le C\|D_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \quad \textit{for all } u \in H_0^{\rho,p}(\Omega).$$

(ii) If $\lim_{t\downarrow 0} t^{n-1}\overline{\rho}(t) = \infty$, then $H_0^{\rho,p}(\Omega)$ is compactly embedded into $L^p(\mathbb{R}^n)$.

Proof. Part (i). In light of Lemma 5.4.3 and (5.36), we find that

$$\widehat{Q}_\rho(\xi) \ge C/|\xi| \qquad \text{for } |\xi| \ge 1, \tag{5.47}$$

so it follows from Lemma 5.4.10 that for every $\alpha \in \mathbb{N}^n$,

$$|\partial^\alpha \widehat{Q}_\rho(\xi)| \le C_\alpha |\xi|^{-|\alpha|}\widehat{Q}_\rho(\xi) \quad \text{for } |\xi| \ge 1.$$

Applying Faà di Bruno's formula for the derivatives of a composition,

$$|\partial^\alpha \widehat{Q}_\rho^{-1}| \le C_\alpha \sum_{k=1}^{|\alpha|} \widehat{Q}_\rho^{-1-k} \sum_{j_1,\dots,j_{|\alpha|-k+1}} B_{j_1,\dots,j_{|\alpha|-k+1}},$$

where the indices $j_1, \dots, j_{|\alpha|-k+1} \in \mathbb{N}$ in the sum are taken with the restrictions

$$j_1 + \cdots + j_{|\alpha|-k+1} = k \quad \text{and} \quad j_1 + \cdots + (|\alpha| - k + 1)j_{|\alpha|-k+1} = |\alpha|,$$

and $B_{j_1,\dots,j_{|\alpha|-k+1}}$ is a symbol for all products of the absolute value of the partial derivatives of $\widehat{Q}_\rho$ with j_i derivatives of order i, for $i = 1, \dots, |\alpha| - k + 1$. Thus,

$$B_{j_1,\dots,j_{|\alpha|-k+1}} \le C_\alpha \left(|\xi|^{-1}\widehat{Q}_\rho\right)^{j_1} \cdots \left(|\xi|^{-(|\alpha|-k+1)}\widehat{Q}_\rho\right)^{j_{|\alpha|-k+1}}.$$

Hence, for $|\xi| \ge 1$, using (5.47),

$$|\partial^\alpha \widehat{Q}_\rho^{-1}| \le C_\alpha |\xi|^{-|\alpha|}\widehat{Q}_\rho^{-1} \le C_\alpha |\xi|^{1-|\alpha|}.$$

Now, defining $R(\xi) = \xi|\xi|^{-2}$, we have that

$$|\partial^\alpha R(\xi)| \le C_\alpha |\xi|^{-1-|\alpha|}.$$

Let $\chi \in C_c^\infty(\mathbb{R}^n)$ and $M \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ be as in the proof of Theorem 5.4.5, so that identity (5.35) holds. With the calculations above, thanks to Leibniz' rule we can estimate for all $\alpha \in \mathbb{N}^n$ that

$$|\partial^\alpha M(\xi)| \le C_\alpha |\xi|^{-1-|\alpha|}\widehat{Q}_\rho^{-1} \le C_\alpha |\xi|^{-|\alpha|}, \quad \text{for } |\xi| \ge 1. \tag{5.48}$$

As such, the Mihlin-Hörmander theorem (cf. [122, Th. 6.2.7]) implies that the operator T_M of (5.37) can be extended to a bounded operator from $L^p(\mathbb{R}^n, \mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$. The Poincaré inequality can now be argued as in Theorem 5.4.5.

Part (ii). We saw in the proof of Theorem 5.4.5 (ii) that the operator T_M is compact from $L^2(\Omega_\delta, \mathbb{R}^n)$ to $L^2(\Omega)$, since $\lim_{t\downarrow 0} t^{n-1}\overline{\rho}(t) = \infty$. As a consequence of the proof of part (i), T_M is bounded from $L^p(\Omega_\delta, \mathbb{R}^n)$ to $L^p(\Omega)$ for all $p \in (1, \infty)$ and, hence, by Krasnoselskii's interpolation theorem [41, Th. IV.2.9], also compact for all $p \in (1, \infty)$. The compact embedding now follows as in the proof of Theorem 5.4.5 (ii). □

Example 5.4.12. Let $p \in (1, \infty)$ and let $\Omega \subset \mathbb{R}^n$ be open and bounded.

(a) Let $0 \le s < 1$, $\chi \in C_c^\infty(\mathbb{R}^n)$ and ρ be as in Example 5.4.9 (b). By Theorem 5.4.11, if $s = 0$ we have a Poincaré inequality, while for $s > 0$ we obtain compactness. This constitutes an alternative proof as well as a generalization of the results [31, Thms. 6.2 and 7.3], where it was assumed that χ is constant around the origin and radially decreasing.

(b) Let $0 \le s < 1$ and ρ be as in Example 5.4.9 (c). By Theorem 5.4.11, we have both a Poincaré inequality and compactness.

(c) Let $0 \le s < 1$ and ρ be as in Example 5.4.9 (d). By Theorem 5.4.11, if $s > 0$ we have both a Poincaré inequality and compactness.

(d) Let $s : [0, \infty) \to (0, 1)$ and ρ be as in Example 5.4.9 (e). By Theorem 5.4.11, we have both a Poincaré inequality and compactness.

5.5 Fundamental theorem of calculus

In this section, we will study when the inverse Fourier transform of W_ρ (see (5.33)) is a locally integrable function, and identify its behavior at the origin. This function will then give us an analogue of the fundamental theorem of calculus for the nonlocal gradient; precisely, applying the inverse Fourier transform to (5.34) yields $u = V_\rho * \mathcal{G}_\rho u$ with $V_\rho := W_\rho^\vee$.

For $W_\rho^\vee$ to be a locally integrable function, we require the following mild assumptions, saying that the kernel ρ lies between two fractional kernels:

(H3) The function $g_\rho : (0, \infty) \to \mathbb{R}$, $t \mapsto t^{n+\sigma-1}\overline{\rho}(t)$ is almost decreasing on $(0, \varepsilon)$ for some $\sigma \in (0, 1)$;

(H4) the function $h_\rho : (0, \infty) \to \mathbb{R}$, $t \mapsto t^{n+\gamma-1}\overline{\rho}(t)$ is almost increasing on $(0, \varepsilon)$ for some $\gamma \in (0, 1)$.

Recall from the notation section the definition of *almost decreasing/increasing*. We will use, as a consequence of (H3) and Lemma 5.4.3 that

$$\widehat{Q}_\rho(\xi) \ge C|\xi|^{\sigma-1}, \quad |\xi| \ge 1. \tag{5.49}$$

Note also that (H3)–(H4) require $\sigma \le \gamma$; this can be seen directly or invoking (5.58) below.

Example 5.5.1. Classes of kernels ρ satisfying (H0)–(H4) are:

(a) Given $s \in (0, 1)$,

$$\rho(x) = \frac{1}{|x|^{n+s-1}}.$$

In this case, one can take $\sigma = \gamma = s$. The same conclusion holds for

$$\rho(x) = \frac{\chi(x)}{|x|^{n+s-1}} \tag{5.50}$$

with $\chi \in C_c^\infty(\mathbb{R}^n)$ a non-negative radial function such that $\chi(0) > 0$ and $\chi(x)/|x|^{1+s}$ is radially decreasing.

(b) Given $s \in (0, 1)$,

$$\rho(x) = \frac{\chi(x)(-\log|x|)}{|x|^{n+s-1}} \tag{5.51}$$

with $\chi \in C_c^\infty(B_1)$ a non-negative radial function such that $\chi(0) > 0$ and $\chi(x)(-\log|x|)/|x|^{1+s}$ is radially decreasing. In this case one can take $\sigma = s$ and any $\gamma \in (s, 1)$.

(c) Given $s \in (0,1)$,

$$\rho(x) = \frac{\chi(x)}{|x|^{n+s-1}(-\log|x|)} \tag{5.52}$$

with $\chi \in C_c^\infty(B_1)$ a non-negative radial function such that $\chi(0) > 0$ and $\chi(x)/(-|x|^{1+s}\log|x|)$ is radially decreasing. In this case one can take any $\sigma \in (0,s)$ and $\gamma = s$.

(d) Given a smooth $s : [0,\infty) \to (0,1)$ and a non-negative radial function $\chi \in C_c^\infty(\mathbb{R}^n)$ such that $\chi(0) > 0$ and $\chi(x)/|x|^{1+s(|x|)}$ is radially decreasing,

$$\rho(x) = \frac{\chi(x)}{|x|^{n+s(|x|)-1}}.$$

In this case, one can take $\sigma = \min_{[0,\varepsilon]} s$ and $\gamma = \max_{[0,\varepsilon]} s$ for any $\varepsilon > 0$.

(e) If ρ_1, ρ_2 satisfy (H0)–(H4) and $\alpha_1, \alpha_2 > 0$ then $\alpha_1\rho_1 + \alpha_2\rho_2$ satisfies (H0)–(H4). In fact, let σ_1, σ_2 be the exponents of (H3) for ρ_1, ρ_2, respectively; let γ_1, γ_2 be the exponents of (H4) for ρ_1, ρ_2, respectively. Then, (H3) holds for $\alpha_1\rho_1 + \alpha_2\rho_2$ with the exponent $\min\{\sigma_1, \sigma_2\}$, while (H4) holds with the exponent $\max\{\gamma_1, \gamma_2\}$.

In the following result, the proof that V_ρ is a function is adapted from that of [31, Th. 5.9], while the bounds around the origin on V_ρ require different arguments since we cannot compare it with the Riesz kernel.

Theorem 5.5.2. *Let ρ have compact support and satisfy (H0)–(H4). Then, there exists a vector radial function $V_\rho \in C^\infty(\mathbb{R}^n \setminus \{0\}, \mathbb{R}^n) \cap L^1_{\text{loc}}(\mathbb{R}^n, \mathbb{R}^n)$ such that for all $\varphi \in C_c^\infty(\mathbb{R}^n)$,*

$$\varphi(x) = \int_{\mathbb{R}^n} V_\rho(x-y) \cdot \mathcal{G}_\rho\varphi(y)\, dy \quad \text{for all } x \in \mathbb{R}^n. \tag{5.53}$$

Moreover, there is a constant $C = C(n,\rho) > 0$ such that for $x \in B_\varepsilon \setminus \{0\}$,

$$|V_\rho(x)| \le \frac{C}{|x|^{2n-1}\rho(x)} \quad \text{and} \quad |\nabla V_\rho(x)| \le \frac{C}{|x|^{2n}\rho(x)}. \tag{5.54}$$

Proof. In light of Proposition 5.4.4 and the well-known interaction between Fourier transforms and multiplication and convolution, it suffices to show that the inverse Fourier transform $V_\rho := W_\rho^\vee$ agrees with a locally integrable function with the stated properties.

Step 1: V_ρ is a function. Let $\chi \in C_c^\infty(\mathbb{R}^n)$ be a radial cut-off function with $\chi \equiv 1$ on $B_{2/\varepsilon}$. As in the proof of Theorem 5.4.5 (i), we can write

$$W_\rho = \chi W_\rho + (1-\chi)W_\rho =: W_\rho^1 + W_\rho^2.$$

Since W_ρ^1 has compact support, it follows from the Paley-Wiener theorem that $(W_\rho^1)^\vee$ is analytic. We also observe that W_ρ^2 is actually a smooth locally integrable function, namely,

$$W_\rho^2(\xi) = (1-\chi(\xi))\frac{-i\xi}{2\pi|\xi|^2\widehat{Q}_\rho(\xi)},$$

as in (5.35). From Lemma 5.4.10 and its consequence (5.48), and (5.49) we obtain that for any $\alpha \in \mathbb{N}^n$,

$$|\partial^\alpha W_\rho^2(\xi)| \le \frac{C_\alpha}{(1+|\xi|^{1+|\alpha|})\widehat{Q}_\rho(\xi)} \le \frac{C_\alpha}{1+|\xi|^{|\alpha|+\sigma}}, \quad \xi \in \mathbb{R}^n. \tag{5.55}$$

If $|\alpha| \geq n+m$ for some $m \in \mathbb{N}$, we deduce from [122, Exercise 2.4.1] that

$$(\partial^\alpha W_\rho^2)^\vee = (-2\pi i\cdot)^\alpha (W_\rho^2)^\vee, \tag{5.56}$$

lies in $C^m(\mathbb{R}^n)$. In particular, we find that $(W_\rho^2)^\vee$ coincides with a smooth function $K : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n$ outside the origin. We show in step 2 below that K is integrable, which implies that

$$(W_\rho^2)^\vee = K + V_0,$$

where V_0 is a tempered distribution supported at the origin. Therefore, by [122, Prop. 2.4.1], V_0 can be written as a linear combination of derivatives of Dirac deltas, i.e.,

$$V_0 = \sum_{|\alpha|\leq k} c_\alpha \partial^\alpha \delta_0 \quad \text{and} \quad \widehat{V_0} = \sum_{|\alpha|\leq k} c_\alpha (2\pi i\cdot)^\alpha$$

for some $k \in \mathbb{N}$ and $c_\alpha \in \mathbb{C}$. We then obtain that

$$W_\rho^2 = \widehat{K} + \sum_{|\alpha|\leq k} c_\alpha (2\pi i\cdot)^\alpha.$$

However, W_ρ^2 vanishes at infinity (thanks to (5.55)) and so does $\widehat{K}$ as the Fourier transform of an integrable function, so we must have $c_\alpha = 0$ for all $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq k$. This shows that $(W_\rho^2)^\vee = K$, and hence

$$V_\rho = W_\rho^\vee = (W_\rho^1)^\vee + (W_\rho^2)^\vee \in C^\infty(\mathbb{R}^n \setminus \{0\}; \mathbb{R}^n),$$

is a locally integrable function. Note that V_ρ is real-valued and vector radial since W_ρ is imaginary-valued and vector radial.

Step 2: K is integrable. By choosing $\alpha = (n, 0, \dots, 0)$ and $\alpha = (n+1, 0, \cdots, 0)$ in (5.56) and using symmetry considerations, we have for $x \neq 0$ that

$$|K(x)| = |K(|x|e_1)| = \frac{|(\partial_1^n W_\rho^2)^\vee(|x|e_1)|}{(2\pi|x|)^n} \quad \text{and} \quad |K(x)| = |K(|x|e_1)| = \frac{|(\partial_1^{n+1} W_\rho^2)^\vee(|x|e_1)|}{(2\pi|x|)^{n+1}}.$$

The function $(\partial_1^{n+1} W_\rho^2)^\vee$ is bounded since $\partial_1^{n+1} W_\rho^2$ is integrable by (5.55); consequently, we deduce from the second identity that

$$|K(x)| \leq \frac{C}{|x|^{n+1}} \quad x \in \mathbb{R}^n \setminus \{0\},$$

which shows that K is integrable away from the origin. Near the origin, we use that $\partial_1^n W_\rho^2$ is also integrable (cf. (5.55)). Hence, we may utilize the standard formula of the Fourier transform and partial integration for $0 < |x| < 2/\varepsilon$, to find

$$\begin{aligned}
\left|(\partial_1^n W_\rho^2)^\vee(|x|e_1)\right| &= \left|\int_{\mathbb{R}^n} \partial_1^n W_\rho^2(\xi) e^{2\pi i|x|\xi_1}\, d\xi\right| \\
&\leq 2\pi|x| \left|\int_{B_{1/|x|}\setminus B_{2/\varepsilon}} \partial_1^{n-1} W_\rho^2(\xi) e^{2\pi i|x|\xi_1}\, d\xi\right| \\
&\quad + \left|\int_{\partial B_{1/|x|}} \partial_1^{n-1} W_\rho^2(\xi) e^{2\pi i|x|\xi_1}\, d\mathcal{H}^{n-1}(\xi)\right| + \left|\int_{B^c_{1/|x|}} \partial_1^n W_\rho^2(\xi) e^{2\pi i|x|\xi_1}\, d\xi\right| \\
&\leq C|x| \int_{B_{1/|x|}\setminus B_{2/\varepsilon}} \frac{d\xi}{|\xi|^n \widehat{Q}_\rho(\xi)} + C\int_{\partial B_{1/|x|}} \frac{d\mathcal{H}^{n-1}(\xi)}{|\xi|^n \widehat{Q}_\rho(\xi)} + C\int_{B^c_{1/|x|}} \frac{d\xi}{|\xi|^{n+1} \widehat{Q}_\rho(\xi)}.
\end{aligned}$$

The first inequality is integration by parts in $B_{1/|x|}$ and using that $(1-\chi)$ and its derivatives are zero on $B_{2/\varepsilon}$, whereas the second inequality uses (5.55). We now estimate each of the three terms of the right-hand side of the last inequality. We may use Lemma 5.4.3 and (H4) to find

$$|x|\int_{B_{1/|x|}\setminus B_{2/\varepsilon}}\frac{d\xi}{|\xi|^n\widehat{Q}_\rho(\xi)}\le C|x|\int_{B_{1/|x|}\setminus B_{2/\varepsilon}}\frac{1}{\overline{\rho}(1/|\xi|)}\,d\xi=C|x|\int_{B_{1/|x|}\setminus B_{2/\varepsilon}}\frac{1}{|\xi|^{n+\gamma-1}h_\rho(1/|\xi|)}\,d\xi$$
$$\le\frac{C|x|}{h_\rho(|x|)}\int_{B_{1/|x|}}\frac{1}{|\xi|^{n+\gamma-1}}\,d\xi=\frac{C|x|}{|x|^{1-\gamma}h_\rho(|x|)}=\frac{C}{|x|^{n-1}\rho(x)}.$$

For the second term we use only Lemma 5.4.3 and find

$$\int_{\partial B_{1/|x|}}\frac{d\mathcal{H}^{n-1}(\xi)}{|\xi|^n\widehat{Q}_\rho(\xi)}\le C\int_{\partial B_{1/|x|}}\frac{1}{\overline{\rho}(1/|\xi|)}\,d\mathcal{H}^{n-1}(\xi)=\frac{C}{|x|^{n-1}\rho(x)}.$$

Finally, for the last term we compute with Lemma 5.4.3 and (H3)

$$\int_{B^c_{1/|x|}}\frac{d\xi}{|\xi|^{n+1}\widehat{Q}_\rho(\xi)}\le C\int_{B^c_{1/|x|}}\frac{1}{|\xi|\overline{\rho}(1/|\xi|)}\,d\xi=C\int_{B^c_{1/|x|}}\frac{1}{|\xi|^{n+\sigma}g_\rho(1/|\xi|)}\,d\xi$$
$$\le\frac{C}{g_\rho(|x|)}\int_{B^c_{1/|x|}}\frac{1}{|\xi|^{n+\sigma}}\,d\xi=\frac{C|x|^\sigma}{g_\rho(|x|)}=\frac{C}{|x|^{n-1}\rho(x)}.$$

All in all, and using also (H3), this shows that for $0<|x|<\varepsilon/2$

$$|K(x)|=\frac{|(\partial_1^n W_\rho^2)^\vee(|x|e_1)|}{(2\pi|x|)^n}\le\frac{C}{|x|^{2n-1}\rho(x)}\le\frac{C}{|x|^{n-\sigma}},\tag{5.57}$$

which proves that K is also integrable around the origin.

Step 3: Bounds on V_ρ. Since V_ρ coincides with K up to a smooth function we deduce from (5.57) that the first inequality in (5.54) holds on any bounded set on which ρ is positive, in particular, on $B_\varepsilon\setminus\{0\}$. The bound on the gradient of V_ρ follows from analogous calculations to that of step 2, since

$$\nabla K=((2\pi i\cdot)\otimes W_\rho^2)^\vee.$$

□

When we apply Theorem 5.5.2 to Example 5.5.1 (a), we recover and generalize the nonlocal fundamental theorem of calculus of [31, Th. 4.5]. We can also extend (5.53) to the setting of Sobolev spaces.

Corollary 5.5.3. *Let ρ have compact support and satisfy (H0)–(H4). Let $p\in[1,\infty]$ and $\Omega\subset\mathbb{R}^n$ be open and bounded. Then, for all $u\in H_0^{\rho,p}(\Omega)$ it holds that*

$$u(x)=\int_{\mathbb{R}^n}V_\rho(x-y)\cdot D_\rho u(y)\,dy\quad\text{for a.e. } x\in\mathbb{R}^n.$$

Proof. Since Ω is bounded, we find $H_0^{\rho,p}(\Omega)\subset H_0^{\rho,1}(\Omega)$, so it suffices to prove the statement for $u\in H_0^{\rho,1}(\Omega)$. This can be done by a simple mollification argument. Indeed, we can find a sequence $(\varphi_j)_j\subset C_c^\infty(\Omega')$ for some open and bounded $\Omega'\subset\mathbb{R}^n$, such that $\varphi_j\to u$ in $L^1(\mathbb{R}^n)$ and $\mathcal{G}_\rho\varphi_j\to D_\rho u$ in $L^1(\mathbb{R}^n,\mathbb{R}^n)$. Then, the fact that $(\mathcal{G}_\rho\varphi_j)_j$ is supported in a fixed compact set and V_ρ is locally integrable yields by Young's convolution inequality

$$\varphi_j=\int_{\mathbb{R}^n}V_\rho(\cdot-y)\cdot\mathcal{G}_\rho\varphi_j(y)\,dy\to\int_{\mathbb{R}^n}V_\rho(\cdot-y)\cdot D_\rho u(y)\,dy\text{ in }L^1_{\text{loc}}(\mathbb{R}^n)\text{ as }j\to\infty.$$

Hence, u must coincide with the right-hand side, which proves the statement. □

5.6 Embeddings

The aim of this section is to apply the nonlocal fundamental theorem of calculus (Theorem 5.5.2) to prove embeddings, Poincaré inequalities and compactness results. One of the advantages of this approach is that some of them can also be proven for $p = 1$ and $p = \infty$, which is not possible with purely Fourier arguments, as in Sections 5.4.2 and 5.4.3. Moreover, the embeddings shown are not restricted to Lebesgue or Hölder spaces, but to the more general Orlicz spaces and spaces with a prescribed modulus of continuity. Thus, the proof of those embeddings cannot be obtained by reducing to the fractional setting.

Throughout this section, we assume ρ satisfies (H0)–(H4). To start the analysis of the embeddings, we introduce the modulus of continuity $\omega : [0, \varepsilon) \to [0, \infty)$

$$\omega(t) = \begin{cases} \frac{1}{t^{n-1}\overline{\rho}(t)} & \text{if } t \in (0, \varepsilon), \\ 0 & \text{if } t = 0. \end{cases}$$

It is a modulus of continuity in the sense that it is continuous by (H2) and (H3) and almost increasing by (H3). In view of (H3) and (H4), we have that there exists $C > 0$ such that for all $t \in (0, \varepsilon)$ and $\lambda \in [1, \varepsilon/t)$,

$$\frac{t^{\gamma}}{C} \leq \omega(t) \leq C t^{\sigma} \quad \text{and} \quad \frac{\lambda^{\sigma}\omega(t)}{C} \leq \omega(\lambda t) \leq C\lambda^{\gamma}\omega(t). \tag{5.58}$$

The second inequalities show that we may take scaling factors outside ω, which we will often use without mention.

5.6.1 Embeddings into Orlicz spaces

Our first result (Theorem 5.6.2) is an embedding into Orlicz spaces, together with its associated Poincaré inequality. Its proof uses an inequality of the style of Hardy-Littlewood-Sobolev for generalized Riesz potentials proved in [138].

As usual in the theory of Orlicz spaces, we say that $A : [0, \infty) \to [0, \infty)$ is a Young function if it is continuous, strictly increasing, convex, with $A(0) = 0$ and $\lim_{t\to\infty} A(t) = \infty$; note that any Young function is invertible. Then, for $\Omega \subset \mathbb{R}^n$ open one can define the Orlicz space

$$L^A(\Omega) = \{u : \Omega \to \mathbb{R} \text{ measurable} \, : \, \|u\|_{L^A(\Omega)} < \infty\},$$

with the Luxemburg norm

$$\|u\|_{L^A(\Omega)} = \inf\left\{\lambda > 0 \, : \, \int_{\Omega} A\left(\frac{|u(x)|}{\lambda}\right) dx \leq 1\right\}.$$

Of course, if $A(t) = t^p$ with $p \in (1, \infty)$ then $L^A(\Omega) = L^p(\Omega)$ with the same norm.

In the proof of Theorem 5.6.2 below, given a measurable function $\tilde{\omega} : (0, \infty) \to (0, \infty)$ we will use the operator $I_{\tilde{\omega}}$, sending measurable functions $u : \mathbb{R}^n \to \mathbb{R}$ to measurable functions $I_{\tilde{\omega}}(u) : \mathbb{R}^n \to \mathbb{R}$, defined as

$$I_{\tilde{\omega}}(u)(x) := \int_{\mathbb{R}^n} \frac{\tilde{\omega}(|x-y|)}{|x-y|^n} u(y) \, dy, \qquad x \in \mathbb{R}^n.$$

In particular, we will use the boundedness properties of this operator between Lebesgue and Orlicz spaces proved in [138, Cor. 3.8 (i)], which we reproduce for ease of reference.

Proposition 5.6.1. *Let $p \in (1, \infty)$. Assume that:*

(a) $\displaystyle\int_0^1 \frac{\tilde{\omega}(t)}{t} \, dt < \infty.$

(b) There exist $C, K_1, K_2 > 0$ *with* $K_1 < K_2$ *such that for all* $r > 0$,

$$\sup_{r \le t \le 2r} \tilde{\omega}(t) \le C \int_{K_1 r}^{K_2 r} \frac{\tilde{\omega}(t)}{t}\, dt.$$

(c) There exists $C > 0$ *such that for all* $r > 0$,

$$\frac{1}{r^{n/p}} \int_0^r \frac{\tilde{\omega}(t)}{t}\, dt + \int_r^\infty \frac{\tilde{\omega}(t)}{t^{1+n/p}}\, dt \le CA^{-1}(1/r^n). \tag{5.59}$$

Then, $I_{\tilde{\omega}}$ *is bounded from* $L^p(\mathbb{R}^n)$ *to* $L^A(\mathbb{R}^n)$.

With the aid of Theorem 5.5.2, we have the following embedding of $H_0^{\rho,p}(\Omega)$ into Orlicz spaces.

Theorem 5.6.2. *Let* ρ *satisfy (H0)–(H4) and have compact support. Let* Ω *be bounded. Assume* $p \in (1, \infty)$ *satisfies* $\gamma p < n$*, and* A *is a Young function such that*

$$\liminf_{t \to \infty} \frac{A^{-1}(t)}{\omega\left(t^{-1/n}\right) t^{1/p}} > 0. \tag{5.60}$$

Then, $H_0^{\rho,p}(\Omega)$ *is embedded into* $L^A(\Omega)$ *and there is a constant* $C = C(\Omega, A, n, \rho) > 0$ *such that*

$$\|u\|_{L^A(\Omega)} \le C\|D_\rho u\|_{L^p(\mathbb{R}^n, \mathbb{R}^n)} \quad \text{for all } u \in H_0^{\rho,p}(\Omega). \tag{5.61}$$

Proof. Define $\tilde{\omega} : (0, \infty) \to (0, \infty)$ as

$$\tilde{\omega}(t) = \begin{cases} \omega(t) & \text{for } t \in (0, \varepsilon), \\ e^{-t} & \text{for } t \in [\varepsilon, \infty). \end{cases}$$

Bound (5.54) and the fact that V_ρ is locally bounded away from the origin allows us to establish the estimate

$$V_\rho(x) \le C \frac{\tilde{\omega}(|x|)}{|x|^n}, \qquad x \in B_{\operatorname{diam}\Omega + \delta} \setminus \{0\}, \tag{5.62}$$

for a suitable constant $C > 0$ and $\delta > 0$ given by $\operatorname{supp} \rho = \overline{B}_\delta$.

Let $u \in H_0^{\rho,p}(\Omega)$. By Corollary 5.5.3 and (5.62) we can estimate for a.e. $x \in \Omega$

$$|u(x)| \le \int_{\Omega_\delta} |V_\delta(x-y)||D_\rho u(y)|\, dy \le C \int_{\mathbb{R}^n} \frac{\tilde{\omega}(|x-y|)}{|x-y|^n} |D_\rho u(y)|\, dy = CI_{\tilde{\omega}}(|D_\rho u|)(x).$$

On the other hand, the bound $|u| \le CI_{\tilde{\omega}}(|D_\rho u|)$ is obvious in Ω^c. Therefore, it suffices to show that the operator $I_{\tilde{\omega}}$ is bounded from $L^p(\mathbb{R}^n)$ into $L^A(\Omega)$, since we then find

$$\|u\|_{L^A(\Omega)} \le C\|I_{\tilde{\omega}}(|D_\rho u|)\|_{L^A(\Omega)} \le C\||D_\rho u|\|_{L^p(\mathbb{R}^n)} = C\|D_\rho u\|_{L^p(\mathbb{R}^n, \mathbb{R}^n)}.$$

For the boundedness of $I_{\tilde{\omega}}$ from $L^p(\mathbb{R}^n)$ into $L^A(\Omega)$, we shall check (a)–(c) of Proposition 5.6.1. Condition (a) holds thanks to (5.58), since

$$\int_0^\varepsilon \frac{\tilde{\omega}(t)}{t}\, dt \le C \int_0^\varepsilon \frac{1}{t^{1-\sigma}}\, dt < \infty.$$

As for (b), we calculate, for $r < \varepsilon/2$,

$$\int_{r/2}^{2r} \frac{\tilde{\omega}(t)}{t}\, dt \ge C \sup_{r \le t \le 2r} \tilde{\omega}(t) \int_{r/2}^{2r} \frac{1}{t}\, dt = C \log(4) \sup_{r \le t \le 2r} \tilde{\omega}(t),$$

where we have used the second property in (5.58), while for $r > 2\varepsilon$

$$\int_{r/2}^{2r} \frac{\tilde{\omega}(t)}{t}\,dt \geq \frac{1}{2r}\int_{r/2}^{2r} e^{-t}\,dt = \frac{e^{r/2}-e^{-r}}{2r}e^{-r} \geq C \sup_{r\leq t\leq 2r} \tilde{\omega}(t).$$

In closing, the bound

$$\int_{r/2}^{2r} \frac{\tilde{\omega}(t)}{t}\,dt \geq C \sup_{r\leq t\leq 2r} \tilde{\omega}(t), \qquad \frac{\varepsilon}{2} \leq r \leq 2\varepsilon$$

holds trivially as $\tilde{\omega}$ is a piecewise smooth positive function. Thus, condition (b) is proved.

Finally, we check (c). In fact, since we are only interested in the embedding into $L^A(\Omega)$ for Ω bounded, it suffices to verify inequality (5.59) for $r < r_0$ for some $r_0 < \varepsilon$. Indeed, for $r \geq r_0$ the left-hand side of (5.59) is bounded by a constant times $r^{-n/p}$, so that we can change A around zero such that that the inequality is satisfied everywhere (cf. [6, Lemma 4.5]), which leads to an equivalent Orlicz space [180, Th. V.1.3]. For $r < r_0$ we compute

$$\frac{1}{r^{n/p}}\int_0^r \frac{\tilde{\omega}(t)}{t}\,dt + \int_r^\infty \frac{\tilde{\omega}(t)}{t^{1+n/p}}\,dt \leq C\frac{\omega(r)}{r^{n/p+\sigma}}\int_0^r \frac{1}{t^{1-\sigma}}\,dt + C\frac{\omega(r)}{r^\gamma}\int_r^\infty \frac{1}{t^{1-\gamma+n/p}}\,dt = \frac{C\omega(r)}{r^{n/p}},$$

where we have used that $\tilde{\omega}(t)/t^\sigma$ is almost increasing on $(0, r_0)$, whereas $\tilde{\omega}(t)/t^\gamma$ is almost decreasing on (r, ∞); note that the last inequality also uses that $\gamma < n/p$. All in all, this shows that it is sufficient to have

$$\frac{\omega(r)}{r^{n/p}} \leq CA^{-1}(1/r^n),$$

for all $r < r_0$. The validity of such an inequality for some $r_0 > 0$ is a consequence of (5.60). □

Example 5.6.3. We consider the following applications of Theorem 5.6.2:

(a) Let ρ satisfy (H0)–(H4) and have compact support. Assume $p, q \in (1, \infty)$ satisfy $\gamma p < n$, and

$$\liminf_{t\to 0} t^{1-\frac{1}{q}+\frac{1}{p}-\frac{1}{n}}\overline{\rho}(t)^{\frac{1}{n}} > 0. \tag{5.63}$$

By Theorem 5.6.2, $H_0^{\rho,p}(\Omega)$ is embedded into $L^q(\Omega)$. The assumptions are satisfied for the kernel (5.50) of Example 5.5.1 (a). In this case, $q = \frac{np}{n-sp}$ and $\sigma = \gamma = s$. Thus, the embedding [31, Th. 6.1] is recovered.

(b) Let ρ satisfy (H0)–(H4) and have compact support. Assume $p \in (1, \infty)$ satisfies $\gamma p < n$. Using (H3), we find that (5.63) holds with $q = \frac{np}{n-\sigma p}$. Therefore, by (a), $H_0^{\rho,p}(\Omega)$ is embedded into $L^q(\Omega)$.

(c) Consider the kernel (5.51) of Example 5.5.1 (b). Assume, in addition, that $p > 1$ with $sp < n$. Then, we may pick any $\gamma > s$ such that $\gamma p < n$. Theorem 5.6.2 shows that $H_0^{\rho,p}(\Omega)$ embeds into the Orlicz space $L^A(\Omega)$, where

$$A(t) = (t\,\mathrm{lm}(t))^{p_*},$$

$p_* = \frac{np}{n-sp}$ and $\mathrm{lm}(t)$ is the modified logarithm function

$$\mathrm{lm}(t) = \begin{cases} 1/(1-\log t) & \text{for } t \leq 1, \\ 1+\log t & \text{for } t > 1. \end{cases}$$

Indeed, checking that A is a Young function is a routine calculation. The inverse of A for $t \geq 1$ is given by

$$A^{-1}(t) = \frac{t^{1/p_*}}{W(et^{1/p_*})},$$

with W the Lambert W function (the inverse function of $r \mapsto re^r$ for $r \geq 0$). Since W behaves like log at infinity (i.e., $W(t)/\log t \to 1$ as $t \to \infty$), we can see that

$$A^{-1}(t) \geq C\frac{t^{1/p_*}}{\log t} \qquad \text{for } t \text{ large.}$$

On the other hand, the corresponding ω satisfies

$$\omega(t) = \frac{t^s}{\bar{\chi}(t)(-\log t)} \leq C\frac{t^s}{-\log t}, \qquad t \in (0, \min\{\varepsilon, 1\}).$$

The two inequalities above imply (5.60) at once, so Theorem 5.6.2 concludes the embedding of $H_0^{\rho,p}(\Omega)$ into $L^A(\Omega)$ and the validity of (5.61).

5.6.2 Embeddings into spaces of continuous functions

The opposite case to that of Theorem 5.6.2 that we present is not $\gamma p \geq n$, as one would desire, but $\sigma p > n$. In this case, we actually have embeddings into spaces of continuous functions. To show this, we first prove the following estimates for the integrals of V_ρ.

Lemma 5.6.4. *Let ρ have compact support and satisfy (H0)–(H4), $p \in (1, \infty]$ with $\sigma p > n$ and $R > 0$. Then, there is a constant $C = C(n, p, \rho, R) > 0$ such that:*

(i) For all $r \in (0, \varepsilon)$ and $|\zeta| \leq r/2$,

$$\|V_\rho\|_{L^{p'}(B_r)} \leq C\omega(r)r^{-n/p} \quad \text{and} \quad \|V_\rho - V_\rho(\cdot + \zeta)\|_{L^{p'}(B_R \setminus B_r)} \leq C\omega(r)r^{-n/p}.$$

(ii) For $|\zeta| \leq \varepsilon/3$,

$$\|V_\rho - V_\rho(\cdot + \zeta)\|_{L^{p'}(B_R)} \leq C\omega(|\zeta|)|\zeta|^{-n/p}.$$

Proof. Part (i). For the first bound we compute using (5.54)

$$\|V_\rho\|_{L^{p'}(B_r)} \leq C\left(\int_{B_r}\left(\frac{\omega(|x|)}{|x|^n}\right)^{p'} dx\right)^{1/p'} \leq \frac{C\omega(r)}{r^\sigma}\left(\int_{B_r}\frac{1}{|x|^{(n-\sigma)p'}}\,dx\right)^{1/p'} = C\omega(r)r^{-n/p},$$

where the second inequality uses that $\omega(r)/r^\sigma$ is almost increasing by (H3).

For the second inequality of the statement, we first note that we may restrict to integration over $B_{\varepsilon/2} \setminus B_r$, since V_ρ is Lipschitz continuous on $B_R \setminus B_{\varepsilon/2}$; indeed, we have

$$\|V_\rho - V_\rho(\cdot + \zeta)\|_{L^{p'}(B_R \setminus B_{\varepsilon/2})} \leq C|\zeta|\|\mathbb{1}_{B_R}\|_{L^{p'}(\mathbb{R}^n)} \leq CR^{\frac{n}{p'}}r \leq Cr^{\gamma - n/p} \leq C\omega(r)r^{-n/p},$$

where the third inequality holds because $\gamma - \frac{n}{p} < 1$, and the last one is due to (5.58). As for the integration in $B_{\varepsilon/2} \setminus B_r$, we first compute for $x \in B_{\varepsilon/2} \setminus B_r$,

$$|V_\rho(x) - V_\rho(x + \zeta)| \leq |\zeta| \int_0^1 |\nabla V_\rho(x + t\zeta)|\,dt \leq r\int_0^1 \frac{\omega(|x + t\zeta|)}{|x + t\zeta|^{n+1}}\,dt,$$

by the fundamental theorem of calculus and (5.54). Therefore,

$$\|V_\rho - V_\rho(\cdot + \zeta)\|_{L^{p'}(B_{\varepsilon/2} \setminus B_r)} \leq r\left(\int_{B_{\varepsilon/2} \setminus B_r}\left(\int_0^1 \frac{\omega(|x + t\zeta|)}{|x + t\zeta|^{n+1}}\,dt\right)^{p'} dx\right)^{\frac{1}{p'}}$$

and, by Jensen's inequality, Fubini's theorem, and the fact that $x+t\zeta \in B_\varepsilon \setminus B_{r/2}$ for each $x \in B_{\varepsilon/2} \setminus B_r$ and $\zeta \in B_{r/2}$,

$$\int_{B_{\varepsilon/2}\setminus B_r} \left(\int_0^1 \frac{\omega(|x+t\zeta|)}{|x+t\zeta|^{n+1}}\, dt \right)^{p'} dx \leq \int_0^1 \int_{B_{\varepsilon/2}\setminus B_r} \left(\frac{\omega(|x+t\zeta|)}{|x+t\zeta|^{n+1}} \right)^{p'} dx\, dt \leq \int_{B_\varepsilon \setminus B_{r/2}} \left(\frac{\omega(|x|)}{|x|^{n+1}} \right)^{p'} dx.$$

Now, by the last inequality of (5.58), for all $x \in B_\varepsilon \setminus B_{r/2}$,

$$\omega(|x|) \leq C \left(\frac{|x|}{r} \right)^{\gamma} \omega(r),$$

so

$$\left(\int_{B_\varepsilon \setminus B_{r/2}} \left(\frac{\omega(|x|)}{|x|^{n+1}} \right)^{p'} dx \right)^{\frac{1}{p'}} \leq C \frac{\omega(r)}{r^\gamma} \left(\int_{B^c_{r/2}} \frac{1}{|x|^{(n+1-\gamma)p'}} dx \right)^{\frac{1}{p'}} = C\omega(r) r^{-\frac{n}{p}-1},$$

since $(n+1-\gamma)p' > np' \geq n$. Altogether,

$$\|V_\rho - V_\rho(\cdot + \zeta)\|_{L^{p'}(B_{\varepsilon/2}\setminus B_r)} \leq C\omega(r) r^{-\frac{n}{p}}$$

and (i) is proved.

Part (ii). We have

$$\|V_\rho - V_\rho(\cdot+\zeta)\|_{L^{p'}(B_R)} \leq \|V_\rho\|_{L^{p'}(B_{2|\zeta|})} + \|V_\rho(\cdot+\zeta)\|_{L^{p'}(B_{2|\zeta|})} + \|V_\rho - V_\rho(\cdot+\zeta)\|_{L^{p'}(B_R \setminus B_{2|\zeta|})},$$

and each of the terms of the right-hand side can be estimated by part (i) and (5.58) as follows:

$$\|V_\rho\|_{L^{p'}(B_{2|\zeta|})} + \|V_\rho - V_\rho(\cdot+\zeta)\|_{L^{p'}(B_R \setminus B_{2|\zeta|})} \leq 2C\omega(2|\zeta|)|2\zeta|^{-n/p} \leq C\omega(|\zeta|)|\zeta|^{-n/p}$$

and

$$\|V_\rho(\cdot+\zeta)\|_{L^{p'}(B_{2|\zeta|})} \leq \|V_\rho\|_{L^{p'}(B_{3|\zeta|})} \leq C\omega(3|\zeta|)|3\zeta|^{-n/p} \leq C\omega(|\zeta|)|\zeta|^{-n/p},$$

which completes the proof. □

The next step is to show an embedding into spaces of continuous functions. We first make some observations and introduce the notation of the spaces of functions with a prescribed modulus of continuity. For $\alpha \in [0, \sigma)$ we define the function $\omega_\alpha : [0,\infty) \to [0,\infty)$ as

$$\omega_\alpha(t) = \begin{cases} 0 & \text{if } t = 0, \\ \omega(t)t^{-\alpha} & \text{if } t \in (0,\varepsilon), \\ \omega(\varepsilon)\varepsilon^{-\alpha} & \text{if } t \in [\varepsilon,\infty). \end{cases}$$

Thanks to (5.58), ω_α is continuous at 0. In fact, by (H2), it is continuous, and by (H0), it only vanishes at 0. Given $U \subset \mathbb{R}^n$, we define the space $C^{\omega_\alpha}(U)$ as the set of bounded functions $u : U \to \mathbb{R}$ such that there exists $C > 0$ for which

$$|u(x) - u(y)| \leq C\omega_\alpha(|x-y|), \qquad x, y \in U,$$

equipped with the seminorm

$$[u]_{C^{\omega_\alpha}(\Omega)} = \sup_{\substack{x,y\in U \\ x\neq y}} \frac{|u(x)-u(y)|}{\omega_\alpha(|x-y|)}$$

and the norm

$$\|u\|_{C^{\omega_\alpha}(U)} = \|u\|_{L^\infty(U)} + [u]_{C^{\omega_\alpha}(U)}.$$

A standard argument shows that $C^{\omega_\alpha}(U)$ is a Banach space. Moreover, as a consequence of (5.58), $C^{\omega_\alpha}(U)$ is embedded in the space $C^{0,\sigma-\alpha}(U)$ of bounded, Hölder continuous functions of exponent $\sigma - \alpha$.

With this language we prove, as a consequence of Lemma 5.6.4, the following type of Morrey inequality.

Theorem 5.6.5. *Let ρ have compact support and satisfy (H0)–(H4), $p \in (1, \infty]$ with $\sigma p > n$ and $\Omega \subset \mathbb{R}^n$ be open and bounded. Let $\alpha = n/p$. Then, any function in $H_0^{\rho,p}(\Omega)$ admits a representative that is in $C^{\omega_\alpha}(\mathbb{R}^n)$. Moreover, there exists a constant $C = C(n, p, \rho, \Omega) > 0$ such that for all continuous $u \in H_0^{\rho,p}(\Omega)$,*

$$\|u\|_{C^{\omega_\alpha}(\mathbb{R}^n)} \leq C\|D_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)}. \tag{5.64}$$

Proof. Let $u \in H_0^{\rho,p}(\Omega)$ and $R > 0$ be such that $\Omega_\delta \subset B_R$, where $\delta > 0$ is such that $\operatorname{supp}\rho = \overline{B}_\delta$. Then, by Corollary 5.5.3 and Lemma 5.6.4 (ii), we find for a.e. $x, z \in \mathbb{R}^n$ with $r := |x - z| < \varepsilon/3$ that

$$\begin{aligned} |u(x) - u(z)| &\leq \int_{B_R} |V_\rho(x - y) - V_\rho(z - y)||D_\rho u(y)|\, dy \\ &\leq \|V_\rho - V_\rho(\cdot + x - z)\|_{L^{p'}(B_{2R};\mathbb{R}^n)}\|D_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \leq C\omega_\alpha(r)\|D_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)}. \end{aligned} \tag{5.65}$$

In particular, there is a continuous representative $\bar{u} \in C(\mathbb{R}^n)$ of u, which also satisfies (5.65). Since $\bar{u} = 0$ in Ω^c, we even find that

$$|\bar{u}(x) - \bar{u}(z)| \leq C\omega_\alpha(|x - z|)\|D_\rho \bar{u}\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)}, \qquad \text{for all } x, z \in \mathbb{R}^n.$$

Taking $z \in \Omega^c$, also yields

$$|\bar{u}(x)| \leq C\|D_\rho \bar{u}\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)}, \qquad \text{for all } x \in \mathbb{R}^n,$$

which finishes the proof. □

Example 5.6.6. Let $p \in (1, \infty]$ and $\Omega \subset \mathbb{R}^n$ be open and bounded. Then, we have the inequality (5.64) for the following modulus of continuity ω_α defined for $t \in (0, \varepsilon)$:

(a) For ρ given by (5.50), as in Example 5.5.1 (a), we obtain for $sp > n$,

$$\omega_\alpha(t) = t^{s-\frac{n}{p}}.$$

This is a generalization of [31, Th. 6.3].

(b) For ρ given by (5.51), as in Example 5.5.1 (b), we obtain for $sp > n$,

$$\omega_\alpha(t) = \frac{t^{s-\frac{n}{p}}}{-\log t}.$$

(c) For ρ given by (5.52), as in Example 5.5.1 (c), we obtain for $sp > n$,

$$\omega_\alpha(t) = t^{s-\frac{n}{p}}(-\log t).$$

5.6.3 Compact embeddings

An immediate consequence of Theorem 5.6.5 is the compact inclusion of $H_0^{\rho,p}(\Omega)$ into $L^p(\mathbb{R}^n)$ in the regime $\sigma p > n$, since the space of $C^{\omega_\alpha}(\mathbb{R}^n)$ functions vanishing in Ω^c is compactly embedded into $L^p(\Omega)$. Thus, we recover part of the conclusion of Theorem 5.4.11, but with stronger assumptions. Here, we present an approach to prove compactness from $H_0^{\rho,p}(\Omega)$ into $L^p(\mathbb{R}^n)$ for any $p \in [1,\infty]$.

A similar proof to the Morrey inequality of Theorem 5.6.5, using Lemma 5.6.4 as well, yields the following bounds on translations.

Proposition 5.6.7. *Let ρ have compact support and satisfy (H0)–(H4). Let $p \in [1,\infty]$ and let $\Omega \subset \mathbb{R}^n$ be open and bounded. Then, there exists a constant $C = C(n,p,\rho,\Omega) > 0$ such that for all $|\zeta| < \varepsilon/3$,*

$$\|u - u(\cdot + \zeta)\|_{L^p(\mathbb{R}^n)} \le C\omega(|\zeta|)\|D_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \quad \textit{for all } H_0^{\rho,p}(\Omega).$$

Proof. Let $u \in H_0^{\rho,p}(\Omega)$ and $R > 0$ be such that $\Omega_\delta \subset B_R$ with $\delta > 0$ such that $\operatorname{supp}\rho = \overline{B}_\delta$. As in the first inequality of (5.65), we find for a.e. $x \in \Omega$

$$|u(x) - u(x+\zeta)| \le \int_{B_{2R}} |V_\rho(y) - V_\rho(y+\zeta)||\mathcal{G}_\rho u(x-y)|\,dy.$$

Consequently, by Minkowski's integral inequality,

$$\begin{aligned}\|u - u(\cdot+\zeta)\|_{L^p(\mathbb{R}^n)} &\le \int_{B_{2R}} |V_\rho(y) - V_\rho(y+\zeta)|\|\mathcal{G}_\rho u(\cdot - y)\|_{L^p(\mathbb{R}^n)}\,dy \\ &= \int_{B_{2R}} |V_\rho(y) - V_\rho(y+\zeta)|\,dy\|\mathcal{G}_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \le C\omega(|\zeta|)\|\mathcal{G}_\rho u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)},\end{aligned}$$

where the last inequality is due to Lemma 5.6.4 (ii). □

This result allows for an alternative proof of the compact inclusion from $H_0^{\rho,p}(\Omega)$ into $L^p(\mathbb{R}^n)$, via the Fréchet-Kolmogorov criterion (when $p < \infty$) or the Ascoli-Arzelà theorem (when $p = \infty$), as in [31, Th. 7.3]. We do not provide a proof since it is standard. Except for the cases $p = 1, \infty$, the following result requires stronger assumptions than those of Theorem 5.4.11.

Corollary 5.6.8. *Let ρ have compact support and satisfy (H0)–(H4). Let $p \in [1,\infty]$ and let $\Omega \subset \mathbb{R}^n$ be open and bounded. Then the inclusion from $H_0^{\rho,p}(\Omega)$ into $L^p(\mathbb{R}^n)$ is compact.*

5.7 Inclusion between spaces for different kernels

In this section, we compare the nonlocal Sobolev spaces induced by different kernels. We first prove the following upper bound on $\widehat{Q}_\rho$, complementing Lemma 5.4.3.

Lemma 5.7.1. *Let ρ have compact support, be differentiable outside the origin and satisfy (H0), (H1) and (H4). Then, there is a $C = C(n,\rho) > 0$ such that*

$$\widehat{Q}_\rho(\xi) \le C\frac{\overline{\rho}(1/|\xi|)}{|\xi|^n} \quad \textit{for all } \xi \in B^c_{1/\varepsilon}.$$

Proof. For $\theta > 1/\varepsilon$ and $0 < r < \frac{1}{2\theta}$ we compute, thanks to (H4),

$$\begin{aligned} f_\rho(r) - f_\rho\left(r + \frac{1}{2\theta}\right) &= h_\rho(r)\left(\frac{1}{r^{1+\gamma}} - \frac{1}{\left(r+\frac{1}{2\theta}\right)^{1+\gamma}}\right) + \left(h_\rho(r) - h_\rho\left(r+\frac{1}{2\theta}\right)\right)\frac{1}{\left(r+\frac{1}{2\theta}\right)^{1+\gamma}} \\ &\le C h_\rho\left(\frac{1}{2\theta}\right)\left(\frac{1}{r^{1+\gamma}} - \frac{1}{\left(r+\frac{1}{2\theta}\right)^{1+\gamma}}\right) + C\frac{h_\rho\left(r+\frac{1}{2\theta}\right)}{\left(r+\frac{1}{2\theta}\right)^{1+\gamma}} \\ &= C\frac{1}{(2\theta)^{1+\gamma}} f_\rho\left(\frac{1}{2\theta}\right)\left(\frac{1}{r^{1+\gamma}} - \frac{1}{\left(r+\frac{1}{2\theta}\right)^{1+\gamma}}\right) + C f_\rho\left(r+\frac{1}{2\theta}\right). \end{aligned}$$

Consequently,

$$\begin{aligned} &\left(f_\rho(r) - f_\rho\left(r+\frac{1}{2\theta}\right)\right)\sin(2\pi\theta r) \\ &\qquad \le \frac{C}{\theta^{1+\gamma}} f_\rho\left(\frac{1}{2\theta}\right)\left(\frac{1}{r^{1+\gamma}} - \frac{1}{\left(r+\frac{1}{2\theta}\right)^{1+\gamma}}\right)\sin(2\pi\theta r) + C f_\rho\left(r+\frac{1}{2\theta}\right). \end{aligned}$$

Since, as in (5.32),

$$\int_0^{\frac{1}{2\theta}}\left(\frac{1}{r^{1+\gamma}} - \frac{1}{\left(r+\frac{1}{2\theta}\right)^{1+\gamma}}\right)\sin(2\pi\theta r)\,dr = \int_0^{\frac{1}{\theta}}\frac{1}{r^{1+\gamma}}\sin(2\pi\theta r)\,dr \le C\theta\int_0^{\frac{1}{\theta}}\frac{1}{r^\gamma}dr \le C\theta^\gamma$$

and, by (H1),

$$\int_0^{\frac{1}{2\theta}} f_\rho\left(r+\frac{1}{2\theta}\right)\,dr \le \frac{1}{2\theta} f_\rho\left(\frac{1}{2\theta}\right),$$

we conclude that

$$\int_0^{\frac{1}{\theta}} f_\rho(r)\sin(2\pi\theta r)\,dr = \int_0^{\frac{1}{2\theta}}\left(f_\rho(r) - f_\rho\left(r+\frac{1}{2\theta}\right)\right)\sin(2\pi\theta r)\,dr \le \frac{C}{\theta} f_\rho\left(\frac{1}{2\theta}\right).$$

Furthermore, with integration by parts we find

$$\begin{aligned} \int_{\frac{1}{\theta}}^\infty f_\rho(r)\sin(2\pi\theta r)\,dr &= \frac{1}{2\pi\theta} f_\rho\left(\frac{1}{\theta}\right) + \frac{1}{2\pi\theta}\int_{\frac{1}{\theta}}^\infty \frac{d}{dr} f_\rho(r)\cos(2\pi|\xi| r z_1)\,dr \\ &\le \frac{C}{\theta}\left(f_\rho\left(\frac{1}{\theta}\right) + \int_{\frac{1}{\theta}}^\infty -\frac{d}{dr} f_\rho(r)\,dr\right) = \frac{C}{\theta} f_\rho\left(\frac{1}{\theta}\right) \le \frac{C}{\theta} f_\rho\left(\frac{1}{2\theta}\right). \end{aligned}$$

Summing the previous two inequalities, we find that for all $\theta > 1/\varepsilon$,

$$\int_0^\infty f_\rho(r)\sin(2\pi\theta r)\,dr \le \frac{C}{\theta} f_\rho\left(\frac{1}{2\theta}\right). \tag{5.66}$$

On the other hand, recalling (5.30), we obtain that for $\xi \neq 0$,

$$\begin{aligned} \pi|\xi|\widehat{Q}_\rho(\xi) &= \int_{\mathbb{S}^{n-1}_+} z_1 \int_0^\infty f_\rho(r)\sin(2\pi|\xi| r z_1)\,dr\,d\mathcal{H}^{n-1}(z) \\ &\le \int_{\mathbb{S}^{n-1}_+\cap\{z_1|\xi|\le 1/\varepsilon\}} \frac{1}{|\xi|\varepsilon}\int_0^\infty f_\rho(r)\sin(2\pi|\xi| r z_1)\,dr\,d\mathcal{H}^{n-1}(z) \\ &\quad + \int_{\mathbb{S}^{n-1}_+\cap\{z_1|\xi|> 1/\varepsilon\}} z_1\int_0^\infty f_\rho(r)\sin(2\pi|\xi| r z_1)\,dr\,d\mathcal{H}^{n-1}(z). \end{aligned}$$

For the first term of the right-hand side, we use the inequality $\sin t \le t$ (for $t > 0$) as well as (5.10) and the compact support of f_ρ, to obtain that when $0 < z_1|\xi| \le 1/\varepsilon$,

$$\frac{1}{|\xi|\varepsilon}\int_0^\infty f_\rho(r)\sin(2\pi|\xi|rz_1)\,dr \le C\frac{1}{|\xi|\varepsilon}|\xi|z_1 \le C\frac{1}{|\xi|}.$$

Hence, using that $\mathcal{H}^{n-1}(\mathbb{S}^{n-1}_+ \cap \{z_1|\xi| \le 1/\varepsilon\}) \le C/|\xi|^{n-1}$, we have, for $\xi \in B^c_{1/\varepsilon}$,

$$\int_{\mathbb{S}^{n-1}_+\cap\{z_1|\xi|\le 1/\varepsilon\}} \frac{1}{|\xi|\varepsilon}\int_0^\infty f_\rho(r)\sin(2\pi|\xi|rz_1)\,dr\,d\mathcal{H}^{n-1}(z) \le \frac{C}{|\xi|^n} \le C\frac{\overline{\rho}(1/|\xi|)}{|\xi|^{n-1}},$$

where in the last equality we have used the assumption $\inf_{B_\varepsilon}\rho > 0$ of (H0). For the second term, we apply (5.66) to obtain that, when $z_1|\xi| > 1/\varepsilon$,

$$\begin{aligned} z_1\int_0^\infty f_\rho(r)\sin(2\pi|\xi|rz_1)\,dr &\le \frac{C}{|\xi|}f_\rho\left(\frac{1}{2|\xi|z_1}\right) \le \frac{C}{|\xi|}f_\rho\left(\frac{1}{2|\xi|}\right) \\ &= \frac{C}{|\xi|^{n-1}}\overline{\rho}\left(\frac{1}{2|\xi|}\right) \le \frac{C}{|\xi|^{n-1}}\overline{\rho}\left(\frac{1}{|\xi|}\right), \end{aligned}$$

where we have used that f_ρ is decreasing and, for the last inequality, (H4). The proof is concluded. □

As a consequence of this lemma, we may prove embeddings of our nonlocal spaces for different kernels.

Theorem 5.7.2. *Let $\Omega \subset \mathbb{R}^n$ be open. Let ρ_1, ρ_2 have compact support and satisfy (H0)–(H1); let ρ_1 satisfy (H4). Assume $\liminf_{t\downarrow 0}\overline{\rho_2}(t)/\overline{\rho_1}(t) > 0$ and some of the following:*

(i) $p \in (1,\infty)$, ρ_1, ρ_2 satisfy (H2) and $\liminf_{t\downarrow 0} t^{n-1}\overline{\rho_1}(t) > 0$.

(ii) $p = 2$ and ρ_1 is differentiable outside the origin.

Then, the continuous inclusion $H^{\rho_2,p}_0(\Omega) \subset H^{\rho_1,p}_0(\Omega)$ holds and there is a $C = C(n,\rho_1,\rho_2) > 0$ such that

$$\|D_{\rho_1}u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \le C\|D_{\rho_2}u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \quad \text{for all } u \in H^{\rho_2,p}_0(\Omega).$$

Proof. Let $\varphi \in C^\infty_c(\Omega)$. By Proposition 5.2.6 (ii) and Lemma 5.4.3, we can write

$$\widehat{\mathcal{G}_{\rho_1}\varphi} = \frac{\widehat{Q}_{\rho_1}}{\widehat{Q}_{\rho_2}}\widehat{\mathcal{G}_{\rho_2}\varphi}.$$

Therefore, if

$$m := \frac{\widehat{Q}_{\rho_1}}{\widehat{Q}_{\rho_2}}$$

is an L^p Fourier multiplier, the result follows readily from a density argument based on Theorem 5.3.9 (i). In the case $p = 2$ (part (ii)), we only have to check that m is bounded, while for general p (part (i)), we shall see that m satisfies the hypotheses of the Mihlin-Hörmander theorem (cf. [122, Th. 6.2.7]).

To show that m is bounded, we invoke Lemma 5.7.1 for ρ_1 and Lemma 5.4.3 for ρ_2 to obtain for all $\xi \in B^c_{2/\varepsilon}$,

$$m(\xi) \le C\frac{\overline{\rho_1}(1/|\xi|)}{\overline{\rho_2}(1/|\xi|)} \le C,$$

where in the last inequality we have used $\liminf_{t\downarrow 0} \overline{\rho_2}(t)/\overline{\rho_1}(t) > 0$. The fact that m is bounded in $B_{2/\varepsilon}$ is a consequence of Lemma 5.4.3. This completes the proof under assumption (ii).

For $p \in (1, \infty)$, to check the hypotheses of the Mihlin-Hörmander theorem we need to verify that

$$|\partial^\alpha m(\xi)| \leq C_\alpha |\xi|^{-|\alpha|},$$

for all multiindices $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq n/2 + 1$. We note that this condition holds for $\xi \in B_1$, since both $\widehat{Q}_{\rho_1}$ and $\widehat{Q}_{\rho_2}$ are smooth and $\widehat{Q}_{\rho_2}$ is positive. Furthermore, the assumptions imply

$$\liminf_{t\downarrow 0} t^{n-1}\overline{\rho_1}(t) > 0, \qquad \liminf_{t\downarrow 0} t^{n-1}\overline{\rho_2}(t) > 0,$$

so we can use (5.47) as well as Lemma 5.4.10 to obtain that, for any $\beta \in \mathbb{N}^n$ and $\xi \in B_1^c$,

$$|\partial^\beta \widehat{Q}_{\rho_1}(\xi)| \leq C_\beta |\xi|^{-|\beta|} |\widehat{Q}_{\rho_1}(\xi)| \quad \text{and} \quad |\partial^\beta \widehat{Q}_{\rho_2}(\xi)| \leq C_\beta |\xi|^{-|\beta|} |\widehat{Q}_{\rho_2}(\xi)|.$$

A straightforward yet tedious calculation now shows that for $|\xi| \geq 1$

$$\left| \partial^\alpha \left(\frac{\widehat{Q}_{\rho_1}}{\widehat{Q}_{\rho_2}} \right)(\xi) \right| \leq C_\alpha |\xi|^{-|\alpha|} \left| \frac{\widehat{Q}_{\rho_1}(\xi)}{\widehat{Q}_{\rho_2}(\xi)} \right| = C_\alpha |\xi|^{-|\alpha|} m(\xi) \leq C_\alpha |\xi|^{-|\alpha|},$$

where in the last inequality we use that m is bounded. This completes the proof under assumption (i). □

Example 5.7.3. Let $\Omega \subset \mathbb{R}^n$ be open, $s \in (0,1)$ and $p \in (1, \infty)$. As in Examples 5.5.1 (a)–(c), consider ρ_1 as in (5.50), ρ_2 as in (5.51) and ρ_3 as in (5.52). Then $H_0^{\rho_2,p}(\Omega) \subset H_0^{\rho_1,p}(\Omega) \subset H_0^{\rho_3,p}(\Omega)$. Moreover, if ρ_2' is as in (5.51) but for a exponent $s' \in (0, s)$, then $H_0^{\rho_3,p}(\Omega) \subset H_0^{\rho_2',p}(\Omega)$.

Of course, changing the roles of ρ_1 and ρ_2 in Theorem 5.7.2 gives rise to a criterion of equality of spaces, which complements that of Proposition 5.3.10.

Corollary 5.7.4. *Let $\Omega \subset \mathbb{R}^n$ be open. Let ρ_1, ρ_2 have compact support and satisfy (H0)–(H1) and (H4). Assume*

$$0 < \liminf_{t\downarrow 0} \frac{\overline{\rho_2}(t)}{\overline{\rho_1}(t)} \leq \limsup_{t\downarrow 0} \frac{\overline{\rho_2}(t)}{\overline{\rho_1}(t)} < \infty$$

and some of the following:

(i) $p \in (1, \infty)$, *ρ_1, ρ_2 satisfy (H2) and* $\liminf_{t\downarrow 0} t^{n-1}\overline{\rho_1}(t) > 0$.

(ii) $p = 2$ *and ρ_1, ρ_2 are differentiable outside the origin.*

Then $H_0^{\rho_1,p}(\Omega) = H_0^{\rho_2,p}(\Omega)$ and there is a $C = C(n, p, \rho_1, \rho_2) > 0$ such that

$$\frac{1}{C} \|D_{\rho_2} u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \leq \|D_{\rho_1} u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \leq C \|D_{\rho_2} u\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} \quad \text{for all } u \in H_0^{\rho_1,p}(\Omega).$$

We finish this section by showing a partial converse of Theorem 5.4.5.

Proposition 5.7.5. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded. Let ρ have compact support, be differentiable outside the origin and satisfy (H0), (H1) and (H4). Then, the following two statements hold:*

(i) If there is a $C = C(\Omega, n, \rho) > 0$ such that

$$\|u\|_{L^2(\Omega)} \leq C \|D_\rho u\|_{L^2(\mathbb{R}^n,\mathbb{R}^n)} \quad \text{for all } u \in H_0^\rho(\Omega),$$

then $\limsup_{t\downarrow 0} t^{n-1}\overline{\rho}(t) > 0$.

(ii) If $H_0^\rho(\Omega)$ is compactly embedded into $L^2(\mathbb{R}^n)$, then $\limsup_{t\downarrow 0} t^{n-1}\overline{\rho}(t) = \infty$.

Proof. Part (i). We assume to the contrary that $\lim_{t\downarrow 0} t^{n-1}\overline{\rho}(t) = 0$. Then, we obtain thanks to Lemma 5.7.1 that the function $\lambda_\rho(\xi) = 2\pi i\xi\widehat{Q}_\rho(\xi)$ (see Proposition 5.2.6) is a bounded function with $|\lambda_\rho(\xi)| \to 0$ as $|\xi| \to \infty$. Therefore, in light of Lemma 5.2.4, we find

$$\lim_{R\to\infty} \sup_{\substack{u\in C_c^\infty(\mathbb{R}^n)\\ \|u\|_{L^2(\mathbb{R}^n)}\leq 1}} \|\widehat{\mathcal{G}_\rho u}\|_{L^2(B_R^c)} \leq \lim_{R\to\infty} \sup_{|\xi|\geq R} |\lambda_\rho(\xi)| = 0,$$

which implies by the Fréchet-Kolmogorov criterion in Fourier space (cf. [175, Theorem 3]) and the compact support of ρ, that D_ρ is a compact operator from $L^2(\Omega)$ (extended as zero in Ω^c) to $L^2(\mathbb{R}^n, \mathbb{R}^n)$. Therefore, a Poincaré inequality is not possible.

Part (ii). If $\limsup_{t\downarrow 0} t^{n-1}\overline{\rho}(t) < \infty$, then the function λ_ρ is bounded. We infer that D_ρ is a bounded operator from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n, \mathbb{R}^n)$, as a composition of the following bounded operators in L^2, initially defined for $u \in C_c^\infty(\mathbb{R}^n)$:

$$u \mapsto \widehat{u} \mapsto \lambda_\rho\widehat{u} \mapsto (\lambda_\rho\widehat{u})^\vee = \mathcal{G}_\rho u;$$

see Lemma 5.2.4. Since D_ρ is bounded from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n, \mathbb{R}^n)$, a compact embedding is not possible. □

Chapter 6

Γ-convergence involving nonlocal gradients with varying horizon: Recovery of local and fractional models

This chapter coincides with the preprint

[73] J. Cueto, C. Kreisbeck and H. Schönberger. Γ-convergence involving nonlocal gradients with varying horizon: Recovery of local and fractional models. Preprint arXiv:2404.18509, 2024.

6.1 Introduction

Nonlocal-to-local limits constitute a pivotal aspect in nonlocal modeling. They can provide a useful consistency check in confirming the compatibility of a (new) nonlocal model with a local counterpart that is covered by well-established theories. More generally speaking, the asymptotic analysis of critical parameter regimes can yield new insights about limit models from their approximations and vice versa. In this paper, we address these topics in the context of models with nonlocal gradients, which have seen a great rise in interest in the last few years, see e.g., [28, 31, 66, 92, 140, 161, 193, 208].

For a general radial kernel ρ, the nonlocal gradient of a function $u : \mathbb{R}^n \to \mathbb{R}^m$ associated to ρ is defined as

$$D_\rho u(x) = \int_{\mathbb{R}^n} \frac{u(x) - u(y)}{|x - y|} \otimes \frac{x - y}{|x - y|} \rho(x - y)\, dy \quad \text{for } x \in \mathbb{R}^n, \tag{6.1}$$

whenever this integral exists. A widely studied special case is the Riesz fractional gradient given by $D^s := D_{\rho^s}$ with $\rho^s = |\cdot|^{-(n+s-1)}$ for the fractional parameter $s \in (0, 1)$. It satisfies natural physical invariance requirements [208] and was brought to the spotlight by Shieh & Spector [193, 194], who established useful counterparts of results from classical Sobolev space theory and showed that the associated function spaces coincide with the Bessel potential spaces $H^{s,p}(\mathbb{R}^n; \mathbb{R}^m)$. This paved the way for the forthcoming works on fractional variational problems [28, 29, 140], which were proposed as alternatives to the standard models in continuum mechanics. Due to the reduced regularity requirements imposed by the fractional derivatives in comparison to the classical ones, these models admit a broader class of admissible deformations, which allows to account also for discontinuity effects, such as fracture and cavitation.

Despite the desirable properties of the Riesz fractional gradient, an intrinsic drawback from the perspective of continuum mechanical modeling is that it involves integration over the whole space $\mathbb{R}^n$, which is not suitable for models on bounded domains. This shortcoming can be resolved by considering kernel functions ρ with compact support, meaning that the range of interactions between individual points, called the horizon $\delta > 0$, is finite. In [30,31], Bellido, Cueto & Mora-Corral used finite-horizon fractional gradients as a basis to propose models of nonlocal hyperelasticity.

Note that the concept of a horizon stems originally from peridynamics [195,196], a nonlocal formulation of continuum mechanics that, in contrast to classical modeling, avoids the use of derivatives, and instead considers the interaction between individual particles that are not necessarily at an infinitesimal distance. Since its introduction in the 2000s, it has led to a vast literature, ranging from applied to theoretical contributions, see e.g. [44, 91, 154]. While energetic approaches in the context of bond-based peridynamics typically involve double-integrals, the energy functionals of nonlocal hyperelasticity, which are integrals depending on nonlocal gradients, can be interpreted in the context of state-based peridynamics [197]; one of the advantages of this framework is that it allows to model a broad range of material properties, e.g., general Poisson ratios in isotropic elastic materials, in contrast to the bond-based formulation [197].

A first step in putting the nonlocal hyperelastic models on a solid mathematical foundation is to guarantee the existence of solutions, which has been addressed in [30, 31, 72] for the case of finite-horizon gradients. The arguments rely essentially on two key techniques: a nonlocal version of the fundamental theorem of calculus [31], which is needed to prove Poincaré inequalities and compact embeddings, and a translation method established in [31, 72] that expresses the nonlocal gradient as the classical gradient composed with a convolution. These tools were recently extended in [36] to general nonlocal gradients with compactly supported radial kernels, and we utilize this in Section 6.2.3 to develop an existence theory in this broadened setting.

Our focus in this work lies on the study of the critical parameter regimes for the horizon δ, which can be seen as an important next step towards understanding and validating nonlocal hyperelasticity. While the widely used bond-based models are only able to recover a considerably restrictive class of models through a nonlocal-to-local limit passage [27, 160], we establish in this paper that the models involving nonlocal gradients are compatible with their local counterpart via a vanishing horizon limit. For a complete picture of the horizon-dependence, we also analyze the other extreme regime of diverging horizon, providing a rigorous connection with purely fractional models.

In the following, we adopt the framework for general nonlocal gradients from the recent paper [36] by Bellido, Mora-Corral & Schönberger. Starting with a fixed radial kernel ρ satisfying the hypotheses of [36] (see (H0)-(H4) in Section 6.2.2) with horizon equal to 1, i.e., $\operatorname{supp}\rho = \overline{B_1(0)}$, we apply an isotropic rescaling to obtain the kernels

$$\rho_\delta = c_\delta \rho\left(\frac{\cdot}{\delta}\right) \tag{6.2}$$

with horizon $\delta > 0$ and scaling constants $c_\delta > 0$ to be chosen suitably depending on the targeted parameter regime for δ, cf. (i) and (ii) below; examples of admissible realizations of kernel functions ρ can be found in Example 6.2.5.

Our aim in this work is to study the asymptotic behavior of the nonlocal gradients associated to the kernels ρ_δ in (6.2) for both limits of vanishing and diverging horizon, that is,

$$\text{(i) } \delta \to 0 \qquad \text{and} \qquad \text{(ii) } \delta \to \infty,$$

and prove the convergence of minimizers via Γ-convergence (cf. [49, 80]) for the corresponding families of δ-dependent functionals $\mathcal{F}_\delta$. They consist of vectorial integrals of the form

$$\mathcal{F}_\delta(u) := \int_{\Omega_\delta} f(x, D_{\rho_\delta} u)\, dx, \tag{6.3}$$

where the involved quantities are given as follows: The set $\Omega \subset \mathbb{R}^n$ is a bounded domain, $\Omega_\delta = \Omega + B_\delta(0)$ is its expansion by the horizon parameter, and the integrand $f : \mathbb{R}^n \times \mathbb{R}^{n\times m} \to \mathbb{R}$ is assumed to have p-growth for $p \in (1,\infty)$ and to be quasiconvex in its second argument. The space of admissible functions for (6.3) is $H_0^{\rho,p,\delta}(\Omega;\mathbb{R}^m)$, the natural nonlocal Sobolev space associated to the gradient D_{ρ_δ} with a zero complementary-value conditions, meaning that the functions vanish in Ω^c, see Section 6.2 for more details. Let us now give a brief overview of our findings on the two limit passages $\delta \to 0$ and $\delta \to \infty$.

(i) Localization via shrinking horizon $\boldsymbol{\delta \to 0}$. With the scaling factors $c_\delta = \delta^{-n}$ in (6.2), which preserves the mass of the kernel ρ_δ, we confirm for each $u \in W^{1,p}(\mathbb{R}^n;\mathbb{R}^m)$ the convergence of the nonlocal gradients to the classical one, precisely,

$$D_{\rho_\delta} u \to \nabla u \quad \text{in } L^p(\mathbb{R}^n;\mathbb{R}^{m\times n}) \text{ as } \delta \to 0,$$

see Lemma 6.3.1; for sufficiently smooth functions, the convergence is uniform and the optimal rate of convergence is explicitly determined by δ^2.

As our main result within (i), Theorem 6.3.7 states the Γ-convergence of $(\mathcal{F}_\delta)_\delta$ from (6.3) with respect to the strong L^p-topology as $\delta \to 0$ to the limit functional $\mathcal{F}_0$ given by

$$\mathcal{F}_0(u) = \int_\Omega f(x,\nabla u)\,dx \quad \text{for } u \in W_0^{1,p}(\Omega;\mathbb{R}^m),$$

and provides along with this, the corresponding equi-coercivity of $(\mathcal{F}_\delta)_\delta$; the latter constitutes the major novelty of Theorem 6.3.7, as explained below. Consequently, the minimizers of $\mathcal{F}_\delta$, which exist by Theorem 6.2.11, converge (up to subsequences) in L^p to a minimizer of $\mathcal{F}_0$.

The key technical ingredient for our proof of equi-coercivity for $(\mathcal{F}_\delta)_\delta$ is the estimate (6.4), which can be seen as an enhanced Poincaré-type inequality. The proof of (6.4) uses the isotropic scaling from (6.2) to identify the dependence of the Fourier symbol of D_{ρ_δ} on δ. Together with the results on the Fourier symbol of D_ρ for a fixed kernel [36], we then deduce from the Mihlin-Hörmander theorem that the spaces $H^{\rho,p,\delta}(\Omega;\mathbb{R}^m)$ do not change with $\delta > 0$ and that there is a δ-independent constant $C > 0$ such that

$$\|u\|_{H^{\sigma,p}(\mathbb{R}^n;\mathbb{R}^m)} \le C\|D_{\rho_\delta} u\|_{L^p(\mathbb{R}^n;\mathbb{R}^{m\times n})} \quad \text{for all } u \in H^{\rho,p,\delta}(\Omega;\mathbb{R}^m) \text{ and } \delta \in (0,1], \tag{6.4}$$

where $\sigma > 0$ is related to the kernel ρ, see Theorem 6.3.3 and Corollary 6.3.4.

To set our contribution into context with the existing literature, we mention that closely related localization results for nonlocal gradients in various relevant topologies can be found in [161] or deduced from the results in the recent paper [16] by Arroyo-Rabasa, which addresses more general nonlocal first-order linear operators. Mengesha & Spector in [161, Theorem 1.7] also present a first Γ-convergence result for scalar and convex variational problems in their setting. Beyond the fact that we consider more general quasiconvex integrands in the vectorial case, the only minor difference with our set-up (see (6.1) and (6.3)) lies in the definition of the nonlocal gradient, where they consider interactions only between points within the domain Ω. Accordingly, some of our arguments regarding the liminf-inequality and the construction of recovery sequences share similarities with [161]. However, in contrast to our work, [161] does not contain any compactness results (uniformly in δ) or equi-coercivity results, and, therefore, cannot guarantee the existence or the convergence of minimizers for the involved integral functionals. On the other hand, in a different setting of non-symmetric half-space nonlocal gradients for the case $p = 2$, such compactness results have recently been obtained uniformly in δ [128].

(ii) Fractional models via diverging horizon $\boldsymbol{\delta \to \infty}$. For the regime of large horizons, we identify the scaling factors $c_\delta = \overline{\rho}(1/\delta)^{-1}$ in (6.2), with $\overline{\rho}$ the radial representation of ρ, and assume additionally that the kernels ρ_δ converge pointwise to a function ρ_∞; the scaling factors ensure that ρ_∞ is equal to 1 on the unit sphere. It turns out that the limit kernel ρ_∞ will always be a fractional kernel, that is,

$$\rho_\infty = |\cdot|^{-(n+s_\infty-1)} \tag{6.5}$$

with some $s_\infty \in (0,1)$ characteristic for ρ, see Lemma 6.4.2. Even though this observation may be surprising, given that the strong singularity of the kernel at the origin will generally not be of a fractional type, it follows because of the fact that the radial representation of ρ_∞ gains multiplicativity through the limit process, and is thus, a power function.

In view of (6.5), we then deduce in Proposition 6.4.6 the convergence of the nonlocal gradients to a Riesz fractional gradient. Precisely, it holds for $u \in W^{1,p}(\mathbb{R}^n;\mathbb{R}^n)$ that

$$D_{\rho_\delta} u \to D^{s_\infty} u \quad \text{in } L^p(\mathbb{R}^n;\mathbb{R}^{m\times n}) \text{ as } \delta \to \infty. \tag{6.6}$$

Using a similar strategy as in (i), we can specify the dependence of the Fourier symbol of D_{ρ_δ} on δ and show the analogue of (6.4) for large δ, see Proposition 6.4.7. This facilitates along with (6.6) the proof of our main theorem on Γ-convergence and equi-coercivity of the family $(\mathcal{F}_\delta)_\delta$ (cf. (6.3)) as $\delta \to \infty$; explicitly, Theorem 6.4.10 yields the Γ-limit $\Gamma(L^p)\text{-}\lim_{\delta\to\infty} \mathcal{F}_\delta = \mathcal{F}_\infty$ with

$$\mathcal{F}_\infty(u) = \int_{\mathbb{R}^n} f(x, D^{s_\infty} u)\, dx \qquad \text{for } u \in H_0^{s_\infty,p}(\Omega;\mathbb{R}^m). \tag{6.7}$$

In particular, we have that the minimizers of $\mathcal{F}_\delta$ converge (up to subsequence) in L^p to a minimizer of a variational integral depending on Riesz fractional gradients.

Note that the resulting limit objects, that is, fractional functionals of the type (6.7), have been well-studied in the last years under different assumptions on the integrand, see e.g. [28,140,189,193]. The aspects addressed include weak lower semicontinuity, relaxation, existence of minimizers, and Euler-Lagrange equations.

The prototypical example that illustrates our results is a truncated version of the Riesz fractional kernel, that is,

$$\rho = \frac{w}{|\cdot|^{n+s-1}} \qquad \text{for } s \in (0,1),$$

where $w : \mathbb{R}^n \to [0,\infty)$ is a suitable smooth, radial cut-off function compactly supported in the closed unit ball of $\mathbb{R}^n$, cf. Example 6.2.5 a). Applying the two scaling choices of (i) and (ii) gives the scaled kernels

$$\rho_\delta = \delta^{s-1}\frac{w(\,\cdot\,/\delta)}{|\cdot|^{n+s-1}} \quad \text{and} \quad \rho_\delta = \overline{w}(1/\delta)^{-1}\frac{w(\,\cdot\,/\delta)}{|\cdot|^{n+s-1}},$$

respectively, where $\overline{w}$ denotes the radial representation of w. Both these kernels are supported in the ball $\overline{B_\delta(0)}$ of radius δ around the origin and give rise to the finite-horizon fractional gradients D^s_δ studied in [26, 30, 31, 72, 141]; we remark that in those references the dependence of the kernel on δ is not made explicit, since the horizon was always considered fixed. As a consequence of the results in this paper, variational problems involving the gradients D^s_δ are confirmed to approximate their local analogue with classical gradients in the limit of vanishing horizon. As $\delta \to \infty$, we justify the intuitive connection with the fractional case, which reflects an infinite range of interaction.

Let us close the introduction by briefly pointing out some interesting further connections for broader context. We observe that for fractional gradients, localization also occurs by letting the fractional index s tend to 1. This was demonstrated for the Riesz fractional gradient D^s in [29] and for its finite-horizon version D^s_δ in [72, 141]. Among the first works on nonlocal-to-local limit passages were those in the context of bond-based peridynamics [8, 35, 47, 158], which successfully recover various classical models. However, as shown in [27, 160], there are classical energies that cannot be obtained from double-integral functionals as δ goes to 0, as opposed to the setting of this paper. Finally, the limit of diverging horizon is less common in the literature. Nevertheless, the convergence of finite-horizon versions of the fractional p-Laplacian has been established in [37,38] for both $\delta \to 0$ and $\delta \to \infty$, recovering the classical and fractional p-Laplacian, respectively.

The manuscript is organized as follows. After introducing in Section 6.2 the nonlocal gradients and associated function spaces that we are going to work with, we specify the required conditions on the kernel ρ and collect some technical tools and preliminary results. The core of the paper are Sections 6.3 and 6.4, where we address the limit analysis of $\delta \to 0$ and $\delta \to \infty$ to recover the classical and fractional models, respectively. These two sections, which are each presented in a self-contained way, share a parallel structure: First, showing the convergence of the varying horizon nonlocal gradients, then, compactness statements uniformly in δ, and, finally, the Γ-convergence of the energy functionals in (6.3).

6.2 Preliminaries

6.2.1 Notation

We write $|x| = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$ for the Euclidean norm of a vector $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ and $|A|$ for the Frobenius norm of a matrix $A \in \mathbb{R}^{m\times n}$. The ball centered at $x \in \mathbb{R}^n$ and with radius $r > 0$ is denoted by $B_r(x) = \{y \in \mathbb{R}^n : |x - y| < r\}$ and the distance between $x \in \mathbb{R}^n$ and a set $E \subset \mathbb{R}^n$ is written as $\operatorname{dist}(x, E)$. For an open set $\Omega \subset \mathbb{R}^n$ and $\delta > 0$, we define

$$\Omega_\delta := \Omega + B_\delta(0) = \{x \in \mathbb{R}^n \,:\, \operatorname{dist}(x, \Omega) < \delta\} \tag{6.8}$$

and set $\Omega_{-\delta} := \{x \in \Omega \,:\, \operatorname{dist}(x, \Omega^c) > \delta\}$. The complement of a set $E \subset \mathbb{R}^n$ is indicated by $E^c := \mathbb{R}^n \setminus E$, its closure by $\overline{E}$, and its boundary by ∂E. We take

$$\mathbb{1}_E(x) = \begin{cases} 1 & \text{for } x \in E, \\ 0 & \text{otherwise,} \end{cases} \qquad x \in \mathbb{R}^n,$$

to be the indicator function of a set $E \subset \mathbb{R}^n$. The extension of a function $u : E \to \mathbb{R}$ to $\mathbb{R}^n$ as zero is sometimes explicitly denoted as $\mathbb{1}_E\, u$. For a function $u : \mathbb{R}^n \to \mathbb{R}$, we denote its support by $\operatorname{supp} u$, and, if u is Lipschitz continuous, its Lipschitz constant by $\operatorname{Lip}(u)$.

For $U \subset \mathbb{R}^n$ open, we adopt the standard notation for the space of smooth functions with compact support $C_c^\infty(U)$, the Lebesgue space $L^p(U)$ and the Sobolev space $W^{1,p}(U)$ with $p \in [1, \infty]$. The spaces can be extended componentwise to vector-valued functions; the target space is explicitly mentioned in the notation, like, for example, $L^p(U; \mathbb{R}^m)$. We use the usual multi-index notation for partial derivatives ∂^α with $\alpha \in \mathbb{N}_0^n$. Our convention for the Fourier transform of $f \in L^1(\mathbb{R}^n)$ is

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i x\cdot\xi}\, dx, \quad \xi \in \mathbb{R}^n,$$

see e.g., [122] for more details.

For real functions, we use the monotonicity properties of being increasing and decreasing in the non-strict sense. A function $f : \mathbb{R} \to \mathbb{R}$ is called almost decreasing if there is a $C > 0$ such that $f(t) \geq Cf(s)$ for $t \leq s$, and an analogous definition holds for almost increasing. For a radial function $p : \mathbb{R}^n \to \mathbb{R}$, we denote its radial representation by $\overline{p} : [0, \infty) \to \mathbb{R}$, i.e., $p(x) = \overline{p}(|x|)$ for $x \in \mathbb{R}^n$, and call p radially decreasing or increasing, if its radial representation is decreasing or increasing, respectively.

Finally, throughout the manuscript, we use generic constants, which may change from line to line without renaming.

6.2.2 Nonlocal gradients

We now introduce the key elements of our setting, that is, the nonlocal gradients for general kernels ρ as recently studied in [36]; for related work see also [92,93,102], as well as [30,31,72] on the special case of finite-horizon fractional gradients.

Assume throughout that $\rho : \mathbb{R}^n \setminus \{0\} \to [0,\infty)$ is a radial kernel such that

(H0) $\inf_{\overline{B_\varepsilon(0)}} \rho > 0$ for some $\varepsilon > 0$ and $\rho \min\{1, |\cdot|^{-1}\} \in L^1(\mathbb{R}^n)$.

Under this condition, the kernel gives rise to an associated nonlocal gradient.

Definition 6.2.1 (Nonlocal gradient). *The nonlocal gradient $D_\rho\varphi : \mathbb{R}^n \to \mathbb{R}^n$ for $\varphi \in C_c^\infty(\mathbb{R}^n)$ is given by*

$$D_\rho\varphi(x) = \int_{\mathbb{R}^n} \frac{\varphi(x) - \varphi(y)}{|x-y|} \frac{x-y}{|x-y|} \rho(x-y)\, dy, \quad x \in \mathbb{R}^n.$$

We collect here some key properties of D_ρ that will be used later on. First of all, with $\overline{\rho} : (0,\infty) \to [0,\infty)$ the radial representation of ρ, i.e., $\rho = \overline{\rho}(|\cdot|)$, the nonlocal gradient can be written as the convolution of the classical gradient with the locally integrable function

$$Q_\rho(x) := \int_{|x|}^\infty \frac{\overline{\rho}(r)}{r}\, dr, \quad x \in \mathbb{R}^n \setminus \{0\},$$

that is,

$$D_\rho\varphi = Q_\rho * \nabla\varphi = \nabla(Q_\rho * \varphi) \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^n), \tag{6.9}$$

see [36, Propositions 2.6]. When $\rho \in L^1(\mathbb{R}^n)$, then also $Q_\rho \in L^1(\mathbb{R}^n)$ and one obtains after taking the Fourier transform that

$$\widehat{D_\rho\varphi}(\xi) = 2\pi i\xi \widehat{Q}_\rho(\xi)\widehat{\varphi}(\xi) \quad \text{for } \xi \in \mathbb{R}^n, \tag{6.10}$$

see [36, Propositions 2.5 (*iii*) and 2.6].

An important property of the nonlocal gradient is the presence of a duality relation with the nonlocal divergence, as conveyed by the following integration by parts formula, cf. [36, Proposition 3.2].

Lemma 6.2.2 (Integration by parts, [36]). *Let $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\psi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$. Then*

$$\int_{\mathbb{R}^n} D_\rho\varphi \cdot \psi\, dx = -\int_{\mathbb{R}^n} \varphi \operatorname{div}_\rho \psi\, dx, \tag{6.11}$$

where

$$\operatorname{div}_\rho \psi(x) := \int_{\mathbb{R}^n} \frac{\psi(x) - \psi(y)}{|x-y|} \cdot \frac{x-y}{|x-y|} \rho(x-y)\, dy.$$

Note that the integration over the whole space in (6.11) is sufficient for our purpose, since we will be working mainly with compactly supported functions; an alternative version of integration by parts over a bounded domain giving rise to boundary terms was recently proven in [26].

The previous lemma motivates a distributional definition of nonlocal gradients.

Definition 6.2.3 (Weak nonlocal gradients). *Let $u \in L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$. We say that $V \in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^n)$ is the weak nonlocal gradient of u, and write $D_\rho u = V$, if*

$$\int_{\mathbb{R}^n} V \cdot \psi\, dx = -\int_{\mathbb{R}^n} u \operatorname{div}_\rho \psi\, dx \quad \textit{for all } \psi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n).$$

In analogy to classical Sobolev spaces, one introduces for $p \in (1,\infty)$ the ρ-nonlocal Sobolev spaces as

$$H^{\rho,p}(\mathbb{R}^n) := \{u \in L^p(\mathbb{R}^n) \,:\, D_\rho u \in L^p(\mathbb{R}^n;\mathbb{R}^n)\},$$

endowed with the norm

$$\|u\|_{H^{\rho,p}(\mathbb{R}^n)} := \|u\|_{L^p(\mathbb{R}^n)} + \|D_\rho u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)}, \tag{6.12}$$

see [36, Definition 3.4]. Note that these spaces can be equivalently characterized as the closure of $C_c^\infty(\mathbb{R}^n)$ under the norm in (6.12) in light of [36, Theorem 3.9 (i)]. Additionally, we define for an open $\Omega \subset \mathbb{R}^n$ the subspaces

$$H_0^{\rho,p}(\Omega) := \overline{C_c^\infty(\Omega)}^{\|\cdot\|_{H^{\rho,p}(\mathbb{R}^n)}},$$

where the elements of $C_c^\infty(\Omega)$ are interpreted as extended to $\mathbb{R}^n$ by zero. If Ω is a bounded Lipschitz domain, $H_0^{\rho,p}(\Omega)$ agrees with the complementary-value space of the functions in $H^{\rho,p}(\mathbb{R}^n)$ that are zero in Ω^c, see [36, Theorem 3.9 (iii)]. Prescribed complementary values can be viewed as the nonlocal analogue of Dirichlet boundary conditions in the local setting.

Example 6.2.4 (Riesz fractional gradient and Bessel potential spaces). The special choice of kernel function

$$\rho^s := \frac{1}{|\cdot|^{n+s-1}} \quad \text{with } s \in (0,1) \tag{6.13}$$

gives rise to the Riesz s-fractional gradient $D^s := D_{\rho^s}$, given for $\varphi \in C_c^\infty(\mathbb{R}^n)$ by

$$D^s\varphi(x) = \int_{\mathbb{R}^n} \frac{\varphi(x)-\varphi(y)}{|x-y|^{n+s}}\frac{x-y}{|x-y|}\,dy, \quad x \in \mathbb{R}^n,$$

cf. [193,194]. Commonly, D^s features a normalization constant $c_{n,s}$, which we omit here for the sake of a cleaner presentation in Section 6.4. The associated nonlocal Sobolev space $H^{\rho^s,p}(\mathbb{R}^n)$ coincides with the Bessel potential space $H^{s,p}(\mathbb{R}^n)$ as shown in [193, Theorem 1.7]. A property we will often exploit is that

$$H^{s,p}(\mathbb{R}^n) \text{ is compactly embedded into } L^p_{\mathrm{loc}}(\mathbb{R}^n), \tag{6.14}$$

see e.g., [194, Theorem 2.2] or [28, Theorem 2.3]. Moreover, we set $H_0^{s,p}(\Omega) = \{u \in H^{s,p}(\mathbb{R}^n) : u = 0 \text{ a.e. in } \Omega^c\}$.

Besides the previously introduced hypothesis (H0) on the kernel ρ, used for the definition of the nonlocal gradient, we require a few more properties in order to have a wider variety of technical tools, such as compact embeddings and Poincaré inequalities, at our disposal. In accordance with [36] (see also [36, Remark 4.1]), we make the following assumptions:

Let ε be as in (H0), $\nu > 0$ and $0 < \sigma \le \gamma < 1$.

- (H1) The function $f_\rho : (0,\infty) \to \mathbb{R},\ r \mapsto r^{n-2}\overline{\rho}(r)$ is decreasing on $(0,\infty)$ and $r \mapsto r^\nu f_\rho(r)$ is decreasing on $(0,\varepsilon)$;
- (H2) f_ρ is smooth outside the origin and for every $k \in \mathbb{N}$ there exists a $C_k > 0$ with

$$\left|\frac{d^k}{dr^k} f_\rho(r)\right| \le C_k \frac{f_\rho(r)}{r^k} \quad \text{for } r \in (0,\varepsilon);$$

- (H3) the function $r \mapsto r^{n+\sigma-1}\overline{\rho}(r)$ is almost decreasing on $(0,\varepsilon)$;

(H4) the function $r \mapsto r^{n+\gamma-1}\overline{\rho}(r)$ is almost increasing on $(0, \varepsilon)$.

Most of the time, we will not work directly with these hypotheses, but instead make use of the results and tools proven in [36]; we refer to that paper for a more detailed discussion of the assumptions (H0)-(H4). The Riesz potential kernel from (6.13) satisfies all these properties, as one can easily check. Beyond that, we list here a few examples with compactly supported kernels from [36, Example 5.1] that fit into the setting. These will be revisited also in the later sections to illustrate our findings.

Example 6.2.5 (Selected kernel functions ρ). Let $w \in C_c^\infty(\mathbb{R}^n)$ be a non-negative radial function with $w(0) > 0$.

a) Let $s \in (0, 1)$ and suppose that $w/|\cdot|^{1+s}$ is radially decreasing. Then,

$$\rho(x) = \frac{w(x)}{|x|^{n+s-1}}, \quad x \in \mathbb{R}^n \setminus \{0\},$$

satisfies (H0)-(H4) with $\sigma = \gamma = s$. The associated nonlocal gradient D_ρ is referred to as a finite-horizon fractional gradient. In fact, it holds that $H^{\rho,p}(\mathbb{R}^n) = H^{s,p}(\mathbb{R}^n)$ with equivalent norms by [36, Proposition 3.10].

b) Let $s \in (0, 1)$ and $\kappa \in \{-1, 1\}$. If $\operatorname{supp}(w) \subset \overline{B_1(0)}$ and $w \log^\kappa(1/|\cdot|)/|\cdot|^{1+s}$ is radially decreasing, then the kernel function given by

$$\rho(x) = \frac{w(x) \log^\kappa(1/|x|)}{|x|^{n+s-1}}, \quad x \in \mathbb{R}^n \setminus \{0\},$$

satisfies (H0)-(H4) with $\sigma = s$ and any $\gamma \in (s, 1)$ if $\kappa = 1$ and with any $\sigma \in (0, s)$ and $\gamma = s$ if $\kappa = -1$.

c) Consider a smooth function $s : [0, \infty) \to (0, 1)$ and let $w/|\cdot|^{1+s(|\cdot|)}$ be radially decreasing. Then,

$$\rho(x) = \frac{w(x)}{|x|^{n+s(|x|)-1}}, \quad x \in \mathbb{R}^n \setminus \{0\},$$

is a kernel with spatially varying fractional parameter satisfying (H0)-(H4) with $\sigma = \min_{[0,\varepsilon]} s$ and $\gamma = \max_{[0,\varepsilon]} s$ for any $\varepsilon > 0$.

The following auxiliary result from [36, Lemma 4.3, 4.10 and 7.1] will be exploited in Sections 6.3 and 6.4 to prove compactness results uniformly in the horizon parameter. It provides bounds on the Fourier transform of Q_ρ and its derivatives in terms of the radial representation of ρ.

Lemma 6.2.6 (Estimates on $\widehat{Q}_\rho$ and its derivatives, [36]). *Let $\rho : \mathbb{R}^n \setminus \{0\} \to [0, \infty)$ be a radial kernel with compact support satisfying (H0)-(H4). Then $\widehat{Q}_\rho$ is smooth, positive, and there exists a constant $C > 0$ such that*

$$\frac{1}{C}\frac{\bar{\rho}(1/|\xi|)}{|\xi|^n} \leq \widehat{Q}_\rho(\xi) \leq C\frac{\bar{\rho}(1/|\xi|)}{|\xi|^n} \quad \text{for all } |\xi| \geq 1/\varepsilon. \tag{6.15}$$

Moreover, for every $\alpha \in \mathbb{N}_0^n$, one has

$$\left|\partial^\alpha \widehat{Q}_\rho(\xi)\right| \leq C_\alpha |\xi|^{-|\alpha|} \left|\widehat{Q}_\rho(\xi)\right| \quad \text{for all } \xi \neq 0 \tag{6.16}$$

with constants $C_\alpha > 0$.

Finally, Poincaré-type inequalities will be indispensible tools for our analysis. We present here a particular consequence of [36, Theorem 4.11], that suffices for our setting.

Lemma 6.2.7 (Poincaré inequalities and compact embedding, [36]). *Let $\Omega \subset \mathbb{R}^n$ be open and bounded and suppose that the radial kernel function ρ satisfies (H0)-(H4) and has compact support. Then there is a $C > 0$ such that*

$$\|u\|_{L^p(\mathbb{R}^n)} \leq C\|D_\rho u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \quad \text{for all } u \in H_0^{\rho,p}(\Omega), \tag{6.17}$$

and $H_0^{\rho,p}(\Omega)$ is compactly embedded into $L^p(\mathbb{R}^n)$.

In fact, comparing ρ as in the previous lemma with the kernel from Example 6.2.5 a) with $s = \sigma$ leads to a stronger estimate that will be utilized several times in what follows. Precisely, by using (H3) and [36, Theorem 7.2], we find that there is a $C > 0$ such that

$$\|u\|_{H^{\sigma,p}(\mathbb{R}^n)} \leq C\|D_\rho u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \quad \text{for all } u \in H_0^{\rho,p}(\Omega). \tag{6.18}$$

Remark 6.2.8 (Relaxed assumptions on ρ). Note that according to [36, Proposition 3.10], there is an equivalence between the function spaces and Poincaré inequalities associated to kernels that agree around the origin. Hence, (6.17) and (6.18) hold even when the smoothness in (H1) only holds locally, or when the assumption of ρ having compact support is dropped. For example, one could replace the function w in Example 6.2.5 a) by an indicator function $\mathbb{1}_{B_\delta(0)}$ with $\delta > 0$ or by an exponentially decaying function $e^{-\alpha|\cdot|}$ with $\alpha > 0$, which leads to the truncated and tempered fractional kernel of [92, Examples 2 and 3], respectively. △

6.2.3 Existence theory for nonlocal variational problems

Let us address next the solvability of vectorial variational problems involving the nonlocal gradients as introduced in the previous section. Besides being of general interest, the existence statement of Theorem 6.2.11 is needed below to conclude the convergence of minimizers for the variational problems with varying horizon in Sections 6.3 and 6.4. We remark that the results presented here are new in this generality, but can be derived by following closely the techniques of [72], where the direct method in the calculus of variations is applied to the special case of functionals depending on finite-horizon fractional gradients. The adaptation of the proofs is straightforward and left to the reader.

Throughout this section, we assume that $p \in (1, \infty)$, $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain, and the kernel ρ satisfies (H0)-(H4) and has compact support. The following result, which allows us to translate the nonlocal gradients into classical gradients, can be proven by extending (6.9) via density as in [72, Theorem 2 (i)].

Lemma 6.2.9 (From nonlocal to local gradients). *The linear map $\mathcal{Q}_\rho : H^{\rho,p}(\mathbb{R}^n) \to W^{1,p}(\mathbb{R}^n)$, $u \mapsto Q_\rho * u$ is bounded and it holds for all $u \in H^{\rho,p}(\mathbb{R}^n)$ that*

$$D_\rho u = \nabla(\mathcal{Q}_\rho u).$$

Another ingredient is the strong convergence of nonlocal gradients in the complement of Ω, which follows as in [72, Lemma 3] by utilizing the compact embedding of $H_0^{\rho,p}(\Omega)$ into $L^p(\mathbb{R}^n)$ (see Lemma 6.2.7) and the Leibniz rule in [36, Lemma 3.8].

Lemma 6.2.10 (Strong convergence in the complement). *Let $(u_j)_j \subset H_0^{\rho,p}(\Omega)$ be a sequence that converges weakly to u in $H^{\rho,p}(\mathbb{R}^n)$. Then, for any $\eta > 0$ it holds that*

$$D_\rho u_j \to D_\rho u \quad \text{in } L^p((\Omega_\eta)^c;\mathbb{R}^n) \text{ as } j \to \infty,$$

recalling the definition $\Omega_\eta = \Omega + B_\eta(0)$, see (6.8).

With these two technical tools and the Poincaré inequality from Lemma 6.2.7 at hand, one can argue as in the sufficiency part of [72, Theorem 5] and [72, Corollary 2] to obtain the existence of minimizers for vectorial variational problems with quasiconvex integrands.

Theorem 6.2.11 (Existence of minimizers). *Let $\delta > 0$ be such that* $\operatorname{supp}\rho = \overline{B_\delta(0)}$ *and let $f : \Omega_\delta \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory integrand such that*

$$c|A|^p - C \leq f(x,A) \leq C(1+|A|^p) \quad \textit{for a.e. } x \in \Omega_\delta \textit{ and all } A \in \mathbb{R}^{m\times n}.$$

If $A \mapsto f(x,A)$ is quasiconvex for a.e. $x \in \Omega$, then the functional

$$\mathcal{F} : H_0^{\rho,p}(\Omega;\mathbb{R}^m) \to \mathbb{R}, \quad \mathcal{F}(u) := \int_{\Omega_\delta} f(x, D_\rho u)\,dx \tag{6.19}$$

admits a minimizer.

Note that taking $\delta > 0$ such that $\operatorname{supp}\rho = \overline{B_\delta(0)}$ ensures that $D_\rho u$ is zero in Ω_δ^c for all $u \in H_0^{\rho,p}(\Omega)$. Hence, the functional $\mathcal{F}$ in (6.19) defined as an integral over the bounded set Ω_δ captures all the non-trivial parts of D_ρ.

Quasiconvexity, which is well-known to characterize the weak lower semicontinuity of integral functionals in the classical case [75, 165], is indeed the natural convexity notion also in the context of variational integrals depending on nonlocal gradients. This observation can be seen as a generalization of [140, Theorem 1.1] and [72, Theorem 5] and relies on Lemma 6.2.9 and the following inverse translation operator.

Lemma 6.2.12 (From local to nonlocal gradients). *There is a bounded linear operator $\mathcal{P}_\rho : W^{1,p}(\mathbb{R}^n) \to H^{\rho,p}(\mathbb{R}^n)$ such that $\mathcal{P}_\rho = (\mathcal{Q}_\rho)^{-1}$. In particular, for all $v \in W^{1,p}(\mathbb{R}^n)$ we have*

$$\nabla v = D_\rho(\mathcal{P}_\rho v).$$

Proof. We define the operator

$$\mathcal{P}_\rho : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n), \quad \mathcal{P}_\rho v := \left(\widehat{v}/\widehat{Q}_\rho\right)^\vee,$$

which is well-defined given that $1/\widehat{Q}_\rho$ is smooth and has polynomially bounded derivatives by Lemma 6.2.6 and (H3). It is also clear that $\mathcal{P}_\rho = (\mathcal{Q}_\rho)^{-1}$ and $D_\rho \circ \mathcal{P}_\rho = \nabla$ on the space $\mathcal{S}(\mathbb{R}^n)$, so it is sufficient to prove that $\mathcal{P}_\rho$ extends to a bounded operator from $W^{1,p}(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$. By Lemma 6.2.6, (H3) and the Mihlin-Hörmander theorem (cf. e.g. [122, Theorem 6.2.7]), it can be verified that $\langle\cdot\rangle^{\sigma-1}/\widehat{Q}_\rho$ is an L^p-multiplier, where $\langle\xi\rangle := \sqrt{1+|\xi|^2}$ for $\xi \in \mathbb{R}^n$. Hence, arguing as in [141, Section 2.3], we find that $\mathcal{P}_\rho$ extends to a bounded operator from $H^{1-\sigma,p}(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$ and, in particular, it is also bounded from $W^{1,p}(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$. □

The following can now be proven as in [72, Remark 8], given that we have both translation operators $\mathcal{Q}_\rho, \mathcal{P}_\rho$, the Leibniz rule from [36, Lemma 3.8], and the compact embedding of $H_0^{\rho,p}(\Omega)$ into $L^\infty(\mathbb{R}^n)$ for $p > n/\sigma$ (cf. [36, Theorem 6.5]).

Corollary 6.2.13 (Nonlocal representation of quasiconvexity). *Let $\delta > 0$ be such that* $\operatorname{supp}\rho = \overline{B_\delta(0)}$. *A continuous function $h : \mathbb{R}^{m\times n} \to \mathbb{R}$ is quasiconvex if and only if*

$$f(A) \leq \frac{1}{|\Omega_\delta|}\int_{\Omega_\delta} f(A + D_\rho u)\,dx \quad \textit{for all } A \in \mathbb{R}^{m\times n} \textit{ and } u \in H_0^{\rho,\infty}(\Omega;\mathbb{R}^m),$$

with $H_0^{\rho,\infty}(\Omega;\mathbb{R}^m) := \{u \in L^\infty(\mathbb{R}^n;\mathbb{R}^m) \,:\, D_\rho u \in L^\infty(\mathbb{R}^n;\mathbb{R}^{m\times n}),\ u = 0 \text{ a.e. in } \Omega^c\}$.

6.2.4 Scaled kernels

We introduce here the setting and notations for varying horizon nonlocal gradients obtained via scaling of a fixed nonlocal gradient, as they will be used in the limits of vanishing and diverging horizon in Sections 6.3 and 6.4.

Our starting point is a radial kernel ρ that satisfies (H0)-(H4) and is normalized in the sense that

$$\operatorname{supp} \rho = \overline{B_1(0)} \quad \text{and} \quad \int_{\mathbb{R}^n} \rho \, dx = n. \tag{6.20}$$

One can then compute that

$$\int_{\mathbb{R}^n} Q_\rho \, dx = 1, \quad \text{or equivalently,} \quad \widehat{Q}_\rho(0) = 1. \tag{6.21}$$

Notice that the kernels from Example 6.2.5 can all be rescaled and normalized to satisfy (6.20).

The rescaled family of kernels $(\rho_\delta)_\delta$ for horizons $\delta > 0$ is then defined by

$$\rho_\delta(x) = c_\delta \rho\left(\frac{x}{\delta}\right), \quad x \in \mathbb{R}^n, \tag{6.22}$$

with $(c_\delta)_\delta \subset (0, \infty)$ a suitable sequence of scaling factors. Precisely, they are chosen as $c_\delta = \delta^{-n}$ for the limit $\delta \to 0$ and as $c_\delta = \overline{\rho}(1/\delta)^{-1}$ for the limit $\delta \to \infty$.

We collect here a few general observations about the rescaled kernels and associated gradients. First, it follows that $\operatorname{supp} \rho_\delta = \overline{B_\delta(0)}$, which makes D_{ρ_δ} a nonlocal gradient with horizon δ; in particular, the initial gradient D_ρ corresponds with the gradient D_{ρ_1} with horizon distance equal to 1. Moreover, the rescaling preserves the key properties of the kernel function, that is, for any $\delta > 0$, the kernel ρ_δ also satisfies (H0)-(H4). This makes all the results in the previous sections applicable to these kernels as well, in particular, the existence result of Theorem 6.2.11. Finally, we observe that the kernel associated to ρ_δ satisfies

$$Q_{\rho_\delta} = c_\delta Q_\rho\left(\frac{\cdot}{\delta}\right) \quad \text{and} \quad \widehat{Q}_{\rho_\delta} = c_\delta \delta^n \widehat{Q}_\rho(\delta \cdot), \tag{6.23}$$

where we have used [122, Proposition 2.3.22 (7)] for the interaction between scaling and Fourier transforms.

To highlight the dependence on the horizon parameter, we denote the spaces associated to D_{ρ_δ} by

$$H^{\rho,p,\delta}(\mathbb{R}^n) = \{u \in L^p(\mathbb{R}^n) \, : \, D_{\rho_\delta} u \in L^p(\mathbb{R}^n; \mathbb{R}^n)\},$$

and similarly for $H_0^{\rho,p,\delta}(\Omega)$; it holds specifically that $H^{\rho,p}(\mathbb{R}^n) = H^{\rho,p,1}(\mathbb{R}^n)$.

6.3 Localization when $\delta \to 0$

This section is devoted to the localization process, that is, to the asymptotic analysis in the limit of vanishing horizon. We start by showing that the suitably scaled nonlocal gradients converge to the classical one as $\delta \to 0$, and subsequently prove compactness results uniformly in the horizon parameter δ. Finally, we utilize these tools to establish the Γ-convergence of integral functionals depending on scaled nonlocal gradients to their local counterparts as the horizon tends to zero.

For this analysis, we fix a radial kernel ρ that satisfies (H0)-(H4) and (6.20), and consider for $\delta \in (0, 1]$, the scaled kernels

$$\rho_\delta = \frac{1}{\delta^n} \rho\left(\frac{\cdot}{\delta}\right) \quad \text{and} \quad Q_{\rho_\delta} = \frac{1}{\delta^n} Q_\rho\left(\frac{\cdot}{\delta}\right),$$

which corresponds to (6.22) with the scaling factors $c_\delta = \delta^{-n}$. Observe that this choice of scaling preserves the normalizations $\int_{\mathbb{R}^n} \rho_\delta \, dx = n$ and $\int_{\mathbb{R}^n} Q_{\rho_\delta} \, dx = 1$ for each $\delta \in (0,1]$, and (6.23) specifies to

$$\widehat{Q}_{\rho_\delta} = \widehat{Q}_\rho(\delta \,\cdot\,). \tag{6.24}$$

Throughout this section, we take $p \in (1,\infty)$ and assume Ω to be a bounded Lipschitz domain.

6.3.1 Localization of the nonlocal gradient

Here, we present the convergence of the scaled nonlocal gradients to the classical gradient. Starting with the case of smooth functions, which features an explicit convergence rate, we subsequently extend the analysis to Sobolev functions on bounded domains and the whole space $\mathbb{R}^n$. In the case of bounded domains, the nonlocal gradient is defined on a smaller set than the classical one, but this difference vanishes as $\delta \to 0$. Closely related localization results can be found in [16, 161].

Lemma 6.3.1 (Localization of nonlocal gradients). *The following statements hold:*

(i) *For each $\varphi \in C_c^\infty(\mathbb{R}^n)$ and for all $\delta \in (0,1]$, one has that*

$$\|D_{\rho_\delta}\varphi - \nabla\varphi\|_{L^\infty(\mathbb{R}^n;\mathbb{R}^n)} \leq \delta^2 \mathrm{Lip}(\nabla^2\varphi). \tag{6.25}$$

In particular, $D_{\rho_\delta}\varphi \to \nabla\varphi$ uniformly on $\mathbb{R}^n$ as $\delta \to 0$.

(ii) *For each $u \in W^{1,p}(\Omega)$, it holds that*

$$\mathbb{1}_{\Omega_{-\delta}} D_{\rho_\delta} u \to \nabla u \quad \text{in } L^p(\Omega;\mathbb{R}^n) \text{ as } \delta \to 0;$$

recall that $\Omega_{-\delta} := \{x \in \Omega \,:\, \mathrm{dist}(x,\Omega^c) > \delta\}$.

(iii) *For each $u \in W^{1,p}(\mathbb{R}^n)$, one has that $u \in H^{\rho_\delta,p}(\mathbb{R}^n)$ for all $\delta \in (0,1]$, and*

$$D_{\rho_\delta} u \to \nabla u \quad \text{in } L^p(\mathbb{R}^n;\mathbb{R}^n) \text{ as } \delta \to 0.$$

Proof. Part (i): Let $\psi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$. Then the multivariate version of Taylor's theorem with integral remainder shows for $x \in \mathbb{R}^n$ that

$$\begin{aligned}
|Q_{\rho_\delta} * \psi(x) - \psi(x)| &= \left| \int_{B_\delta(x)} Q_{\rho_\delta}(x-y)\big(\psi(y) - \psi(x)\big)\,dy \right| \\
&\leq \left| \int_{B_\delta(x)} Q_{\rho_\delta}(x-y)\nabla\psi(x)(y-x)\,dy \right| \\
&\quad + \int_{B_\delta(x)} Q_{\rho_\delta}(x-y) \left| \int_0^1 \big(\nabla\psi(x+t(y-x)) - \nabla\psi(x)\big)(y-x)\,dt \right| dy \\
&\leq \delta^2 \mathrm{Lip}(\nabla\psi) \int_{B_\delta(x)} Q_{\rho_\delta}(x-y)\,dy = \delta^2 \mathrm{Lip}(\nabla\psi),
\end{aligned}$$

where we have used that $\|Q_{\rho_\delta}\|_{L^1(\mathbb{R}^n)} = 1$, and also the radiality of Q_{ρ_δ} to cancel the term in the second line. Applying this estimate with $\psi = \nabla\varphi$ for any $\varphi \in C_c^\infty(\mathbb{R}^n)$ proves the claim in light of (6.9).

Part (ii): Since $u \in W^{1,p}(\Omega)$, the nonlocal gradient $D_{\rho_\delta} u$ is well-defined in $\Omega_{-\delta}$ and coincides with $Q_{\rho_\delta} * \nabla u$ on this set (cf. [36, Proposition 3.5]). Let $j \in \mathbb{N}$ and choose $\varphi_j \in C_c^\infty(\mathbb{R}^n)$ such that

$$\|\varphi_j - u\|_{W^{1,p}(\Omega)} \leq \frac{1}{j},$$

which is possible as Ω is a bounded Lipschitz domain by assumption. Additionally, we can choose $\delta = \delta(j)$ small enough in light of Part (i) such that

$$\|D_{\rho_\delta}\varphi_j - \nabla\varphi_j\|_{L^p(\Omega;\mathbb{R}^n)} \leq \frac{1}{j} \quad \text{and} \quad \|\nabla u\|_{L^p(\Omega\setminus\Omega_{-\delta};\mathbb{R}^n)} \leq \frac{1}{j}.$$

The previous estimates along with (6.9) then imply

$$\begin{aligned}\|\mathbb{1}_{\Omega_{-\delta}} D_{\rho_\delta} u - \nabla u\|_{L^p(\Omega;\mathbb{R}^n)} &\leq \|D_{\rho_\delta} u - D_{\rho_\delta}\varphi_j\|_{L^p(\Omega_{-\delta};\mathbb{R}^n)} + \|D_{\rho_\delta}\varphi_j - \nabla\varphi_j\|_{L^p(\Omega_{-\delta};\mathbb{R}^n)} \\ &\quad + \|\nabla\varphi_j - \nabla u\|_{L^p(\Omega_{-\delta};\mathbb{R}^n)} + \|\nabla u\|_{L^p(\Omega\setminus\Omega_{-\delta};\mathbb{R}^n)} \\ &\leq \|Q_{\rho_\delta} * \nabla u - Q_{\rho_\delta} * \nabla\varphi_j\|_{L^p(\Omega_{-\delta};\mathbb{R}^n)} + \frac{3}{j} \\ &\leq \|Q_{\rho_\delta}\|_{L^1(\mathbb{R}^n)}\|\nabla u - \nabla\varphi_j\|_{L^p(\Omega;\mathbb{R}^n)} + \frac{3}{j} \leq \frac{4}{j},\end{aligned}$$

where the last line is due to Young's convolution inequality.

Part (iii): This follows with similar arguments as in Part (ii) or, alternatively, from [16, Theorem C] with $\mathcal{A} = \nabla$. □

Remark 6.3.2. a) In view of estimate (6.25), D_{ρ_δ} converges to ∇ quadratically in δ, given that $\nabla^2\varphi$ is Lipschitz continuous. More generally, if φ is twice differentiable such that $\nabla^2\varphi$ is α-Hölder continuous with $\alpha \in (0,1]$, a similar argument induces the convergence rate $\delta^{1+\alpha}$, while for a differentiable φ with α-Hölder continuous gradient, convergence takes place at a rate of δ^α.

b) Our Γ-convergence result in Section 6.3.3 is formulated for admissible functions with prescribed Dirichlet conditions in the complement of Ω. Therefore, Lemma 6.3.1 (iii) is sufficient for these purposes. However, the sharper result in Part (ii) for bounded domains can be useful in the future for studying vanishing-horizon limits in more general settings, such as the Neumann-type problems considered in [141]. △

6.3.2 Compactness uniformly in $\delta \in (0, 1]$

In this section, we establish a compactness result for the nonlocal gradients that hold uniformly in the horizon parameter. The following theorem, which is also interesting in its own right (cf. (6.27) below), serves as a technical basis by providing a comparison between the norms of the nonlocal gradients with different horizons; this includes also the classical gradient, denoted for consistency by $D_{\rho_0} := \nabla$. Our proof relies on Fourier multiplier theory and takes inspiration from the one of [36, Theorem 7.2] for comparing Sobolev spaces associated to different nonlocal gradients.

Theorem 6.3.3 (Comparison between scaled nonlocal gradients). *Let $\bar{\delta} > 0$ and $(\delta_1, \delta_2) \in [\bar{\delta}, 1] \times [0, 1]$. Then, there exists a constant $C = C(\rho, n, p, \bar{\delta}) > 0$ such that*

$$\|D_{\rho_{\delta_1}}\varphi\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \leq C\|D_{\rho_{\delta_2}}\varphi\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \quad \textit{for all } \varphi \in C_c^\infty(\mathbb{R}^n).$$

Proof. Define the function $m_{\delta_1,\delta_2} : \mathbb{R}^n \setminus \{0\} \to (0, \infty)$ by

$$m_{\delta_1,\delta_2}(\xi) := \frac{\widehat{Q}_\rho(\delta_1\xi)}{\widehat{Q}_\rho(\delta_2\xi)},$$

recalling that $\widehat{Q}_\rho$ is non-negative. Then, we find in view of (6.10) and (6.24) that

$$\widehat{D_{\rho_{\delta_1}}\varphi} = m_{\delta_1,\delta_2}\widehat{D_{\rho_{\delta_2}}\varphi}.$$

for every $\varphi \in C_c^\infty(\mathbb{R}^n)$; in particular, the case $\delta_2 = 0$ is valid given that $\widehat{Q}_\rho(0) = 1$. It now suffices to show with the help the Mihlin-Hörmander theorem (cf. e.g. [122, Theorem 6.2.7]) that m_{δ_1,δ_2} are L^p-multipliers with estimates independent of δ_1 and δ_2. To this aim, we need to prove that for every $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq \frac{n}{2} + 1$,

$$\left|\partial^\alpha m_{\delta_1,\delta_2}(\xi)\right| \leq C|\xi|^{-|\alpha|} \quad \text{for all } \xi \neq 0 \tag{6.26}$$

with a constant $C > 0$ depending only on n, ρ and $\bar{\delta}$.

We note that the second part of Lemma 6.2.6 together with the Leibniz and quotient rules for differentiation imply

$$|\partial^\alpha m_{\delta_1,\delta_2}(\xi)| \leq C|\xi|^{-|\alpha|}|m_{\delta_1,\delta_2}(\xi)| \quad \text{for all } \xi \neq 0$$

with $C = C(n) > 0$. Therefore, it only remains to verify (6.26) for $\alpha = 0$, that is, we need to show m_{δ_1,δ_2} is uniformly bounded independent of δ_1 and δ_2. We prove this by distinguishing two cases.

Case 1: $\delta_1 \geq \delta_2$. For $0 < |\xi| \leq \frac{1}{\delta_2 \varepsilon}$ with $\varepsilon > 0$ the parameter in the hypotheses (H0)-(H4), one can estimate

$$\left|m_{\delta_1,\delta_2}(\xi)\right| = \left|\frac{\widehat{Q}_\rho(\delta_1\xi)}{\widehat{Q}_\rho(\delta_2\xi)}\right| \leq \left(\min_{\overline{B_{1/\varepsilon}(0)}} \widehat{Q}_\rho\right)^{-1},$$

considering that $\|\widehat{Q}_\rho\|_{L^\infty(\mathbb{R}^n)} \leq \|Q_\rho\|_{L^1(\mathbb{R}^n)} = 1$ by (6.21). On the other hand, we infer for $|\xi| \geq \frac{1}{\delta_2\varepsilon}$ from Lemma 6.2.6 that

$$\begin{aligned} |m_{\delta_1,\delta_2}(\xi)| &= \left|\frac{\widehat{Q}_\rho(\delta_1\xi)}{\widehat{Q}_\rho(\delta_2\xi)}\right| \leq C\left(\frac{\delta_2}{\delta_1}\right)^n \frac{\overline{\rho}\left(\frac{1}{\delta_1|\xi|}\right)}{\overline{\rho}\left(\frac{1}{\delta_2|\xi|}\right)} = C\left(\frac{\delta_2}{\delta_1}\right)^{1-\gamma} \frac{\overline{\rho}\left(\frac{1}{\delta_1|\xi|}\right)\left(\frac{1}{\delta_1|\xi|}\right)^{n+\gamma-1}}{\overline{\rho}\left(\frac{1}{\delta_2|\xi|}\right)\left(\frac{1}{\delta_2|\xi|}\right)^{n+\gamma-1}} \\ &\leq C\left(\frac{\delta_2}{\delta_1}\right)^{1-\gamma} \leq C\bar{\delta}^{\gamma-1}, \end{aligned}$$

where the second inequality uses the almost monotonicity in (H4).

Case 2: $\delta_1 \leq \delta_2$. Similarly as in Case 1, we obtain for $0 < |\xi| \leq \frac{1}{\delta_1\varepsilon}$ that

$$\left|m_{\delta_1,\delta_2}(\xi)\right| = \left|\frac{\widehat{Q}_\rho(\delta_1\xi)}{\widehat{Q}_\rho(\delta_2\xi)}\right| \leq \left(\min_{\overline{B_{\delta_2/(\delta_1\varepsilon)}(0)}} \widehat{Q}_\rho\right)^{-1} \leq \left(\min_{\overline{B_{1/(\bar{\delta}\varepsilon)}(0)}} \widehat{Q}_\rho\right)^{-1},$$

and for $|\xi| \geq 1/(\delta_1\varepsilon)$ by Lemma 6.2.6 that

$$|m_{\delta_1,\delta_2}(\xi)| = \left|\frac{\widehat{Q}_\rho(\delta_1\xi)}{\widehat{Q}_\rho(\delta_2\xi)}\right| \leq C\left(\frac{\delta_2}{\delta_1}\right)^n \frac{\overline{\rho}\left(\frac{1}{\delta_1|\xi|}\right)}{\overline{\rho}\left(\frac{1}{\delta_2|\xi|}\right)} \leq C\left(\frac{\delta_2}{\delta_1}\right)^{1-\sigma} \leq C\bar{\delta}^{\sigma-1},$$

with the second inequality due to (H3).

Finally, combining the two cases shows that $\left|m_{\delta_1,\delta_2}(\xi)\right| \leq C = C(\rho,\bar{\delta})$ for all $\xi \neq 0$, which concludes the proof. □

Together with a density argument, the previous theorem shows that $H^{\rho,p,\delta_1}(\mathbb{R}^n) = H^{\rho,p,\delta_2}(\mathbb{R}^n)$ and $H_0^{\rho,p,\delta_1}(\Omega) = H_0^{\rho,p,\delta_2}(\Omega)$ for all $\delta_1, \delta_2 \in (0,1]$ or equivalently, that

$$H^{\rho,p,\delta}(\mathbb{R}^n) = H^{\rho,p}(\mathbb{R}^n) \quad \text{and} \quad H_0^{\rho,p,\delta}(\Omega) = H_0^{\rho,p}(\Omega) \qquad \text{for all } \delta \in (0,1]. \tag{6.27}$$

In fact, by inspecting the proof of Theorem 6.3.3, it is not hard to see that (6.27) holds for all $\delta > 0$, which shows that our nonlocal function spaces do not depend on the horizon parameter δ. Based on this observation, we obtain the following corollary as a consequence of (6.18).

Corollary 6.3.4. *There exists a constant $C = C(\rho, n, \Omega, p) > 0$ such that*

$$\|u\|_{H^{\sigma,p}(\mathbb{R}^n)} \leq C\|D_{\rho_\delta}u\|_{L^p(\Omega_\delta;\mathbb{R}^n)} \quad \text{for all } u \in H_0^{\rho,p,\delta}(\Omega) \text{ and } \delta \in (0,1].$$

Remark 6.3.5. We observe that there is another way of proving Corollary 6.3.4 that does not pass through the stronger statement of Theorem 6.3.3. For this alternative argument, it suffices to require that the kernel ρ satisfies (H0)-(H2) and

$$\liminf_{|x|\to 0} \rho(x)|x|^{n+\sigma-1} > 0.$$

Indeed, one can compare ρ_δ with the kernel from Example 6.2.5 a) for $s = \sigma$ by arguing as in [36, Theorem 7.2] and checking that the constants are independent of δ. When $p = 2$, even (H2) is not necessary (cf. [36, Theorem 7.2]), so that also the truncated fractional gradients in Remark 6.2.8 are covered. △

In order to ensure the existence of convergent subsequences required for the forthcoming Γ-convergence result (see Theorem 6.3.7), we proceed with the following compactness statement.

Lemma 6.3.6 (Convergent subsequences for vanishing horizon). *Let $(\delta_j)_j \subset (0,1]$ be a sequence with $\delta_j \to 0$ and suppose that $u_j \in H_0^{\rho,p,\delta_j}(\Omega)$ for each $j \in \mathbb{N}$ with*

$$\sup_{j\in\mathbb{N}}\|D_{\rho_{\delta_j}}u_j\|_{L^p(\Omega_{\delta_j};\mathbb{R}^n)} < \infty.$$

Then, there is a $u \in W_0^{1,p}(\Omega)$ (extended to $\mathbb{R}^n$ as zero) such that, up to a non-relabeled subsequence,

$$u_j \to u \ \text{ in } L^p(\mathbb{R}^n) \quad \text{and} \quad D_{\rho_{\delta_j}}u_j \rightharpoonup \nabla u \ \text{ in } L^p(\mathbb{R}^n;\mathbb{R}^n) \text{ as } j \to \infty.$$

Proof. By Corollary 6.3.4, the sequence $(u_j)_j$ is bounded in $H^{\sigma,p}(\mathbb{R}^n)$. Since each u_j is supported in $\overline{\Omega}$, we conclude from the compact embedding $H^{\sigma,p}(\mathbb{R}^n) \hookrightarrow\hookrightarrow L^p_{\text{loc}}(\mathbb{R}^n)$ (cf. 6.14) that there is a $u \in L^p(\mathbb{R}^n)$ with $u = 0$ a.e. in Ω^c such that, up to a non-relabeled subsequence,

$$u_j \to u \qquad \text{in } L^p(\mathbb{R}^n) \text{ as } j \to \infty.$$

Moreover, up to extracting a potential further subsequence, we find that $D_{\rho_{\delta_j}}u_j \rightharpoonup V$ in $L^p(\mathbb{R}^n;\mathbb{R}^n)$ for some $V \in L^p(\mathbb{R}^n;\mathbb{R}^n)$. To deduce that $V = \nabla u$, we compute for $\varphi \in C_c^\infty(\mathbb{R}^n)$ that

$$\begin{aligned}\int_{\mathbb{R}^n} V\,\varphi\,dx &= \lim_{j\to\infty}\int_{\mathbb{R}^n} D_{\rho_{\delta_j}}u_j\,\varphi\,dx \\ &= -\lim_{j\to\infty}\int_{\mathbb{R}^n} u_j\,D_{\rho_{\delta_j}}\varphi\,dx \\ &= -\int_{\mathbb{R}^n} u\,\nabla\varphi\,dx,\end{aligned}$$

where the last equality follows from the localization result for the nonlocal gradients in Lemma 6.3.1 (i). This shows that $u \in W_0^{1,p}(\Omega)$ (extended to $\mathbb{R}^n$ as zero) with $\nabla u = V$, which finishes the proof. □

6.3.3 Γ-convergence $\delta \to 0$

We are now in the position to make the conjectured localization of our variational problems in the limit $\delta \to 0$ rigorous, choosing Γ-convergence as a natural framework.

Before stating the theorem, let us collect the relevant objects. The family of vectorial energy functionals $(\mathcal{F}_\delta)_{\delta\in(0,1]}$ with $\mathcal{F}_\delta : L^p(\mathbb{R}^n;\mathbb{R}^m) \to \mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$ for $\delta \in (0,1]$ is given by

$$\mathcal{F}_\delta(u) := \begin{cases} \displaystyle\int_{\Omega_\delta} f(x, D_{\rho_\delta}u)\,dx & \text{for } u \in H_0^{\rho,p,\delta}(\Omega;\mathbb{R}^m), \\ \infty & \text{else,} \end{cases} \tag{6.28}$$

where $f : \Omega_1 \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a suitable Carathéodory integrand; considering that the functions in the domain of $\mathcal{F}_\delta$ are defined on $\mathbb{R}^n$ with zero Dirichlet conditions in Ω^c, we may take, without loss of generality, the integrals over the bounded set Ω_δ.

As prospective limit functional, we introduce $\mathcal{F}_0 : L^p(\mathbb{R}^n;\mathbb{R}^m) \to \mathbb{R}_\infty$ given by

$$\mathcal{F}_0(u) := \begin{cases} \displaystyle\int_{\Omega} f(x, \nabla u)\,dx & \text{for } u \in W_0^{1,p}(\Omega;\mathbb{R}^m), \\ \infty & \text{else;} \end{cases} \tag{6.29}$$

here, functions in $W_0^{1,p}(\Omega;\mathbb{R}^m)$ are identified with their extension to $\mathbb{R}^n$ as zero.

Theorem 6.3.7 (Localization for vanishing horizon via Γ-convergence). *Let $f : \Omega_1\times\mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory integrand such that*

$$c|A|^p - C \leq f(x,A) \leq C(1 + |A|^p) \quad \textit{for a.e. } x \in \Omega_1 \textit{ and all } A \in \mathbb{R}^{m\times n}$$

with $c, C > 0$. If $f(x,\cdot)$ is quasiconvex for a.e. $x \in \Omega$, then the family $(\mathcal{F}_\delta)_{\delta\in(0,1]}$ in (6.28) *Γ-converges with respect to $L^p(\mathbb{R}^n;\mathbb{R}^m)$-convergence to the functional $\mathcal{F}_0$ in* (6.29) *as $\delta \to 0$, that is,*

$$\Gamma(L^p)\text{-}\lim_{\delta\to 0} \mathcal{F}_\delta = \mathcal{F}_0.$$

Additionally, $(\mathcal{F}_\delta)_\delta$ is equi-coercive with respect to convergence in $L^p(\mathbb{R}^n;\mathbb{R}^m)$.

Proof. Let $(\delta_j)_j \subset (0,1]$ be a sequence converging to 0 as $j \to \infty$.

Equi-coercivity: By the growth bound on f from below and Corollary 6.3.4, we deduce that there are constants $c', C' > 0$ such that

$$\mathcal{F}_{\delta_j}(u) \geq c'\|u\|_{H^{\sigma,p}(\mathbb{R}^n)} - C'$$

for all $j \in \mathbb{N}$ and $u \in H_0^{\rho,p,\delta_j}(\Omega;\mathbb{R}^m)$. This yields the equi-coercivity, given the compact embedding of $H^{\sigma,p}(\mathbb{R}^n)$ into $L^p_{\mathrm{loc}}(\mathbb{R}^n)$, cf. (6.14).

Liminf-inequality: Let $(u_j)_j \subset L^p(\mathbb{R}^n;\mathbb{R}^m)$ with $u_j \to u$ in $L^p(\mathbb{R}^n;\mathbb{R}^m)$. Assuming without loss of generality that

$$\liminf_{j\to\infty} \mathcal{F}_{\delta_j}(u_j) = \lim_{j\to\infty} \mathcal{F}_{\delta_j}(u_j) < \infty,$$

we have that $u_j \in H_0^{\rho,p,\delta_j}(\Omega;\mathbb{R}^m)$ for each $j \in \mathbb{N}$ and

$$\sup_{j\in\mathbb{N}} \|D_{\rho_{\delta_j}} u_j\|_{L^p(\Omega_{\delta_j};\mathbb{R}^m)} < \infty,$$

due to the lower bound on f. Lemma 6.3.6 therefore yields a $u \in W_0^{1,p}(\Omega;\mathbb{R}^m)$ (extended to $\mathbb{R}^n$ as zero) such that $D_{\rho_{\delta_j}} u_j \rightharpoonup \nabla u$ in $L^p(\mathbb{R}^n;\mathbb{R}^{m\times n})$. If we use translation operators as in Lemma 6.2.9 to define

$$v_j := \mathcal{Q}_{\rho_{\delta_j}} u_j = Q_{\rho_{\delta_j}} * u_j \quad \text{for } j \in \mathbb{N},$$

then $(v_j)_j$ is a bounded sequence in $W^{1,p}(\mathbb{R}^n;\mathbb{R}^m)$ with $\nabla v_j = D_{\rho_{\delta_j}} u_j \rightharpoonup \nabla u$ in $L^p(\mathbb{R}^n;\mathbb{R}^{m\times n})$ as $j \to \infty$. Consequently, it even holds that $v_j \rightharpoonup u$ in $W^{1,p}(\mathbb{R}^n;\mathbb{R}^m)$, considering that the functions v_j are zero outside of Ω_1 for each $j \in \mathbb{N}$. A standard lower semicontinuity result for functionals with quasiconvex integrands (cf. [75, Theorem 8.11]) then implies

$$\begin{aligned}
\liminf_{j\to\infty} \mathcal{F}_{\delta_j}(u_j) &= \liminf_{j\to\infty} \int_{\Omega_{\delta_j}} f(x, D_{\rho_{\delta_j}} u_j)\,dx \\
&\geq \liminf_{j\to\infty} \int_{\Omega} f(x, D_{\rho_{\delta_j}} u_j)\,dx - C|\Omega_{\delta_j} \setminus \Omega| \\
&= \liminf_{j\to\infty} \int_{\Omega} f(x, \nabla v_j)\,dx \geq \int_{\Omega} f(x, \nabla u)\,dx = \mathcal{F}_0(u),
\end{aligned}$$

which is the desired liminf-inequality.

Recovery sequence: Without loss of generality, consider $u \in W_0^{1,p}(\Omega;\mathbb{R}^m)$. Then, we infer from Lemma 6.3.1 (*iii*) that $u \in H_0^{\rho,p,\delta_j}(\Omega;\mathbb{R}^m)$ for all $j \in \mathbb{N}$ with $D_{\rho_{\delta_j}} u \to \nabla u$ in $L^p(\mathbb{R}^n;\mathbb{R}^{m\times n})$ as $j \to \infty$. The upper bound on f enables the application of Lebesgue's dominated convergence theorem to find

$$\lim_{j\to\infty} \mathcal{F}_{\delta_j}(u) = \lim_{j\to\infty} \int_{\Omega_{\delta_j}} f(x, D_{\rho_{\delta_j}} u)\,dx = \int_{\Omega} f(x, \nabla u)\,dx = \mathcal{F}_0(u).$$

This shows that the constant sequence is a suitable recovery sequence. □

Remark 6.3.8. a) Under the assumptions of Theorem 6.3.7, for every $\delta \in (0,1]$, the functional $\mathcal{F}_\delta$ admits a minimizer according to Theorem 6.2.11. Therefore, standard properties of Γ-convergence imply that these minimizers converge, up to subsequence, to a minimizer of $\mathcal{F}_0$ as $\delta \to 0$.

Referring to the literature, a closely related Γ-convergence result involving similar nonlocal gradients in the case $m = 1$ can be found in [161, Theorem 1.7]. However, the latter does not feature the equi-coercivity required for the convergence of minimizers.

b) Since the definition of $D_{\rho_\delta} u$ on Ω_δ only depends on the values of u in $\Omega_{2\delta}$, the Dirichlet condition in Ω^c can be equivalently replaced by prescribing zero values in $\Omega_{2\delta} \setminus \Omega$. We remark that the papers [30, Theorem 6.1] and [72, Corollary 2] on finite-horizon fractional gradients use a slightly different convention by considering the gradients on Ω and requiring Dirichlet conditions in $\Omega_\delta \setminus \Omega_{-\delta}$. Clearly, both settings are equivalent by a suitable renaming of the domain. The reason for our choice is that only the setting used here is meaningful for both limit passages $\delta \to 0$ and $\delta \to \infty$.

c) Theorem 6.3.7 can readily be extended to non-zero complementary values, that is, to admissible functions in the spaces $g + H_0^{\rho,p,\delta}(\Omega;\mathbb{R}^m)$ for any given $g \in W^{1,p}(\mathbb{R}^n;\mathbb{R}^m)$. Indeed, since $g \in H^{\rho,p,\delta}(\Omega;\mathbb{R}^m)$ for all $\delta > 0$ with

$$\|D_{\rho_\delta} g\|_{L^p(\mathbb{R}^n;\mathbb{R}^{m\times n})} \leq \|Q_{\rho_\delta}\|_{L^1(\mathbb{R}^n)} \|\nabla g\|_{L^p(\mathbb{R}^n;\mathbb{R}^{m\times n})} = \|\nabla g\|_{L^p(\mathbb{R}^n;\mathbb{R}^{m\times n})},$$

the argument follows through in the same manner with the domain of the Γ-limit being $g + W_0^{1,p}(\Omega;\mathbb{R}^m)$. △

Example 6.3.9. By applying Theorem 6.3.7 to the kernels of Example 6.2.5, we obtain the localization of functionals as in (6.28) with nonlocal gradients associated to the following scaled kernels:

a) For ρ as in Example 6.2.5 a), one finds that

$$\rho_\delta(x) = \delta^{-n}\frac{w(x/\delta)}{|x/\delta|^{n+s-1}} = \delta^{s-1}\frac{w(x/\delta)}{|x|^{n+s-1}}, \quad x \in \mathbb{R}^n \setminus \{0\}.$$

These scaled kernels coincide (up to a constant) with finite-horizon fractional gradients D^s_δ from [30,31,72,141]. Theorem 6.3.7 then complements the localization result for $s \uparrow 1$ in [72, Theorem 7].

b) The scaled versions of the kernels ρ in Example 6.2.5 b) read as

$$\rho_\delta(x) = \delta^{-n}\frac{w(x/\delta)\log^\kappa(\delta/|x|)}{|x/\delta|^{n+s-1}} = \delta^{s-1}\frac{w(x/\delta)(\log(\delta)-\log(|x|))^\kappa}{|x|^{n+s-1}}, \quad x \in \mathbb{R}^n \setminus \{0\}.$$

c) With ρ as in Example 6.2.5 c), the scaled kernels are given by

$$\rho_\delta(x) = \delta^{-n}\frac{w(x/\delta)}{|x/\delta|^{n+s(|x/\delta|)-1}} = \delta^{s(|x/\delta|)-1}\frac{w(x/\delta)}{|x|^{n+s(|x/\delta|)-1}}, \quad x \in \mathbb{R}^n \setminus \{0\}.$$

6.4 Γ-convergence $\delta \to \infty$

We focus now on the asymptotics of the nonlocal gradient as the horizon δ diverges to infinity. As proven below, the associated limiting object is the Riesz fractional gradient. While this is to be expected for finite-horizon fractional gradients, surprisingly, the same holds when starting from any general nonlocal gradient within our setting. This section is structured in parallel to Section 6.3, showing first the convergence of the nonlocal gradients to the Riesz fractional gradient, then providing a uniform compactness result, and finally, proving the Γ-convergence of the associated energy functionals.

Let ρ again be a non-negative radial kernel that satisfies (H0)-(H4) and (6.20), and assume throughout that $p \in (1,\infty)$ and Ω is a bounded Lipschitz domain. The lack of integrability of the fractional kernel on $\mathbb{R}^n$ calls for a different scaling procedure compared with Section 6.3. Precisely, for $\delta \in (1/\varepsilon, \infty)$ (so that $\overline{\rho}(1/\delta) \neq 0$), we now define

$$\rho_\delta(x) := \overline{\rho}\left(\frac{1}{\delta}\right)^{-1}\rho\left(\frac{x}{\delta}\right), \tag{6.30}$$

which corresponds to (6.22) with $c_\delta := \overline{\rho}(1/\delta)^{-1}$. In this way, the values of ρ_δ on the unit sphere $\partial B_1(0)$ are normalized to 1 for any δ. In addition, we require that these kernels converge pointwise on $\mathbb{R}^n \setminus \{0\}$ as $\delta \to \infty$, and set

$$\rho_\infty(x) := \lim_{\delta\to\infty}\rho_\delta(x) = \lim_{\delta\to\infty}\overline{\rho}\left(\frac{1}{\delta}\right)^{-1}\rho\left(\frac{x}{\delta}\right), \quad x \in \mathbb{R}^n \setminus \{0\}. \tag{6.31}$$

With the scaling (6.30), the kernel function Q_{ρ_δ} and its Fourier transform satisfy

$$Q_{\rho_\delta} = \overline{\rho}\left(\frac{1}{\delta}\right)^{-1} Q_\rho\left(\frac{\cdot}{\delta}\right) \quad \text{and} \quad \widehat{Q}_{\rho_\delta} = \delta^n\overline{\rho}\left(\frac{1}{\delta}\right)^{-1}\widehat{Q}_\rho(\delta\,\cdot). \tag{6.32}$$

Let us point out that our chosen scaling is, up to a constant, the only relevant one. Indeed, if there is a sequence $(\bar{c}_\delta)_\delta$ of positive reals such that $(\bar{c}_\delta \rho(\cdot/\delta))_\delta$ converges pointwise as $\delta \to \infty$, then for $x \in \partial B_1(0)$,

$$\lim_{\delta\to\infty} \bar{c}_\delta \rho\left(\frac{x}{\delta}\right) = \lim_{\delta\to\infty} \bar{c}_\delta \overline{\rho}\left(\frac{1}{\delta}\right) =: \bar{c}.$$

Therefore, we obtain for all $x \in \mathbb{R}^n \setminus \{0\}$ that

$$\lim_{\delta\to\infty} \bar{c}_\delta \rho\left(\frac{x}{\delta}\right) = \lim_{\delta\to\infty} \bar{c}_\delta \overline{\rho}\left(\frac{1}{\delta}\right) \overline{\rho}\left(\frac{1}{\delta}\right)^{-1} \rho\left(\frac{x}{\delta}\right) = \bar{c}\rho_\infty(x).$$

6.4.1 Convergence of nonlocal gradients as $\delta \to \infty$

This section is about establishing that the scaled nonlocal gradients converge to the Riesz fractional gradient as $\delta \to \infty$. We commence with some bounds on ρ_δ from (6.30) and the limit kernel ρ_∞ that will be used repeatedly later. Recall that $\varepsilon > 0$ is as in (H0)-(H4), and that σ, γ with $0 < \sigma \leq \gamma < 1$ are the parameters appearing in the hypotheses (H3) and (H4), respectively.

Lemma 6.4.1. *There exist constants $C, c > 0$ such that for every $\delta > 1/\varepsilon$ and all $x \in B_{\delta\varepsilon}(0) \setminus \{0\}$,*

$$c \min\left\{\frac{1}{|x|^{n+\sigma-1}}, \frac{1}{|x|^{n+\gamma-1}}\right\} \leq \rho_\delta(x) \leq C \max\left\{\frac{1}{|x|^{n+\sigma-1}}, \frac{1}{|x|^{n+\gamma-1}}\right\}. \tag{6.33}$$

In particular, it holds for all $x \in \mathbb{R}^n \setminus \{0\}$ that

$$c \min\left\{\frac{1}{|x|^{n+\sigma-1}}, \frac{1}{|x|^{n+\gamma-1}}\right\} \leq \rho_\infty(x) \leq C \max\left\{\frac{1}{|x|^{n+\sigma-1}}, \frac{1}{|x|^{n+\gamma-1}}\right\}.$$

Proof. Observe first that by (H3) and (H4), there are constants $c', C' > 0$ such that

$$c'\left(\frac{t}{r}\right)^{n+\gamma-1} \overline{\rho}(t) \leq \overline{\rho}(r) \leq C'\left(\frac{t}{r}\right)^{n+\sigma-1} \overline{\rho}(t) \tag{6.34}$$

for all $t, r \in (0, \varepsilon)$ with $r \geq t$.

Let $x \in B_{\delta\varepsilon}(0)$. If $|x| \geq 1$, we can apply (6.34) with the choice $r = |x|/\delta$ and $t = 1/\delta$ to find

$$\frac{c'}{|x|^{n+\gamma-1}} \leq \overline{\rho}\left(\frac{1}{\delta}\right)^{-1} \overline{\rho}\left(\frac{|x|}{\delta}\right) \leq \frac{C'}{|x|^{n+\sigma-1}}.$$

As for the case $0 < |x| \leq 1$, we resort to (6.34) as well, but take $r = 1/\delta$ and $t = |x|/\delta$ instead, which gives

$$\frac{1}{C'|x|^{n+\sigma-1}} \leq \overline{\rho}\left(\frac{1}{\delta}\right)^{-1} \overline{\rho}\left(\frac{|x|}{\delta}\right) \leq \frac{1}{c'|x|^{n+\gamma-1}}.$$

We conclude that (6.33) holds for suitably chosen constants c, C. □

As the next lemma shows, ρ_∞ must be a fractional kernel, no matter the specific choice of ρ. This finding is a key ingredient for proving that only Riesz fractional gradients can be obtained as the limit of increasing horizon nonlocal gradients with the scaled sequence of kernels ρ_δ.

Lemma 6.4.2 (Limit kernel ρ_∞ is fractional). *There is an $s_\infty \in [\sigma, \gamma]$ such that ρ_∞ in (6.31) satisfies*

$$\rho_\infty(x) = \rho^{s_\infty}(x) := \frac{1}{|x|^{n+s_\infty-1}} \quad \text{for all } x \in \mathbb{R}^n \setminus \{0\}.$$

Proof. Our argument relies on proving that $\overline{\rho}_\infty$ is a multiplicative function. To this aim, we consider $r, t > 0$ and compute that

$$\begin{aligned}\overline{\rho}_\infty(r\cdot t) &= \lim_{\delta\to\infty} \overline{\rho}\left(\frac{1}{\delta}\right)^{-1} \overline{\rho}\left(\frac{r\cdot t}{\delta}\right) = \lim_{\delta\to\infty} \overline{\rho}\left(\frac{1}{\delta}\right)^{-1} \overline{\rho}\left(\frac{r}{\delta}\right)\overline{\rho}\left(\frac{r}{\delta}\right)^{-1} \overline{\rho}\left(\frac{r\cdot t}{\delta}\right)\\ &= \overline{\rho}_\infty(r) \lim_{\delta\to\infty} \overline{\rho}\left(\frac{1}{\delta/r}\right)^{-1} \overline{\rho}\left(\frac{t}{\delta/r}\right) = \overline{\rho}_\infty(r)\overline{\rho}_\infty(t).\end{aligned}$$

Since $\overline{\rho}_\infty$ is also locally bounded away from 0 by Lemma 6.4.1, we deduce that $\overline{\rho}_\infty$ must be a power function (cf. [4, Chapter 3, Proposition 6]). Together with Lemma 6.4.1, it follows therefore that $\rho_\infty(x) = 1/|x|^{n+s_\infty-1}$ for some $s_\infty \in [\sigma, \gamma]$. □

Remark 6.4.3. The parameter s_∞ associated to a limit kernel ρ_∞ can be determined directly from ρ via the limit

$$s_\infty = \log(\overline{\rho}_\infty(1/e)) - n + 1 = \lim_{\delta\to\infty} \log\left(\overline{\rho}(1/\delta)^{-1}\overline{\rho}\left(1/(e\delta)\right)\right) - n + 1.$$

△

Example 6.4.4. One observes that the kernels ρ from Example 6.2.5 satisfy (6.31), i.e., their rescaled versions converge pointwise. We identify the limiting fractional exponent s_∞ for illustration:

a) Let ρ be as in Example 6.2.5 a). Then, for $x \in \mathbb{R}^n \setminus \{0\}$,

$$\rho_\infty(x) = \lim_{\delta\to\infty} \frac{\overline{w}(1/\delta)^{-1}}{\delta^{n+s-1}}\frac{w(x/\delta)}{|x/\delta|^{n+s-1}} = \frac{1}{|x|^{n+s-1}} = \rho^s(x),$$

which yields $s_\infty = s$.

b) For the kernel ρ of Example 6.2.5 b), one obtains the same limit as in a), that is, $\rho_\infty = \rho^s$, and hence, $s_\infty = s$. The detailed calculation reads

$$\begin{aligned}\rho_\infty(x) &= \lim_{\delta\to\infty} \overline{w}(1/\delta)^{-1}\frac{1}{\log^\kappa(\delta)\delta^{n+s-1}}\frac{w(x/\delta)\log^\kappa(\delta/|x|)}{|x/\delta|^{n+s-1}} = \frac{1}{|x|^{n+s-1}}\lim_{\delta\to\infty}\frac{\log^\kappa(\delta/|x|)}{\log^\kappa(\delta)}\\ &= \frac{1}{|x|^{n+s-1}}\lim_{\delta\to\infty}\left(\frac{\log(\delta)-\log(|x|)}{\log(\delta)}\right)^\kappa = \frac{1}{|x|^{n+s-1}}\end{aligned}$$

for $x \in \mathbb{R}^n \setminus \{0\}$.

c) In the case of ρ from Example 6.2.5 c), the limit fractional exponent becomes $s_\infty = s(0)$, as

$$\begin{aligned}\rho_\infty(x) &= \lim_{\delta\to\infty} \overline{w}(1/\delta)^{-1}\frac{1}{\delta^{n+s(1/\delta)-1}}\frac{w(x/\delta)}{|x/\delta|^{n+s(|x|/\delta)-1}}\\ &= \lim_{\delta\to\infty}\frac{1}{\delta^{s(1/\delta)-s(|x|/\delta)}|x|^{n+s(|x|/\delta)-1}} = \frac{1}{|x|^{n+s(0)-1}}\end{aligned}$$

for $x \in \mathbb{R}^n \setminus \{0\}$ shows.

In the next step, we show that the nonlocal gradients converge to the fractional gradient induced by ρ_∞ as $\delta \to \infty$, see Proposition 6.4.6. The proof involves the following auxiliary result, which allows to control the integrability of the kernels during the limit passage.

Lemma 6.4.5. *It holds that*

$$\rho_\delta \min\{1, |\cdot|^{-1}\} \to \rho_\infty \min\{1, |\cdot|^{-1}\} \quad \textit{in } L^1(\mathbb{R}^n) \textit{ as } \delta \to \infty.$$

Proof. We already know the pointwise a.e. convergence by (6.31), and it follows from (6.33) that

$$\mathbb{1}_{B_{\delta\varepsilon}(0)}\rho_\delta \min\{1, |\cdot|^{-1}\} \le C \min\{1, |\cdot|^{-1}\} \max\left\{\frac{1}{|\cdot|^{n+\sigma-1}}, \frac{1}{|\cdot|^{n+\gamma-1}}\right\}.$$

Since the right-hand side is integrable, Lebesgue's dominated convergence theorem implies

$$\mathbb{1}_{B_{\delta\varepsilon}(0)}\rho_\delta \, \min\{1, |\cdot|^{-1}\} \to \rho_\infty \, \min\{1, |\cdot|^{-1}\} \quad \text{in } L^1(\mathbb{R}^n) \text{ as } \delta \to \infty.$$

It remains to show that

$$\mathbb{1}_{B_{\delta\varepsilon}(0)^c}\rho_\delta \min\{1, |\cdot|^{-1}\} \to 0 \quad \text{in } L^1(\mathbb{R}^n) \text{ as } \delta \to \infty. \tag{6.35}$$

Considering that, in light of (H1), $\rho(\cdot/\delta) \le C$ in $B_{\delta\varepsilon}(0)^c$ for some $C > 0$, and $\rho(\cdot/\delta) = 0$ on $B_\delta(0)^c$ by (6.20), we find

$$\mathbb{1}_{B_{\delta\varepsilon}(0)^c}\rho_\delta \le C\mathbb{1}_{B_\delta(0)\setminus B_{\delta\varepsilon}(0)}\overline{\rho}(1/\delta)^{-1} \le C\mathbb{1}_{B_\delta(0)\setminus B_1(0)}\overline{\rho}(1/\delta)^{-1},$$

given that $\delta\varepsilon > 1$. This yields

$$\begin{aligned}\int_{B_{\delta\varepsilon/2}(0)^c} \rho_\delta(x) \min\{1, |x|^{-1}\}\, dx &\le C\overline{\rho}\left(\frac{1}{\delta}\right)^{-1} \int_{B_\delta(0)\setminus B_1(0)} |x|^{-1}\, dx \\ &\le \begin{cases} C\overline{\rho}(1/\delta)^{-1}(\delta^{n-1} - 1) & \text{if } n > 1, \\ C\overline{\rho}(1/\delta)^{-1}\log(\delta) & \text{if } n = 1. \end{cases}\end{aligned} \tag{6.36}$$

In either case, the expression in (6.36) converges to 0 as $\delta \to \infty$ in view of (H3), which gives rise to (6.35) and finishes the proof. □

As a consequence, we now obtain the convergence of the nonlocal gradients to the Riesz fractional gradient as $\delta \to \infty$ in the case of Sobolev functions.

Proposition 6.4.6 (Convergence to fractional gradient as $\delta \to \infty$). *For any $u \in W^{1,p}(\mathbb{R}^n)$ it holds that*

$$D_{\rho_\delta}u \to D_{\rho_\infty}u = D^{s_\infty}u \quad \text{in } L^p(\mathbb{R}^n;\mathbb{R}^n) \text{ as } \delta \to \infty.$$

Proof. In light of [92, Proposition 1] (cf. also the proof of [36, Proposition 3.5]), we deduce the estimate

$$\|D_{\rho_\delta}u - D_{\rho_\infty}u\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \le C\|u\|_{W^{1,p}(\mathbb{R}^n)}\left\|(\rho_\delta - \rho_\infty)\min\{1, |\cdot|^{-1}\}\right\|_{L^1(\mathbb{R}^n)},$$

for all $u \in W^{1,p}(\mathbb{R}^n)$, and the statement follows via Lemma 6.4.5. □

6.4.2 Compactness uniformly in $\delta \in (1/\varepsilon, \infty)$

Next, we address the issue of compactness with the goal of deriving a counterpart of Lemma 6.3.6 in the setting of diverging horizon. This relies on the following analogue of the Poincaré-type inequality in Corollary 6.3.4. The proof is based on the comparison of the scaled nonlocal gradients D_{ρ_δ} with a suitable finite-horizon fractional gradient.

Proposition 6.4.7. *There exists a constant $C = C(\rho, n, \Omega, p) > 0$ such that*

$$\|u\|_{H^{\sigma,p}(\mathbb{R}^n)} \le C\|D_{\rho_\delta}u\|_{L^p(\Omega_\delta;\mathbb{R}^n)} \quad \text{for all } u \in H_0^{\rho,p,\delta}(\Omega) \text{ and } \delta \in (1/\varepsilon, \infty).$$

Proof. Consider $D_1^\sigma := D_{\rho_1^\sigma}$ induced by the kernel function

$$\rho_1^\sigma = \frac{w}{|\cdot|^{n+\sigma-1}}, \tag{6.37}$$

where $w \in C_c^\infty(\mathbb{R}^n)$ is a non-negative radially decreasing function with $w(0) > 0$ and $\operatorname{supp} w = B_1(0)$; note that ρ_1^σ falls into the setting of Example 6.2.5 a) with $s = \sigma$ and recall that σ is the parameter appearing in the hypothesis (H3) on ρ.

Then, by (6.18) there is a constant $C > 0$ such that

$$\|u\|_{H^{\sigma,p}(\mathbb{R}^n)} \leq C\|D_1^\sigma u\|_{L^p(\Omega_1;\mathbb{R}^n)} \quad \text{for all } u \in H_0^{\rho_1^\sigma,p}(\Omega).$$

The remaining proof shows that there is a constant $C > 0$ independent of δ such that

$$\|D_1^\sigma \varphi\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \leq C\|D_{\rho_\delta}\varphi\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^n) \text{ and } \delta \in (1/\varepsilon, \infty), \tag{6.38}$$

from which the claim follows after a density argument.

Let us define $m_\delta : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ as

$$m_\delta(\xi) := \frac{\widehat{Q}_{\rho_1^\sigma}(\xi)}{\widehat{Q}_{\rho_\delta}(\xi)},$$

and observe that, in light of (6.10),

$$\widehat{D_1^\sigma \varphi} = m_\delta \widehat{D_{\rho_\delta}\varphi} \qquad \text{for } \varphi \in C_c^\infty(\mathbb{R}^n).$$

The estimate (6.38) follows then directly from Fourier multiplier theory, once m_δ is confirmed to satisfy the Mihlin-Hörmander condition with uniform constants, that is,

$$|\partial^\alpha m_\delta(\xi)| \leq C|\xi|^{-|\alpha|} \tag{6.39}$$

for all $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq n/2 + 1$ and $C > 0$ a constant independent of δ.

To this aim, observe that (6.16) implies for $\xi \neq 0$ that

$$\begin{aligned}\left|\partial^\alpha \widehat{Q}_{\rho_\delta}(\xi)\right| &= \left|\overline{\rho}(1/\delta)^{-1}\delta^{n+|\alpha|}\partial^\alpha \widehat{Q}_\rho(\delta\xi)\right| \\ &\leq C\overline{\rho}(1/\delta)^{-1}\delta^{n+|\alpha|}|\delta\xi|^{-|\alpha|}\left|\widehat{Q}_\rho(\delta\xi)\right| = C|\xi|^{-|\alpha|}\left|\widehat{Q}_{\rho_\delta}(\xi)\right|,\end{aligned}$$

with $C > 0$ independent of δ. Since the same holds for $\widehat{Q}_{\rho_1^\sigma}$, we deduce via the Leibniz and quotient rule that

$$|\partial^\alpha m_\delta(\xi)| \leq C|\xi|^{-|\alpha|}|m_\delta(\xi)| \quad \text{for all } \xi \neq 0.$$

Therefore, it remains to verify (6.39) for $\alpha = 0$, which corresponds to showing that m_δ is bounded independent of δ. For simpler notation, we write $\langle\xi\rangle := \sqrt{1+|\xi|^2}$ for $\xi \in \mathbb{R}^n$. Since the estimate (6.15) in Lemma 6.2.6 along with (6.37) allows us to deduce

$$\widehat{Q}_{\rho_1^\sigma}(\xi) \leq C\langle\xi\rangle^{\sigma-1} \quad \text{for all } \xi \in \mathbb{R}^n,$$

the proof of (6.39) for $\alpha = 0$ can be reduced to verifying that

$$\widehat{Q}_{\rho_\delta}(\xi) \geq C\langle\xi\rangle^{\sigma-1} \quad \text{for all } \xi \in \mathbb{R}^n. \tag{6.40}$$

Let us first consider $|\xi| \leq 1/(\delta\varepsilon)$. Then, in view of (6.32),

$$\widehat{Q}_{\rho_\delta}(\xi) \geq \overline{\rho}(1/\delta)^{-1}\delta^n \min_{\overline{B_{1/\varepsilon}(0)}} \widehat{Q}_\rho \geq C \geq C\langle\xi\rangle^{\sigma-1}, \tag{6.41}$$

where C is independent of δ because $\overline{\rho}(1/\delta)^{-1}\delta^n \to \infty$ as $\delta \to \infty$ by (H4). For the case $|\xi| > 1/(\delta\varepsilon)$, we use Lemma 6.2.6 and Lemma 6.4.1 to infer

$$\begin{aligned} \widehat{Q}_{\rho_\delta}(\xi) &\geq C\overline{\rho}(1/\delta)^{-1}\delta^n|\delta\xi|^{-n}\overline{\rho}(1/|\delta\xi|) \\ &\geq C|\xi|^{-n}\min\big\{|\xi|^{n+\sigma-1}, |\xi|^{n+\gamma-1}\big\} \\ &= C\min\big\{|\xi|^{\sigma-1}, |\xi|^{\gamma-1}\big\} \geq C\langle\xi\rangle^{\sigma-1}. \end{aligned} \tag{6.42}$$

Finally, (6.41) together with (6.42) gives (6.40), and thus, (6.39). This finishes the proof in light of the Mihlin-Hörmander theorem (see e.g. [122, Theorem 6.2.7]). □

Remark 6.4.8. While the previous proof is built on (6.38), we mention that a statement parallel to Theorem 6.3.3 cannot be expected to hold for an unbounded parameter range of δ. Indeed, this is due to the fact that the singular behavior of ρ_∞ and ρ_δ at the origin may be different, as one can see, for instance, from the two kernels in Example 6.4.4 b); they feature a stronger and weaker singularity than ρ_∞, respectively. △

By combining Proposition 6.4.7 and Proposition 6.4.6, we can now deduce the following compactness statement.

Lemma 6.4.9 (Convergent subsequences for diverging horizon). *Let $(\delta_j)_j \subset (1/\varepsilon, \infty)$ be a sequence with $\delta_j \to \infty$ and suppose that $u_j \in H_0^{\rho,p,\delta_j}(\Omega)$ for each $j \in \mathbb{N}$ with*

$$\sup_{j\in\mathbb{N}} \|D_{\rho_{\delta_j}} u_j\|_{L^p(\Omega_{\delta_j};\mathbb{R}^n)} < \infty.$$

Then, there is a $u \in H_0^{s_\infty,p}(\Omega)$, such that, up to a non-relabeled subsequence,

$$u_j \to u \quad \text{in } L^p(\mathbb{R}^n) \quad \text{and} \quad D_{\rho_{\delta_j}} u_j \rightharpoonup D^{s_\infty} u \quad \text{in } L^p(\mathbb{R}^n;\mathbb{R}^n) \text{ as } j \to \infty.$$

Moreover, for every $\eta > 0$ it holds that

$$D_{\rho_{\delta_j}} u_j \to D^{s_\infty} u \quad \text{in } L^p((\Omega_\eta)^c;\mathbb{R}^n) \text{ as } j \to \infty.$$

Proof. We infer from Proposition 6.4.7 that $(u_j)_j$ is a bounded sequence in $H^{\sigma,p}(\mathbb{R}^n)$. Then, the compact embedding $H^{\sigma,p}(\mathbb{R}^n) \hookrightarrow\hookrightarrow L^p_{\mathrm{loc}}(\mathbb{R}^n)$ (see (6.14)) together with the fact that each u_j is supported in $\overline{\Omega}$ yields the existence of a $u \in L^p(\mathbb{R}^n)$ with $u = 0$ a.e. in Ω^c such that, up to a non-relabeled subsequence,

$$u_j \to u \qquad \text{in } L^p(\mathbb{R}^n) \text{ as } j \to \infty. \tag{6.43}$$

After selecting a potential further subsequence, we find that $D_{\rho_{\delta_j}} u_j \rightharpoonup V$ in $L^p(\mathbb{R}^n;\mathbb{R}^n)$ for some $V \in L^p(\mathbb{R}^n;\mathbb{R}^n)$. One can compute that

$$\begin{aligned} \int_{\mathbb{R}^n} V\,\varphi\,dx &= \lim_{j\to\infty}\int_{\mathbb{R}^n} D_{\rho_{\delta_j}} u_j\,\varphi\,dx \\ &= -\lim_{j\to\infty}\int_{\mathbb{R}^n} u_j\,\operatorname{div}_{\rho_{\delta_j}}\varphi\,dx = -\int_{\mathbb{R}^n} u\,\operatorname{div}^{s_\infty}\varphi\,dx, \end{aligned}$$

for any $\varphi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$, where the last equality employs Proposition 6.4.6 adapted to the nonlocal divergence. This allows us conclude that $V = D^{s_\infty}u$ and $u \in H_0^{s_\infty,p}(\Omega)$.

To prove the second part of the statement, we exploit that the nonlocal gradients on $(\Omega_\eta)^c$ can be expressed as a convolution. Precisely, let us define

$$d_\delta(z) := -\frac{z\rho_\delta(z)}{|z|^2} \quad \text{and} \quad d_\infty(z) := -\frac{z\rho_\infty(z)}{|z|^2} \quad \text{for } z \in \mathbb{R}^n \setminus \{0\}.$$

For $\varphi \in C_c^\infty(\Omega)$, we can compute in view of the radiality of ρ that for any $x \in (\Omega_\eta)^c$,

$$D_{\rho_\delta}\varphi(x) = \int_{\mathbb{R}^n} -\frac{\varphi(y)}{|x-y|}\frac{x-y}{|x-y|}\rho_\delta(x-y)\,dy = (\mathbb{1}_{B_\eta(0)^c}d_\delta) * \varphi(x); \tag{6.44}$$

since $\mathbb{1}_{B_\eta(0)^c}d_\delta \in L^1(\mathbb{R}^n)$, the identity (6.44) can be extended via density to all $u \in H_0^{\rho,p,\delta}(\Omega)$. In the same way, there is an analogous representation when considering the kernels ρ_∞, that is,

$$D^{s_\infty}u = D_{\rho_\infty}u = (\mathbb{1}_{B_\eta(0)^c}d_\infty) * u \qquad \text{on } (\Omega_\eta)^c$$

for $u \in H_0^{s_\infty,p}(\Omega) = H_0^{\rho_\infty,p}(\Omega)$.

Furthermore, we observe that Lemma 6.4.5 induces the convergence

$$\mathbb{1}_{B_\eta(0)^c}d_\delta \to \mathbb{1}_{B_\eta(0)^c}d_\infty \qquad \text{in } L^1(\mathbb{R}^n) \text{ as } \delta \to \infty.$$

This allows us to conclude by Young's convolution inequality and (6.43) that

$$\|D_{\rho_{\delta_j}}u_j - D^{s_\infty}u\|_{L^p((\Omega_\eta)^c;\mathbb{R}^n)} = \|(\mathbb{1}_{B_\eta(0)^c}d_\delta) * u_j - (\mathbb{1}_{B_\eta(0)^c}d_\infty) * u\|_{L^p((\Omega_\eta)^c;\mathbb{R}^n)} \to 0 \text{ as } \delta \to 0.$$

□

6.4.3 Γ-convergence $\delta \to \infty$

Based on the technical foundations provided in the previous sections, we are now in the position to prove the Γ-convergence for diverging horizon. We consider for $\delta \in (1/\varepsilon, \infty)$ the functionals $\mathcal{F}_\delta : L^p(\mathbb{R}^n;\mathbb{R}^m) \to \mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$ given by

$$\mathcal{F}_\delta(u) := \begin{cases} \displaystyle\int_{\Omega_\delta} f(x, D_{\rho_\delta}u)\,dx & \text{for } u \in H_0^{\rho,p,\delta}(\Omega;\mathbb{R}^m), \\ \infty & \text{else,} \end{cases} \tag{6.45}$$

where $f : \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a suitable Carathéodory integrand and ρ_δ is the scaled version of the kernel ρ, cf (6.30). As made precise in Theorem 6.4.10 below, the limiting object for $\delta \to \infty$ is the functional $\mathcal{F}_\infty : L^p(\mathbb{R}^n;\mathbb{R}^m) \to \mathbb{R}_\infty$,

$$\mathcal{F}_\infty(u) := \mathcal{F}^{s_\infty}(u) := \begin{cases} \displaystyle\int_{\mathbb{R}^n} f(x, D^{s_\infty}u)\,dx & \text{for } u \in H_0^{s_\infty,p}(\Omega;\mathbb{R}^m), \\ \infty & \text{else;} \end{cases} \tag{6.46}$$

The fractional parameter s_∞ is here related to the kernel ρ via $\lim_{\delta\to\infty}\rho_\delta = \rho^{s_\infty}$, see (6.31), Lemma 6.4.2, and also Remark 6.4.3.

Theorem 6.4.10 (Γ-convergence for diverging horizon). *Let $f : \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory integrand such that*

$$c|A|^p - a(x) \le f(x, A) \le a(x) + C|A|^p \quad \textit{for a.e. } x \in \mathbb{R}^n \textit{ and all } A \in \mathbb{R}^{m\times n}$$

with $c, C > 0$ and $a \in L^1(\mathbb{R}^n)$. If $f(x,\cdot)$ is quasiconvex for a.e. $x \in \Omega$, then the family $(\mathcal{F}_\delta)_{\delta\in(1/\varepsilon,\infty)}$ in (6.45) *Γ-converges with respect to $L^p(\mathbb{R}^n;\mathbb{R}^m)$-convergence to the functional $\mathcal{F}_\infty$ in* (6.46) *as $\delta \to \infty$, that is,*

$$\Gamma(L^p)\text{-}\lim_{\delta\to\infty} \mathcal{F}_\delta = \mathcal{F}_\infty.$$

Additionally, the sequence $(\mathcal{F}_\delta)_\delta$ is equi-coercive with respect to convergence in $L^p(\mathbb{R}^n;\mathbb{R}^m)$.

Proof. Let $(\delta_j)_j \subset (1/\varepsilon, \infty)$ be a sequence converging to ∞ as $j \to \infty$.

Equi-coercivity: From Proposition 6.4.7 and the lower bound on f, we deduce that

$$\mathcal{F}_{\delta_j}(u) \geq C\|u\|_{H^{\sigma,p}(\mathbb{R}^n)} - \|a\|_{L^1(\mathbb{R}^n)}$$

for all $j \in \mathbb{N}$ and $u \in H_0^{\rho,p,\delta_j}(\Omega;\mathbb{R}^m)$. The embedding (6.14) now immediately gives the stated equi-coercivity.

Liminf-inequality: Consider a sequence $(u_j)_j \subset L^p(\mathbb{R}^n;\mathbb{R}^m)$ with $u_j \to u$ in $L^p(\mathbb{R}^n;\mathbb{R}^m)$ satisfying, without loss of generality,

$$\liminf_{j\to\infty} \mathcal{F}_{\delta_j}(u_j) = \lim_{j\to\infty} \mathcal{F}_{\delta_j}(u_j) < \infty.$$

Then, $u_j \in H_0^{\rho,p,\delta_j}(\Omega;\mathbb{R}^m)$ for each $j \in \mathbb{N}$ and $\sup_{j\in\mathbb{N}}\|D_{\rho_{\delta_j}} u_j\|_{L^p(\Omega_{\delta_j};\mathbb{R}^m)} < \infty$, so that Lemma 6.4.9 yields that u lies in $H_0^{s_\infty,p}(\Omega;\mathbb{R}^m)$ with

$$D_{\rho_{\delta_j}} u_j \rightharpoonup D^{s_\infty} u \qquad \text{in } L^p(\mathbb{R}^n;\mathbb{R}^{m\times n}) \text{ as } j \to \infty. \tag{6.47}$$

Similarly to the proof of Theorem 6.3.7, we perform a translation to the classical gradient setting in order to estimate the integral contribution over Ω. With $v_j := \mathcal{Q}_{\rho_{\delta_j}} u_j \in W^{1,p}(\mathbb{R}^n;\mathbb{R}^m)$ for $j \in \mathbb{N}$, we have that $\nabla v_j = D_{\rho_{\delta_j}} u_j$ due to Lemma 6.2.9. Moreover, there exists a $v \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^m)$ with $\nabla v = D^{s_\infty} u$ by [140, Proposition 3.1 (i)]. We therefore obtain in view of (6.47) that $\nabla v_j \rightharpoonup \nabla v$ in $L^p(\Omega;\mathbb{R}^{m\times n})$, and (up to translation by constants) that $v_j \rightharpoonup v$ in $W^{1,p}(\Omega;\mathbb{R}^m)$. A standard lower semicontinuity result for quasiconvex integrands (cf. [75, Theorem 8.11]) then yields

$$\begin{aligned}\liminf_{j\to\infty} \int_\Omega f(x, D_{\rho_{\delta_j}} u_j)\,dx &= \liminf_{j\to\infty} \int_\Omega f(x, \nabla v_j)\,dx \\ &\geq \int_\Omega f(x, \nabla v)\,dx = \int_\Omega f(x, D^{s_\infty} u)\,dx.\end{aligned} \tag{6.48}$$

Regarding the integral contributions over Ω^c, observe that for any $\eta > 0$,

$$D_{\rho_{\delta_j}} u_j \to D^{s_\infty} u \qquad \text{in } L^p((\Omega_\eta)^c;\mathbb{R}^n) \text{ as } j \to \infty.$$

Together with the upper and lower bound on f and Lebesgue's dominated convergence theorem, we obtain

$$\begin{aligned}\liminf_{j\to\infty} \int_{\Omega_{\delta_j}\setminus\Omega} f(x, D_{\rho_{\delta_j}} u_j)\,dx &\geq \liminf_{j\to\infty} \int_{\Omega_{\delta_j}\setminus\Omega_\eta} f(x, D_{\rho_{\delta_j}} u_j)\,dx - \int_{\Omega_\eta\setminus\Omega} a(x)\,dx \\ &= \int_{\mathbb{R}^n\setminus\Omega_\eta} f(x, D^{s_\infty} u)\,dx - \int_{\Omega_\eta\setminus\Omega} a(x)\,dx \\ &\geq \int_{\mathbb{R}^n\setminus\Omega} f(x, D^{s_\infty} u)\,dx - \int_{\Omega_\eta\setminus\Omega} 2a(x) + C|D^{s_\infty} u|^p\,dx.\end{aligned}$$

Letting $\eta \to 0$ under consideration of $|\partial\Omega| = 0$ results in

$$\liminf_{j\to\infty} \int_{\Omega_{\delta_j}\setminus\Omega} f(x, D_{\rho_{\delta_j}} u_j)\, dx \geq \int_{\mathbb{R}^n\setminus\Omega} f(x, D^{s_\infty} u)\, dx. \tag{6.49}$$

The desired liminf-inequality follows from adding (6.48) and (6.49).

Recovery sequence: It suffices to consider $u \in H_0^{s_\infty,p}(\Omega;\mathbb{R}^m)$. Let $(u_k)_k \subset C_c^\infty(\Omega;\mathbb{R}^m)$ be a sequence such that $u_k \to u$ in $H_0^{s_\infty,p}(\Omega;\mathbb{R}^m)$ as $k \to \infty$. From Proposition 6.4.6 and the second part of Lemma 6.4.9, we deduce that

$$D_{\rho_{\delta_j}} u_k \to D^{s_\infty} u_k \quad \text{in } L^p(\mathbb{R}^n;\mathbb{R}^{m\times n}) \text{ as } j \to \infty \text{ for all } k \in \mathbb{N}.$$

Hence, the upper bound on f and a twofold application of Lebesgue's dominated convergence theorem shows

$$\lim_{k\to\infty} \lim_{j\to\infty} \int_{\Omega_{\delta_j}} f(x, D_{\rho_{\delta_j}} u_k)\, dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} f(x, D^{s_\infty} u_k)\, dx = \int_{\mathbb{R}^n} f(x, D^{s_\infty} u)\, dx.$$

Finally, a recovery sequence is obtained by extracting a suitable diagonal sequence in the sense of Attouch [18, Lemma 1.15, Proposition 1.16]. □

Remark 6.4.11. a) We find as a consequence of the Γ-convergence and equi-coercivity proven in Theorem 6.4.10 that the minimizers of the functionals $\mathcal{F}_\delta$ in (6.45), whose existence is guaranteed by Theorem 6.2.11, converge (up to a subsequence) in L^p to a minimizer of $\mathcal{F}^{s_\infty}$. In particular, this result applies to all kernels from Example 6.2.5 with their limiting fractional exponents s_∞ computed in Example 6.4.4.

b) Note that Theorem 6.4.10 can be readily generalized to functionals defined on the spaces $g + H_0^{\rho,p,\delta}(\Omega;\mathbb{R}^m)$ with a given complementary value $g \in W^{1,p}(\mathbb{R}^n;\mathbb{R}^m)$, considering that Proposition 6.4.6 holds for all Sobolev functions in $W^{1,p}(\mathbb{R}^n)$. △

Chapter 7

Structural changes in nonlocal denoising models arising through bi-level parameter learning

This chapter corresponds to the published article

[82] E. Davoli, R. Ferreira, C. Kreisbeck and H. Schönberger. Structural changes in nonlocal denoising models arising through bi-level parameter learning. *Applied Mathematics and Optimization*, 88(1):Paper No. 9, 2023. https://doi.org/10.1007/s00245-023-09982-4.

7.1 Introduction

One of the most widely used methods to solve image restoration problems is the variational regularization approach. This variational approach consists of minimizing a reconstruction functional that decomposes into a fidelity and a regularization term, which give rise to competing effects. While the fidelity term ensures that the reconstructed image is close to the (noisy) data, the regularization term is designed to remove the noise by incorporating prior information on the clean image. In the case of a simple L^2-fidelity term, the reconstruction functional is given by

$$\mathcal{J}(u) = \|u - u^\eta\|^2_{L^2(\Omega)} + \mathcal{R}(u), \qquad \text{for } u \in L^2(\Omega),$$

where $\Omega \subset \mathbb{R}^n$ is the image domain, $u^\eta \in L^2(\Omega)$ the noisy image, and $\mathcal{R} : L^2(\Omega) \to [0, \infty]$ the regularizer.

A common choice for $\mathcal{R}$ is the total variation (TV) regularization proposed by Rudin, Osher, & Fatemi [186], which penalizes sharp oscillations, but does not exclude edge discontinuities, as they appear in most images. Since its introduction, the TV-model has inspired a variety of more advanced regularization terms, like the infimal-convolution total variation ($ICTV$) [60], the total generalized variation (TGV) [51], and many more, cf. [42] and the references therein. Due to the versatility of the variational formulation, regularizers of a completely different type can be used as well. Recently, a lot of attention has been directed towards regularizers incorporating nonlocal effects, such as those induced by difference quotients [21, 46, 53, 117] and fractional operators [12, 14, 15]. Nonlocal regularizers have the advantage of not requiring the existence of (full) derivatives, allowing to work with functions that are less regular than those in the local counterpart.

With an abundance of available choices, finding a suitable regularization term for a specific application is paramount for obtaining accurate reconstructions. This is often done by fixing a parameter-dependent family of regularizers and tuning the parameter in accordance with the noise and data. Carrying out this process via trial and error can be hard and inefficient, which led to the development of a more structured approach in the form of bi-level optimization. We refer, e.g., to [87, 88] (see also [62, 63, 101, 202]) and to the references therein, as well as to [95] for a detailed overview. The idea behind bi-level optimization is to employ a supervised learning scheme based on a set of training data consisting of noisy images and their corresponding clean versions. To determine an optimal parameter, one minimizes a selected cost functional which quantifies the error with respect to the training data. Overall, this results in a nested variational problem with upper- and lower-level optimization steps related to the cost and reconstruction functional, respectively. Key aspects of the mathematical study of these bi-level learning schemes include establishing the existence of solutions and deriving optimality conditions, which lay the foundation for devising reliable numerical solution methods.

In recent years, there has been a rapid growth in the literature devoted to addressing the above questions. To mention but a few examples, we first refer the paper [135] dealing with learning real-valued weight parameters in front of the regularization terms for a rather general class of inverse problems; in [13, 25], the authors optimize the fractional parameter of a regularizer depending on the spectral fractional Laplacian; spatially dependent weights are determined through training via other nonlocal bi-level schemes (e.g., inside the Gagliardo semi-norm [134] or in a type of fractional gradient [90]), and in classical TV-models [64, 133, 172]; as done in [86], one can also learn the fidelity term instead of the regularizer.

A common denominator in the above references is the presence of certain a priori compactness constraints on the set of admissible parameters, such as box constraints like in [135], where the weights are assumed to lie in some compact interval away from 0 and infinity. These conditions make it possible to prove stability of the lower-level problem and obtain existence of optimal parameters within a class of structurally equivalent regularizers. However, imposing artificial restrictions to the parameter range like these may lead to suboptimal results depending on the given training data.

It is then substantial to consider removing such constraints in order to work on maximal domains naturally associated with the parameters, which is also our focus in this paper. An inherent effect of this approach is that qualitative changes in the structure of the regularizer may occur at the edges of the domain. If optimal parameters are attained at the boundary, this indicates that the chosen class of regularization terms is not well-suited to the training data. To exclude these degenerate cases, it is of interest to provide analytic conditions to guarantee that the optimal parameters are attained in the interior of the domain, thereby preserving the structure of the regularizer. The first work to address the aforementioned tasks is [87] by De Los Reyes, Schönlieb, & Valkonen, where optimization is carried out for weighted sums of local regularizers of different type with each weight factor allowed to take any value in $[0, \infty]$. As such, their bi-level scheme is able to encompass multiple regularization structures at once, like TV and TV^2 and their interpolation TGV. Similarly, the authors in [152] vary the weight factor in the whole range $[0, \infty]$ as well as the underlying finite-dimensional norm of the total variation regularizer. We also mention [84], where the order of a newly introduced nonlocal counterpart of the TGV-regularizer is tuned, and [83], which studies a bi-level scheme covering the cases of TV, TGV^2, and $NsTGV^2$ in a comprehensive way.

In this paper, we introduce a unified framework to deal with parameter learning beyond structural stability in the context of bi-level optimization schemes. In contrast to the above references, where the analysis is tailored to a specifically chosen type of parameter dependence, our regularizers can exhibit a general dependence on parameters in a topological space. Precisely, we consider a parametrized family of regularizers $\mathcal{R}_\lambda : L^2(\Omega) \to [0, \infty]$ with λ ranging over a subset Λ of a topological space X, which is assumed to be first countable. If we focus for brevity on a single data point $(u^c, u^\eta) \in L^2(\Omega) \times L^2(\Omega)$, with u^c and u^η the clean and noisy images (see Section 7.2 for larger data sets), the bi-level optimization problem reads:

$$\text{(Upper-level)} \qquad \text{Minimize } \mathcal{I}(\lambda) := \inf_{w \in K_\lambda} \|w - u^c\|^2_{L^2(\Omega)} \text{ over } \lambda \in \Lambda,$$
$$\text{(Lower-level)} \qquad K_\lambda := \operatorname*{arg\,min}_{u \in L^2(\Omega)} \mathcal{J}_\lambda(u),$$

where $\mathcal{J}_\lambda(u) := \|u - u^\eta\|^2_{L^2(\Omega)} + \mathcal{R}_\lambda(u)$ is the reconstruction functional.

Our approach for studying this general bi-level learning scheme relies on asymptotic tools from the calculus of variations. We define a suitable notion of stability for the lower-level problems that requires the family of functionals $\{\mathcal{J}_\lambda\}_{\lambda \in \Lambda}$ to be closed under taking Γ-limits; see [49, 80] for a comprehensive introduction on Γ-convergence. Since Γ-convergence ensures the convergence of sequences of minimizers, one can conclude that, in the presence of stability, the upper-level functional $\mathcal{I}$ admits a minimizer (Theorem 7.2.3).

A different strategy is required to obtain the existence of solutions when stability fails. Especially relevant here is the case of real-valued parameters when box constraints are disposed of and non-closed intervals Λ are considered; clearly, stability is then lost for the simple fact that a sequence of parameters can converge to the boundary of Λ. To overcome this issue, we propose a natural extension $\overline{\mathcal{I}} : \overline{\Lambda} \to [0, \infty]$ of $\mathcal{I}$, now defined on the closure of our parameter domain, and identified via Γ-convergence of the lower-level functionals. Precisely,

$$\text{(Upper-level)} \qquad \text{Minimize } \overline{\mathcal{I}}(\lambda) := \inf_{w \in \overline{K}_\lambda} \|w - u^c\|^2_{L^2(\Omega)} \text{ over } \lambda \in \overline{\Lambda},$$
$$\text{(Lower-level)} \qquad \overline{K}_\lambda := \operatorname*{arg\,min}_{u \in L^2(\Omega)} \overline{\mathcal{J}}_\lambda(u),$$

where the functionals $\overline{\mathcal{J}}_\lambda : L^2(\Omega) \to [0, \infty]$ are characterized as L^2-weak Γ-limits (if they exist) of functionals $\mathcal{J}_{\lambda'}$ with $\lambda' \to \lambda$. To justify the choice of this particular extension, we derive an intrinsic connection with relaxation theory in the calculus of variations (for an introduction, see, e.g., [75, Chapter 9] and the references therein). Explicitly, the relaxation of the upper-level functional $\mathcal{I}$ is given by its lower semicontinuous envelope (after the trivial extension to $\overline{\Lambda}$ by ∞),

$$\mathcal{I}^{\mathrm{rlx}}(\lambda) := \inf \left\{ \liminf_{k \to \infty} \mathcal{I}(\lambda_k) : (\lambda_k)_k \subset \Lambda, \lambda_k \to \lambda \text{ in } \overline{\Lambda} \right\} \quad \text{for } \lambda \in \overline{\Lambda}.$$

This relaxed version of $\mathcal{I}$ has the desirable property that it admits a minimizer (if $\overline{\Lambda}$ is compact) and minimizing sequences of $\mathcal{I}$ have subsequences that converge to an optimal parameter of $\mathcal{I}^{\mathrm{rlx}}$. Our main theoretical result (Theorem 7.2.5) shows that the extension $\overline{\mathcal{I}}$ coincides with the relaxation $\mathcal{I}^{\mathrm{rlx}}$ under suitable assumptions and therefore inherits the same properties (cf. Corollary 7.2.8).

Besides the generic conditions that each $\mathcal{R}_\lambda$ is weakly lower semicontinuous and has non-empty domain (see (H)), which ensure that $\mathcal{J}_\lambda$ possesses a minimizer, we work under two main assumptions:

- (i) The Mosco-convergence of the regularizers, i.e., Γ-convergence with respect to the strong and weak L^2-topology, and
- (ii) the uniqueness of minimizers of $\overline{\mathcal{J}}_\lambda$ for $\lambda \in \overline{\Lambda} \setminus \Lambda$.

We demonstrate in Example 7.2.7 that these assumptions are in fact optimal. Due to (i), the Γ-limits $\overline{\mathcal{J}}_\lambda$ preserve the additive decomposition into the L^2-fidelity term and a regularizer, and coincide with $\mathcal{J}_\lambda$ inside Λ. As a consequence of the latter, it follows that $\overline{\mathcal{I}} = \mathcal{I}$ in Λ, making $\overline{\mathcal{I}}$ a true extension of $\mathcal{I}$. For the parameter values at the boundary, $\lambda \in \overline{\Lambda} \setminus \Lambda$, however, the regularizers present in $\overline{\mathcal{J}}_\lambda$ can have a completely different structure from the family of regularizers $\{\mathcal{R}_\lambda\}_{\lambda \in \Lambda}$ that we initially started with. When the optimal parameter of the extended problem is attained inside Λ, one recovers instead a solution to the original training scheme, yielding structure preservation. For a discussion on related results in the context of optimal control problems [40, 55, 56], we refer to the end of Section 7.2.

To demonstrate the applicability of our abstract framework, we investigate a quartet of practically relevant scenarios with families of nonlocal regularizers that induce qualitatively different structural changes; namely, learning the optimal weight, varying the amount of nonlocality, optimizing the integrability exponent, and tuning the fractional parameter. More precisely, in all these four applications, our starting point is a non-closed real interval $\Lambda \subset [-\infty, \infty]$ and we seek to determine the extension $\overline{\mathcal{I}}$ on the closed interval $\overline{\Lambda}$, which admits a minimizer by the theory outlined above. The first step is to calculate the Mosco-limits of the regularizers, which reveals the type of structural change occurring at the boundary points. Subsequently, we study for which training sets of clean and noisy images the optimal parameters are attained either inside Λ or at the edges. In two cases, we determine explicit analytic conditions on the data that guarantee structure preservation for the optimization process.

The first setting involves a rather general nonlocal regularizer $\mathcal{R} : L^2(\Omega) \to [0, \infty]$ multiplied by a weight parameter α in $\Lambda = (0, \infty)$. Inside the domain, we observe structural stability as $\overline{\mathcal{J}}_\alpha = \mathcal{J}_\alpha$ for all $\alpha \in \Lambda$; in contrast, the regularization disappears when $\alpha = 0$ and forces the solutions to be constant when $\alpha = \infty$. Moreover, we derive sufficient conditions in terms of the data that prevent the optimal parameter from being attained at the boundary points; for a single data point (u^c, u^η), they specify to

$$\mathcal{R}(u^c) < \mathcal{R}(u^\eta) \quad \text{and} \quad \|u^\eta - u^c\|^2_{L^2(\Omega)} < \left\| \fint_\Omega u^\eta \, dx - u^c \right\|^2_{L^2(\Omega)},$$

see Theorem 7.3.2. Notice that the first of these two conditions is comparable to the one in [87, Eq. (10)] and shows positivity of optimal weights.

Inspired by the use of different L^p-norms in image processing, such as in the form of quadratic, TV, and Lipschitz regularization [176, Section 4], we focus our second case on the integrability exponent of nonlocal regularizers of double-integral type; precisely, functionals of the form

$$\mathcal{R}_p(u) = \left(\frac{1}{|\Omega \times \Omega|} \int_\Omega \int_\Omega f^p(x, y, u(x), u(y)) \, dx \, dy \right)^{1/p} \quad \text{for } p \in \Lambda = [1, \infty),$$

with a suitable $f : \Omega \times \Omega \times \mathbb{R} \times \mathbb{R} \to [0, \infty)$. Possible choices for the integrand f include bounded functions or functions of difference-quotient type. We prove stability of the lower-level problem in Λ, and determine the Mosco-limit for $p \to \infty$ via L^p-approximation techniques as in [61, 139]. In particular, we show that it is given by a double-supremal functional of the form

$$\mathcal{R}_\infty(u) = \operatorname{ess\,sup}_{(x,y) \in \Omega \times \Omega} f(x, y, u(x), u(y)).$$

In order to see how this structural change affects the image reconstruction, we highlight examples of training data for which the supremal regularizer performs better or worse than the integral counterparts.

As a third application, we consider two families of nonlocal regularizers $\{\mathcal{R}_\delta\}_{\delta \in \Lambda}$ with $\Lambda = (0, \infty)$, which were introduced by Aubert & Kornprobst [21] and Brezis & Nguyen in [53], respectively, and are closely related to nonlocal filters frequently used in image processing. The parameter δ reflects the amount of nonlocality in the regularizer. It is known that the functionals $\mathcal{R}_\delta$ tend, as $\delta \to 0$, to a multiple of the total variation in the sense of Γ-convergence. Based on these results, we prove in both cases that the reconstruction functional of our bi-level optimization scheme turns into the classical TV-denoising model when $\delta = 0$, whereas the regularization vanishes at the other boundary value, $\delta = \infty$. As such, the extended bi-level schemes encode simultaneously nonlocal and total variation regularizations. We round off the discussion by presenting some instances of training data where the optimal parameters are attained either at the boundary or in the interior of Λ.

Our final bi-level optimization problem features a different type of nonlocality arising from fractional operators; to be precise, we consider, in the same spirit as in [12], the L^2-norm of the spectral fractional Laplacian as a regularizer. The parameter of interest here is the order $s/2$ of the fractional Laplacian, which is taken in the fractional range $s \in \Lambda = (0, 1)$. At the values $s = 0$ and $s = 1$, we recover local models with regularizers equal to the L^2-norm of the function and its gradient, respectively. Thus, one expects the fractional model to perform better than the two local extremes. We quantify this presumption by deriving analytic conditions in terms of the eigenfunctions and eigenvalues of the classical Laplacian on Ω ensuring the optimal parameters to be attained in the truly fractional regime. These conditions on the training data are established by proving and exploiting the differentiability of the extended upper-level functional $\overline{\mathcal{I}}$.

For completeness, we mention that practically relevant scenarios when Λ is a topological space include those in which the reconstruction parameters are space-dependent, and thus described by functions. The analysis of this class of applications is left open for future investigations.

The outline of the paper is as follows. In Section 7.2, we present the general abstract bi-level framework, and prove the results regarding the existence of optimal parameters and the two types of extensions of bi-level optimization schemes. Sections 3-6 then deal with the four different, practically relevant applications mentioned in the previous paragraph. As a note, we point out that they are each presented in a self-contained way, allowing the readers to move directly to the sections that correspond best to their interests.

7.2 Establishing the unified framework

Let $\Omega \subset \mathbb{R}^n$ be an open bounded set, and let

$$\bigcup_{j=1}^{N} (u_j^c, u_j^\eta) \subset L^2(\Omega) \times L^2(\Omega), \qquad N \in \mathbb{N},$$

be a set of available square-integrable training data, where each u_j^c represents a clean image and u_j^η a distorted version thereof, which can be obtained, for instance, by applying some noise to u_j^c. These data are collected in the vector-valued functions $u^c := (u_1^c, \dots, u_N^c) \in L^2(\Omega; \mathbb{R}^N)$ and $u^\eta := (u_1^\eta, \dots, u_N^\eta) \in L^2(\Omega; \mathbb{R}^N)$. As for notation, $\|v\|^2_{L^2(\Omega;\mathbb{R}^N)} = \sum_{j=1}^N \|v_j\|^2_{L^2(\Omega)}$ stands for the L^2-norm of a function $v \in L^2(\Omega; \mathbb{R}^N)$.

To reconstruct each damaged image, u_j^η, we consider denoising models that consist of a simple fidelity term and a (possibly nonlocal) regularizer; precisely, we minimize functionals $\mathcal{J}_{\lambda,j} : L^2(\Omega) \to [0, \infty]$ of the form

$$\mathcal{J}_{\lambda,j}(u) = \|u - u_j^\eta\|^2_{L^2(\Omega)} + \mathcal{R}_\lambda(u), \qquad u \in L^2(\Omega), \tag{7.1}$$

where the regularizer $\mathcal{R}_\lambda : L^2(\Omega) \to [0, \infty]$, with $\operatorname{Dom} \mathcal{R}_\lambda = \{v \in L^2(\Omega) : \mathcal{R}_\lambda(u) < \infty\}$, is a (possibly nonlocal) functional parametrized over $\lambda \in \Lambda$ with Λ a subset of a topological space X satisfying the first axiom of countability. Throughout the paper, we always assume that for every $\lambda \in \Lambda$, we have

$$\begin{cases} \operatorname{Dom} \mathcal{R}_\lambda \text{ is non-empty,} \\ \mathcal{R}_\lambda \text{ is weakly } L^2\text{-lower semicontinuous.} \end{cases} \tag{H}$$

Observe that the functionals $\mathcal{J}_{\lambda,j}$ then have a minimizer by the direct method in the calculus of variations.

The result of the reconstruction process, meaning the quality of the reconstructed image resulting as a minimizer of (7.1), is known to depend on the choice of the regularizing term $\mathcal{R}_\lambda$. Our goal is to set up a training scheme that is able to learn how to select a "good" parameter λ within a corresponding given family $\{\mathcal{R}_\lambda\}_{\lambda\in\Lambda}$ of regularizers. Here, as briefly described in the Introduction for the single data point case ($N = 1$), we follow the approach introduced in [87, 88] in the spirit of machine learning optimization schemes, where training the regularization term means to solve the nested variational problem

$$\begin{aligned} &\text{(Upper-level)} \qquad \text{Minimize } \mathcal{I}(\lambda) := \inf_{w\in K_\lambda} \|w - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} \text{ over } \lambda \in \Lambda, \\ &\text{(Lower-level)} \qquad K_\lambda := \Big\{ w \in L^2(\Omega;\mathbb{R}^N) : w_j \in \operatorname*{arg\,min}_{u\in L^2(\Omega)} \mathcal{J}_{\lambda,j}(u) \text{ for all } j \in \{1,\dots,N\} \Big\}, \end{aligned} \tag{$\mathcal{T}$}$$

with $\mathcal{J}_{\lambda,j}$ as in (7.1). Notice that $K_\lambda \neq \emptyset$ because for all $j \in \{1,\dots,N\}$, we have

$$K_{\lambda,j} := \operatorname{argmin}_{u\in L^2(\Omega)} \mathcal{J}_{\lambda,j}(u) \neq \emptyset \tag{7.2}$$

by Assumption (H).

To study the training scheme ($\mathcal{T}$), we start by introducing a notion of weak L^2-stability for the family $\{\mathcal{J}_\lambda\}_{\lambda\in\Lambda}$, with

$$\mathcal{J}_\lambda := (\mathcal{J}_{\lambda,1},\dots,\mathcal{J}_{\lambda,N}) : L^2(\Omega) \to [0,\infty]^N \quad \text{for } \lambda \in \Lambda. \tag{7.3}$$

This notion relies on the concept of Γ-convergence and is related to the notion of (weak) stability as in [135, Definition 2.3], which is defined in terms of minimizers of the lower-level problem.

Definition 7.2.1 (Weak L^2-stability). *The family in* (7.3) *is called weakly L^2-stable if for every sequence $(\lambda_k)_k \subset \Lambda$ such that $(\mathcal{J}_{\lambda_k,j})_k$ Γ-converges with respect to the weak L^2-topology for all $j \in \{1,\dots,N\}$, there exists $\lambda \in \Lambda$ such that*

$$\Gamma(w\text{-}L^2)\text{-}\lim_{k\to\infty} \mathcal{J}_{\lambda_k,j} = \mathcal{J}_{\lambda,j}$$

for all $j \in \{1,\dots,N\}$.

Before proceeding, we briefly recall the definition and some properties of Γ-convergence in the setting relevant to us; for more on this topic, see [49, 80] for instance.

Definition 7.2.2 (Γ- and Mosco-convergence). *Let $\mathcal{F}_k : L^2(\Omega) \to [0,\infty]$ for $k \in \mathbb{N}$ and $\mathcal{F} : L^2(\Omega) \to [0,\infty]$ be functionals. The sequence $(\mathcal{F}_k)_k$ (sequentially) Γ-converges to $\mathcal{F}$ with respect to the weak L^2-topology, written $\mathcal{F} = \Gamma(w\text{-}L^2)\text{-}\lim_{k\to\infty} \mathcal{F}_k$, if:*

- *(Liminf inequality) For every sequence $(u_k)_k \subset L^2(\Omega)$ and $u \in L^2(\Omega)$ with $u_k \rightharpoonup u$ in $L^2(\Omega)$, it holds that*
$$\mathcal{F}(u) \le \liminf_{k\to\infty} \mathcal{F}_k(u_k).$$
- *(Limsup inequality) For every $u \in L^2(\Omega)$, there exists a sequence $(u_k)_k \subset L^2(\Omega)$ such that $u_k \rightharpoonup u$ in $L^2(\Omega)$ and*
$$\mathcal{F}(u) \ge \limsup_{k\to\infty} \mathcal{F}_k(u_k).$$

The sequence $(\mathcal{F}_k)_k$ converges in the sense of Mosco-convergence in $L^2(\Omega)$ to $\mathcal{F}$, written $\mathcal{F} = \mathrm{Mosc}(L^2)\text{-}\lim_{k\to\infty} \mathcal{F}_k$, if, in addition, the limsup inequality can be realised by a sequence converging strongly in $L^2(\Omega)$.

If the liminf inequality holds, then the sequence from the limsup inequality automatically satisfies $\lim_{k\to\infty} \mathcal{F}_k(u_k) = \mathcal{F}(u)$, and is therefore often called a recovery sequence. We note that the above sequential definition of Γ-convergence coincides with the topological definition [80, Proposition 8.10] for equi-coercive sequences $(\mathcal{F}_k)_k$, i.e., $\mathcal{F}_k \geq \Psi$ for all $k \in \mathbb{N}$ and for some $\Psi : L^2(\Omega) \to [0,\infty]$ with $\Psi(u) \to \infty$ as $\|u\|_{L^2(\Omega)} \to \infty$. In particular, the theory implies that the Γ-limit $\mathcal{F}$ is (sequentially) L^2-weakly lower semicontinuous. The Γ-convergence has the key property of yielding the convergence of solutions (if they exist) to those of the limit problem, which makes it a suitable notion of variational convergence. Precisely, if u_k is a minimizer of $\mathcal{F}_k$ for all $k \in \mathbb{N}$ and u a cluster point of the sequence $(u_k)_k$, then u is a minimizer of $\mathcal{F}$ and $\min_{L^2(\Omega)} \mathcal{F}_k = \mathcal{F}_k(u_k) \to \mathcal{F}(u) = \min_{L^2(\Omega)} \mathcal{F}$, see [80, Corollary 7.20]. Notice that the existence of cluster points is implied by the assumption of equi-coercivity. In the special case when $(\mathcal{F}_k)_k$ is a constant sequence of functionals, say $\mathcal{F}_k = \mathcal{G}$ for all $k \in \mathbb{N}$, the Γ-limit corresponds to the relaxation of $\mathcal{G}$, i.e., its L^2-weakly lower semicontinuous envelope. Observe that replacing each $\mathcal{F}_k$ by its relaxation does not affect the Γ-limit of $(\mathcal{F}_k)_k$, see [80, Proposition 6.11].

As we discuss next, weak L^2-stability provides existence of solutions to the training scheme $(\mathcal{T})$. We note that the family of functionals $\{\mathcal{J}_\lambda\}_{\lambda\in\Lambda}$ as in (7.3) is equi-coercive in a componentwise sense.

Theorem 7.2.3. *Let $\mathcal{J}_\lambda : L^2(\Omega) \to [0,\infty]^N$ be given by* (7.3) *for each $\lambda \in \Lambda$. If the family $\{\mathcal{J}_\lambda\}_{\lambda\in\Lambda}$ is weakly L^2-stable, then $\mathcal{I}$ in $(\mathcal{T})$ has a minimizer.*

Proof. The statement follows directly from the direct method and the classical properties of Γ-convergence.

Let $(\lambda_k)_k \subset \Lambda$ be a minimizing sequence for $\mathcal{I}$. Then, for each $k \in \mathbb{N}$, there is $w_k \in K_{\lambda_k}$ such that

$$\lim_{k\to\infty} \|w_k - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} = \inf_{\lambda\in\Lambda} \mathcal{I}(\lambda). \tag{7.4}$$

In particular, $(w_k)_k$ is uniformly bounded in $L^2(\Omega;\mathbb{R}^N)$; hence, extracting a subsequence if necessary, one may assume that $w_k \rightharpoonup w$ in $L^2(\Omega;\mathbb{R}^N)$ as $k \to \infty$ for some $w \in L^2(\Omega;\mathbb{R}^N)$. Using the equi-coercivity, we apply the compactness result for Γ-limits [80, Corollary 8.12] to find a further subsequence of $(\lambda_k)_k$ (not relabeled) such that $(\mathcal{J}_{\lambda_k,j})_k$ $\Gamma(w\text{-}L^2)$-converges for all $j \in \{1,\dots,N\}$. Consequently, by the weak L^2-stability assumption and the properties of Γ-convergence on minimizing sequences, there exists $\tilde{\lambda} \in \Lambda$ such that $w \in K_{\tilde{\lambda}}$. Then, along with (7.4),

$$\mathcal{I}(\tilde{\lambda}) \leq \|w - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} \leq \liminf_{k\to\infty} \|w_k - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} = \inf_{\lambda\in\Lambda} \mathcal{I}(\lambda) \leq \mathcal{I}(\tilde{\lambda}),$$

which finishes the proof. □

Remark 7.2.4. We give a simple counterexample to illustrate that minimizers for $\mathcal{I}$ may not exist in general. Take $\Lambda = (0,\infty) \subset \mathbb{R}$, a single data point (u^c, u^η) with $u^c = u^\eta \neq 0$, and $\mathcal{R}_\lambda(u) = \lambda\|u\|^2_{L^2(\Omega)}$ for $\lambda \in \Lambda$. Then, $\mathcal{J}_\lambda(u) = \|u - u_\eta\|^2_{L^2(\Omega)} + \lambda\|u\|^2_{L^2(\Omega)}$ for $u \in L^2(\Omega)$ and $K_\lambda = \{u^\eta/(1+\lambda)\} = \{u^c/(1+\lambda)\}$, so that

$$\mathcal{I}(\lambda) = \left(\frac{\lambda}{1+\lambda}\right)^2 \|u^c\|^2_{L^2(\Omega)},$$

which does not have a minimizer on $\Lambda = (0,\infty)$. By the previous theorem, the family must fail to be weakly L^2-stable. Indeed, $\Gamma(w\text{-}L^2)\text{-}\lim_{\lambda\to 0} \mathcal{J}_\lambda$ coincides with the pointwise limit and is equal to $\|\cdot - u_\eta\|^2_{L^2(\Omega)}$, which is not an element of $\{\mathcal{J}_\lambda\}_{\lambda\in(0,\infty)}$. △

Theorem 7.2.3 is useful in many situations, including the basic case when the parameter set Λ is a compact real interval. However, weak L^2-stability is not always guaranteed, as Remark 7.2.4

illustrates. If, for instance, we have a sequence $(\lambda_k)_k$ converging to a point in X outside Λ, then there is no reason to expect that

$$\Gamma(w\text{-}L^2)\text{-}\lim_{k\to\infty} \mathcal{J}_{\lambda_k,j} = \mathcal{J}_{\lambda,j}$$

holds for some $\lambda \in \Lambda$.

To overcome this issue and provide a more general existence framework, we will look at a suitable replacement of the bi-level scheme. In the following, we denote by $\overline{\Lambda}$ the closure of Λ and suppose that for each $j \in \{1, \ldots, N\}$ and $\lambda \in \overline{\Lambda}$, the Γ-limits

$$\overline{\mathcal{J}}_{\lambda,j} := \Gamma(w\text{-}L^2)\text{-}\lim_{\lambda'\to\lambda} \mathcal{J}_{\lambda',j} \tag{7.5}$$

exist, where λ' takes values on an arbitrary sequence in Λ. We further set

$$\overline{\mathcal{J}}_\lambda := (\overline{\mathcal{J}}_{\lambda,1}, \ldots, \overline{\mathcal{J}}_{\lambda,N}) : \overline{\Lambda} \to [0,\infty]^N.$$

Based on these definitions, we introduce $\overline{\mathcal{I}} : \overline{\Lambda} \to [0,\infty]$ as the extension of the upper level functional $\mathcal{I}$ given by

$$\overline{\mathcal{I}}(\lambda) := \min_{w\in\overline{K}_\lambda} \|w - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)}, \tag{7.6}$$

where $\overline{K}_{\lambda,j} := \operatorname{argmin}_{u\in L^2(\Omega)} \overline{\mathcal{J}}_{\lambda,j}(u)$ and $\overline{K}_\lambda := \overline{K}_{\lambda,1} \times \overline{K}_{\lambda,2} \times \cdots \times \overline{K}_{\lambda,N}$ for $\lambda \in \overline{\Lambda}$. Observe that $\overline{K}_{\lambda,j}$ is L^2-weakly closed because the functional $\overline{\mathcal{J}}_{\lambda,j}$, as a $\Gamma(w\text{-}L^2)$-limit by (7.5), is L^2-weakly lower semicontinuous. Hence, the minimum in the definition of $\overline{\mathcal{I}}$ is actually attained. Notice that taking constant sequences in the parameter space in (7.5) and using the weak lower semicontinuity of the regularizers $\mathcal{R}_\lambda$ in (H), we conclude that $\overline{\mathcal{J}}_\lambda$ coincides with $\mathcal{J}_\lambda$ whenever $\lambda \in \Lambda$. In that sense, we can think of $\{\overline{\mathcal{J}}_\lambda\}_{\lambda\in\overline{\Lambda}}$ as the extension of the family $\{\mathcal{J}_\lambda\}_{\lambda\in\Lambda}$ to the closure of Λ.

All together, this leads to the extended bi-level problem

$$\begin{aligned} &\text{(Upper-level)} \qquad \text{Minimize } \overline{\mathcal{I}}(\lambda) := \min_{w\in\overline{K}_\lambda} \|w - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} \text{ over } \lambda \in \overline{\Lambda}, \\ &\text{(Lower-level)} \qquad \overline{K}_\lambda := \Big\{ w \in L^2(\Omega;\mathbb{R}^N) : w_j \in \operatorname*{arg\,min}_{u\in L^2(\Omega)} \overline{\mathcal{J}}_{\lambda,j}(u) \text{ for all } j \in \{1,\ldots,N\} \Big\}. \end{aligned} \tag{$\overline{\mathcal{T}}$}$$

The theorem below compares the extended upper level functional $\overline{\mathcal{I}}$ with the relaxation of $\mathcal{I}$ (after trivial extension to $\overline{\Lambda}$ by ∞), that is, with its lower semicontinuous envelope $\mathcal{I}^{\mathrm{rlx}} : \overline{\Lambda} \to [0,\infty]$ given by

$$\mathcal{I}^{\mathrm{rlx}}(\lambda) := \inf\Big\{ \liminf_{k\to\infty} \mathcal{I}(\lambda_k) : (\lambda_k)_k \subset \Lambda, \lambda_k \to \lambda \text{ in } \overline{\Lambda} \Big\}. \tag{7.7}$$

As we will see, the key assumption to obtain the equality between $\overline{\mathcal{I}}$ and $\mathcal{I}^{\mathrm{rlx}}$ is the Mosco-convergence of the family of regularizers in (7.9), which is stronger than the Γ-convergence of the reconstruction functionals in (7.5). It even implies the Mosco-convergence

$$\overline{\mathcal{J}}_{\lambda,j} = \mathrm{Mosc}(L^2)\text{-}\lim_{\lambda'\to\lambda} \mathcal{J}_{\lambda',j}$$

and, in this case, the limit passage can be performed additively in the fidelity and regularizing term; thus, for all $j \in \{1, \ldots, N\}$, we have

$$\overline{\mathcal{J}}_{\lambda,j}(u) = \|u - u_j^\eta\|_{L^2(\Omega)} + \overline{\mathcal{R}}_\lambda(u) \qquad \text{for } u \in L^2(\Omega). \tag{7.8}$$

Theorem 7.2.5. *Consider the bi-level optimization problems* ($\mathcal{T}$) *and* ($\overline{\mathcal{T}}$), *assume* (7.5), *and recall the definitions in* (7.6) *and* (7.7). *Suppose in addition that*

(i) *the Mosco-limits*

$$\overline{\mathcal{R}}_\lambda := \mathrm{Mosc}(L^2)\text{-}\lim_{\lambda'\to\lambda} \mathcal{R}_{\lambda'} \tag{7.9}$$

exist for each $\lambda \in \overline{\Lambda}$, *with* λ' *taking values on sequences in* Λ, *and*

(ii) $\overline{K}_\lambda$ *is a singleton for every* $\lambda \in \overline{\Lambda} \setminus \Lambda$.

Then, the extension $\overline{\mathcal{I}}$ *of* $\mathcal{I}$ *to the closure* $\overline{\Lambda}$ *coincides with the relaxation of* $\mathcal{I}$, *i.e.,* $\overline{\mathcal{I}} = \mathcal{I}^{\mathrm{rlx}}$ *on* $\overline{\Lambda}$.

Proof. To show that $\overline{\mathcal{I}} \le \mathcal{I}^{\mathrm{rlx}}$, we take $\lambda \in \overline{\Lambda}$ and let $(\lambda_k)_k \subset \Lambda$ with $\lambda_k \to \lambda$ in $\overline{\Lambda}$ be an admissible sequence for $\mathcal{I}^{\mathrm{rlx}}(\lambda)$ in (7.7). We may even assume that $\infty > \liminf_{k\to\infty} \mathcal{I}(\lambda_k) = \lim_{k\to\infty} \mathcal{I}(\lambda_k)$. Then, recalling (7.2) and fixing $\delta > 0$, we can find $w_k \in K_{\lambda_k}$ such that

$$\lim_{k\to\infty} \mathcal{I}(\lambda_k) \ge \liminf_{k\to\infty} \|w_k - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} - \delta.$$

In particular, $(w_k)_k$ is uniformly bounded in $L^2(\Omega;\mathbb{R}^N)$, which allows us to extract an L^2-weakly converging subsequence (not relabeled) with limit $\bar{w} \in L^2(\Omega;\mathbb{R}^N)$. By the properties of Γ-convergence on cluster points of minimizing sequences recalled above (see also [80, Corollary 7.20]), we infer from (7.5) that $\bar{w}_j \in \operatorname{argmin}_{u\in L^2(\Omega)} \overline{\mathcal{J}}_{\lambda,j}(u)$ for all $j \in \{1,\cdots,N\}$; in other words, $\bar{w} \in \overline{K}_\lambda$. Thus,

$$\lim_{k\to\infty} \mathcal{I}(\lambda_k) \ge \|\bar{w} - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} - \delta \ge \overline{\mathcal{I}}(\lambda) - \delta.$$

By letting $\delta \to 0$ first, and then taking the infimum over all admissible sequences for $\mathcal{I}^{\mathrm{rlx}}(\lambda)$ in (7.7), it follows that $\overline{\mathcal{I}}(\lambda) \le \mathcal{I}^{\mathrm{rlx}}(\lambda)$.

To prove the reverse inequality, we start by recalling that for $\lambda \in \Lambda$, $\mathcal{J}_\lambda$ is weakly L^2-lower semicontinuous by Assumption (H); thus, (2.5) yields $\overline{\mathcal{J}}_\lambda = \mathcal{J}_\lambda$ for $\lambda \in \Lambda$. Consequently, $\overline{\mathcal{I}}(\lambda) = \mathcal{I}(\lambda) \ge \mathcal{I}^{\mathrm{rlx}}(\lambda)$ for $\lambda \in \Lambda$. We are then left to consider $\lambda \in \overline{\Lambda} \setminus \Lambda$ and find a sequence $(\lambda_k)_k \subset \Lambda$ converging to λ in $\overline{\Lambda}$ and satisfying $\liminf_{k\to\infty} \mathcal{I}(\lambda_k) \le \overline{\mathcal{I}}(\lambda)$. To that end, take any $(\lambda_k)_k \subset \Lambda$ with $\lambda_k \to \lambda$ in $\overline{\Lambda}$, and let $w_k \in K_{\lambda_k}$ for $k \in \mathbb{N}$. Recalling (ii), denote by $w_\lambda = (w_{\lambda,1},\ldots,w_{\lambda,N})$ the unique element in $\overline{K}_\lambda$. Then, using (7.5) and the equi-coercivity of $(\mathcal{J}_\lambda)_{\lambda\in\Lambda}$, we obtain by the theory of Γ-convergence (see [80, Corollary 7.24]) that $(w_k)_k$ converges weakly in $L^2(\Omega;\mathbb{R}^N)$ to w_λ; moreover, it holds for all $j \in \{1,\ldots,N\}$ that

$$\mathcal{J}_{\lambda_k,j}(w_{k,j}) \to \overline{\mathcal{J}}_{\lambda,j}(w_{\lambda,j}) \qquad \text{as } k \to \infty. \tag{7.10}$$

The following shows that $(w_k)_k$ converges even strongly in $L^2(\Omega;\mathbb{R}^N)$. Indeed, fixing $j \in \{1,\ldots,N\}$, we infer from (7.10) along with the Mosco-convergence of the regularizers in (i) and (7.8) that

$$\begin{aligned}\|w_{\lambda,j} - u_j^\eta\|^2_{L^2(\Omega)} + \overline{\mathcal{R}}_\lambda(w_{\lambda,j}) = \overline{\mathcal{J}}_{\lambda,j}(w_{\lambda,j}) &= \lim_{k\to\infty} \mathcal{J}_{\lambda_k,j}(w_{k,j})\\ &= \lim_{k\to\infty}\Big[\|w_{k,j} - u_j^\eta\|^2_{L^2(\Omega)} + \mathcal{R}_{\lambda_k}(w_{k,j})\Big]\\ &\ge \limsup_{k\to\infty} \|w_{k,j} - u_j^\eta\|^2_{L^2(\Omega)} + \overline{\mathcal{R}}_\lambda(w_{\lambda,j}).\end{aligned}$$

Hence, $\|w_{\lambda,j} - u_j^\eta\|^2_{L^2(\Omega)} \ge \limsup_{k\to\infty} \|w_{k,j} - u_j^\eta\|^2_{L^2(\Omega)}$, which together with the weak lower semicontinuity of the L^2-norm yields

$$\lim_{k\to\infty} \|w_{k,j} - u_j^\eta\|^2_{L^2(\Omega)} = \|w_{\lambda,j} - u_j^\eta\|^2_{L^2(\Omega)};$$

thus, $w_k \to w_\lambda$ strongly in $L^2(\Omega;\mathbb{R}^N)$ using the combination of weak convergence and convergence of norms by the Radon–Riesz property. With this, we finally conclude that

$$\liminf_{k\to\infty} \mathcal{I}(\lambda_k) \leq \liminf_{k\to\infty} \|w_k - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} = \|w_\lambda - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} = \min_{w\in\overline{K}_\lambda} \|w - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} = \overline{\mathcal{I}}(\lambda),$$

finishing the proof. □

Remark 7.2.6. By inspecting the proof, it becomes clear that the estimate $\overline{\mathcal{I}} \leq \mathcal{I}^{\mathrm{rlx}}$ holds without the additional assumptions (i) and (ii) from the previous theorem; in other words, $\overline{\mathcal{I}}$ always provides a lower bound for the relaxation of $\mathcal{I}$. △

The identity $\overline{\mathcal{I}} = \mathcal{I}^{\mathrm{rlx}}$ may fail if either of the assumptions (i) or (ii) in Theorem 7.2.5 is dropped as the following example shows.

Example 7.2.7. a) To see why (i) is necessary, consider $\Lambda = (0, 1]$, a single data point (u^c, u^η) with $u^c = u^\eta = 0$, and

$$\mathcal{R}_\lambda = \frac{1}{\lambda}\|\cdot - v_\lambda\|^2_{L^2(\Omega)} \qquad \text{with } v_\lambda = v(\cdot/\lambda) \in L^2(\Omega)$$

for a given $v \in L^\infty(\mathbb{R}^n)$ with the properties that v is $(0,1)^n$-periodic, $v \in \{-1, 1\}$ almost everywhere, and $\int_{(0,1)^n} v\,dx = 0$. Under these specifications, the Γ-limits $\overline{\mathcal{J}}_\lambda = \Gamma(w\text{-}L^2)\text{-}\lim_{\lambda'\to\lambda} \mathcal{J}_{\lambda'}$ (cf. (7.5) and (7.1)) exist and are given by

$$\overline{\mathcal{J}}_\lambda(u) = \begin{cases} \|u\|^2_{L^2(\Omega)} + \frac{1}{\lambda}\|u - v_\lambda\|^2_{L^2(\Omega)} & \text{for } \lambda \in (0, 1], \\ |\Omega| + \chi_{\{0\}}(u) & \text{for } \lambda = 0, \end{cases} \tag{7.11}$$

where χ_E denotes the indicator function of a set $E \subset L^2(\Omega)$, i.e.,

$$\chi_E(u) = \begin{cases} 0 & \text{if } u \in E, \\ \infty & \text{if } u \notin E, \end{cases} \qquad \text{for } u \in L^2(\Omega).$$

The non-trivial case is when $\lambda = 0$. In this case, we observe that we can take $(v_{\lambda'})_{\lambda'}$ as a recovery sequence for $u = 0$ because it converges weakly in $L^2(\Omega)$ as $\lambda' \to 0$ to $\int_{(0,1)^n} v\,dx = 0$ by the Riemann–Lebesgue lemma for periodically oscillating sequences. For the liminf inequality, let $u_{\lambda'} \rightharpoonup u$ as $\lambda' \to 0$ and suppose without loss of generality that $\sup_{\lambda'} \mathcal{R}_{\lambda'}(u_{\lambda'}) < \infty$. Then, $u_{\lambda'} = v_{\lambda'} + r_{\lambda'}$ with $r_{\lambda'} \to 0$ in $L^2(\Omega)$ as $\lambda' \to 0$, which implies $u = 0$ and, recalling that $v \in \{-1, 1\}$ almost everywhere,

$$\liminf_{\lambda'\to 0} \mathcal{J}_{\lambda'}(u_{\lambda'}) \geq \lim_{\lambda'\to 0} \|v_{\lambda'} + r_{\lambda'}\|^2_{L^2(\Omega)} = \lim_{\lambda'\to 0} \|v_{\lambda'}\|^2_{L^2(\Omega)} = |\Omega| = \overline{\mathcal{J}}_0(0),$$

which completes the proof of (7.11) when $\lambda = 0$.

In view of (7.11), one can now read off that $K_\lambda = \overline{K}_\lambda = \{v_\lambda/(1 + \lambda)\}$ for $\lambda \in (0, 1]$ and $\overline{K}_0 = \{0\}$. In particular, condition (ii) on the uniqueness of minimizers of the extended lower-level problem is fulfilled here. Hence,

$$\mathcal{I}(\lambda) = \Big(\frac{1}{1+\lambda}\Big)^2 |\Omega| \tag{7.12}$$

for $\lambda \in (0, 1]$, and

$$\overline{\mathcal{I}}(\lambda) = \begin{cases} \Big(\frac{1}{1+\lambda}\Big)^2 |\Omega| & \text{if } \lambda \in (0, 1], \\ 0 & \text{if } \lambda = 0 \end{cases}$$

for $\lambda \in [0,1]$. It is immediate to see from (7.12) that

$$\overline{\mathcal{I}}(0) = 0 < |\Omega| = \mathcal{I}^{\text{rlx}}(0).$$

Notice that this example hinges on the fact that the minimizers $v_\lambda/(1+\lambda)$ only converge weakly as $\lambda \to 0$, which, in view of the proof of Theorem 7.2.5, implies that the family of regularizers $\{\mathcal{R}_\lambda\}_{\lambda\in\Lambda}$ does not Mosco-converge in $L^2(\Omega)$ in the sense of (7.9), thus failing to satisfy (i).

b) For the necessity of (ii), consider $\Lambda = (0,1]$, a single data point (u^c, u^η) with $u^c = 0$ and $\|u^\eta\|^2_{L^2(\Omega)} = 1$, and

$$\mathcal{R}_\lambda(u) = \begin{cases} \lambda & \text{if } u = 0, \\ 1 & \text{if } u \neq 0. \end{cases}$$

While it is straightforward to check that condition (i) in Theorem 7.2.5 regarding the Mosco-limits of $\{\mathcal{R}_\lambda\}_{\lambda\in\Lambda}$ is satisfied with

$$\overline{\mathcal{R}}_\lambda(u) = \begin{cases} \lambda & \text{if } u = 0, \\ 1 & \text{if } u \neq 0 \end{cases}$$

for $\lambda \in [0,1]$, which clearly coincides with $\mathcal{R}_\lambda$ for $\lambda \in \Lambda = (0,1]$, condition (ii) fails. Indeed, it follows from (7.8) that $\overline{\mathcal{J}}_\lambda(u^\eta) = \overline{\mathcal{R}}_\lambda(u^\eta) = 1$ and $\overline{\mathcal{J}}_\lambda(0) = \|u^\eta\|^2_{L^2(\Omega)} + \lambda = 1 + \lambda$ for all $\lambda \in [0,1]$. Consequently, for $\lambda \in (0,1]$, we have $\overline{\mathcal{J}}_\lambda = \mathcal{J}_\lambda$ and u^η is its unique minimizer; in contrast, for $\lambda = 0$, $\overline{\mathcal{J}}_0$ has two minimizers, namely $\overline{K}_0 = \{u^\eta, 0\} = \{u^\eta, u^c\}$. Finally, we observe that the conclusion of Theorem 7.2.5 fails here because

$$\overline{\mathcal{I}}(0) = 0 \qquad \text{and} \qquad \mathcal{I}(\lambda) = \|u^c - u^\eta\|^2_{L^2(\Omega)} = 1 \;\text{ for all } \lambda \in (0,1],$$

which yields $\overline{\mathcal{I}}(0) = 0 < 1 = \mathcal{I}^{\text{rlx}}(0)$.

The following result is a direct consequence of Theorem 7.2.5 and standard properties of relaxation.

Corollary 7.2.8. *Under the assumptions of Theorem 7.2.5 and if* $\overline{\Lambda}$ *is compact, it holds that:*

(i) *The extension* $\overline{\mathcal{I}}$ *has at least one minimizer and*

$$\min_{\overline{\Lambda}} \overline{\mathcal{I}} = \inf_{\Lambda} \mathcal{I}.$$

(ii) *Any minimizing sequence* $(\lambda_k)_k \subset \Lambda$ *of* $\mathcal{I}$ *converges up to subsequence to a minimizer* $\lambda \in \overline{\Lambda}$ *of* $\overline{\mathcal{I}}$.

(iii) *If* $\lambda \in \Lambda$ *minimizes* $\overline{\mathcal{I}}$*, then* λ *is also a minimizer of* $\mathcal{I}$.

We conclude this section on the theoretical framework with a brief comparison with related works on optimal control problems. By setting $K = \{(w,\lambda) \in L^2(\Omega) \times \Lambda \,:\, w \in K_\lambda\}$, the bi-level optimization problem $(\mathcal{T})$ can be equivalently rephrased into minimizing

$$\widehat{\mathcal{I}}(u,\lambda) = \|u - u^c\|^2_{L^2(\Omega)} + \chi_K(u,\lambda), \qquad (u,\lambda) \in L^2(\Omega) \times \Lambda,$$

as a functional of two variables; observe that

$$\mathcal{I}(\lambda) = \inf_{w \in L^2(\Omega)} \widehat{\mathcal{I}}(w,\lambda).$$

Similar functionals and their relaxations have been studied in the literature, including [40, 55, 56]. Especially the paper [40] by Belloni, Buttazzo, & Freddi, where the authors propose to extend the control space to its closure and find a description of the relaxed optimal control problem, shares many parallels with our results. Apart from some differences in the assumptions and abstract set-up, the main reason why their results are not applicable here is the continuity condition of the cost functional with respect to the state variable [40, Eq. (2.11)]. In our setting, this would translate into weak continuity of the L^2-norm, which is clearly false. The argument in the proof of Theorem 7.2.5 exploiting the Mosco-convergence of the regularizers (see (7.9)) is precisely what circumvents this issue.

7.3 Learning the optimal weight of the regularization term

In this section, we study the optimization of a weight factor, often called tuning parameter, in front of a fixed regularization term. Such tuning parameters are typically employed in practical implementations of variational denoising models to adjust the best level of regularization. This setting constitutes a simple, yet non-trivial, application of our general theory and therefore helps to exemplify the abstract results from the previous section.

As above, $\Omega \subset \mathbb{R}^n$ is a bounded open set and $u^c, u^\eta \in L^2(\Omega; \mathbb{R}^N)$ are the given data representing pairs of clean and noisy images. We take $\Lambda = (0, \infty)$ describing the range of a weight factor and, to distinguish the various parameters throughout this paper, denote by α an arbitrary point in $\overline{\Lambda} = [0, \infty]$. For a fixed map $\mathcal{R} : L^2(\Omega) \to [0, \infty]$ with the properties that

(H1$_\alpha$) $\mathcal{R}$ is convex, vanishes exactly on constant functions, and $\operatorname{Dom} \mathcal{R}$ is dense in $L^2(\Omega)$,

(H2$_\alpha$) $\mathcal{R}$ is lower semicontinuous on $L^2(\Omega)$,

we define the weighted regularizers

$$\mathcal{R}_\alpha = \alpha \mathcal{R} \qquad \text{for } \alpha \in (0, \infty). \tag{7.13}$$

Note that (H1$_\alpha$) and (H2$_\alpha$) imply that the family $\{\mathcal{R}_\alpha\}_{\alpha \in (0,\infty)}$ satisfies (H) because convexity and lower semicontinuity yield weak lower semicontinuity, making this setting match with the framework of Section 7.2.

Following the definition of the training scheme ($\mathcal{T}$), we introduce here for $\alpha \in (0, \infty)$ and $j \in \{1, \dots, N\}$ the reconstruction functionals

$$\mathcal{J}_{\alpha,j}(u) = \|u - u_j^\eta\|^2_{L^2(\Omega)} + \mathcal{R}_\alpha(u) \qquad \text{for } u \in L^2(\Omega),$$

cf. (7.1), and consider accordingly the upper level functional $\mathcal{I} : (0, \infty) \to [0, \infty)$ given by

$$\mathcal{I}(\alpha) = \inf_{w \in K_\alpha} \|w - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} \quad \text{for } \alpha \in (0, \infty), \tag{7.14}$$

with $K_\alpha = K_{\alpha,1} \times \ldots \times K_{\alpha,N}$ and $K_{\alpha,j} = \operatorname{arg\,min}_{u \in L^2(\Omega)} \mathcal{J}_{\alpha,j}(u)$, cf. (7.2). Further, the following set of hypotheses on the training data will play a crucial role for our main result in this section (Theorem 7.3.2):

(H3$_\alpha$) It holds that

$$\sum_{j=1}^N \mathcal{R}(u_j^c) < \sum_{j=1}^N \mathcal{R}(u_j^\eta);$$

($H4_\alpha$) the data u^η and u^c satisfy

$$\|u^\eta - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} < \left\| \fint_\Omega u^\eta \, dx - u^c \right\|^2_{L^2(\Omega;\mathbb{R}^N)}.$$

Remark 7.3.1 (Discussion of the hypotheses ($H1_\alpha$)–($H4_\alpha$)). a) Note that ($H1_\alpha$) implies that the set of minimizers for the reconstruction functionals, K_α, has cardinality one, owing to the convexity of $\mathcal{R}$ and the strict convexity of the fidelity term, considering also that $\mathcal{J}_{\alpha,j} \not\equiv \infty$. In the following, we write $w^{(\alpha)} = (w_1^{(\alpha)}, \dots, w_N^{(\alpha)}) \in L^2(\Omega;\mathbb{R}^N)$ for the single element of K_α, i.e., $K_\alpha = \{w^{(\alpha)}\}$.

b) An example of a nonlocal regularizer satisfying ($H1_\alpha$) and ($H2_\alpha$) is

$$\mathcal{R}(u) := \int_\Omega \int_\Omega a(x,y)\, g(u(x) - u(y)) \, dx \, dy \quad \text{for } u \in L^2(\Omega),$$

where $g : \mathbb{R} \to [0,\infty)$ is a convex function such that $g^{-1}(0) = \{0\}$ and $a : \Omega \times \Omega \to [0,\infty]$ is a suitable kernel ensuring that $C_c^\infty(\Omega) \subset \operatorname{Dom} \mathcal{R}$. As an explicit choice, one can take $g(t) = t^p$ for $t \in \mathbb{R}$ and $a(x,y) = |y-x|^{-n-sp}$ for $x, y \in \Omega$ with some $s \in (0,1)$ and $p \geq 1$, which corresponds to a fractional Sobolev regularization.

c) Assumption ($H3_\alpha$) asserts that the regularizer penalizes the noisy images more than the clean ones on average. This is a natural condition because any good regularizer should reflect the prior knowledge on the training data, favoring the clean images.

d) The second condition on the data, ($H4_\alpha$), means that the noisy image lies closer to the clean image than its mean value, which can be considered a reasonable assumption in the case of moderate noise and a non-trivial ground truth. Indeed, suppose the noise is bounded by $\|u_j^\eta - u_j^c\|_{L^2(\Omega)} \leq \delta$ for all $j \in \{1,\dots,N\}$ and some $\delta > 0$; then, ($H4_\alpha$) is satisfied if

$$\left\| \fint_\Omega u_j^c \, dx - u_j^c \right\|_{L^2(\Omega)} > \delta\big(1 + |\Omega|^{-\frac{1}{2}}\big) \quad \text{for all } j \in \{1,\dots,N\}$$

because

$$\begin{aligned}
\left\| \fint_\Omega u_j^\eta \, dx - u_j^c \right\|_{L^2(\Omega)} &\geq \left\| \fint_\Omega u_j^c \, dx - u_j^c \right\|_{L^2(\Omega)} - \left\| \fint_\Omega (u_j^\eta - u_j^c) \, dx \right\|_{L^2(\Omega)} \\
&> \delta\big(1 + |\Omega|^{-\frac{1}{2}}\big) - |\Omega|^{-\frac{1}{2}} \|u_j^\eta - u_j^c\|_{L^2(\Omega)} \\
&\geq \delta \geq \|u_j^\eta - u_j^c\|_{L^2(\Omega)},
\end{aligned}$$

where the second inequality is due to Jensen's inequality. △

Next, we prove that the assumptions ($H1_\alpha$)–($H4_\alpha$) on the regularization term and on the training set give rise to optimal weight parameters that stay away from the extremal regimes, $\alpha = 0$ and $\alpha = \infty$. Thus, in this case, the bi-level parameter optimization procedure preserves the structure of the original denoising model.

Theorem 7.3.2 (Structure preservation). *Suppose that* ($H1_\alpha$)–($H4_\alpha$) *hold. Then, the learning scheme corresponding to the minimization of* $\mathcal{I}$ *in* (7.14) *admits a solution* $\bar{\alpha} \in (0,\infty)$.

A related statement in the same spirit can be found in [87, Theorem 1], although some of the details of the proof were not entirely clear to us. Our proof of Theorem 7.3.2 is based on a different approach and hinges on the following two lemmas, the first of which determines the Mosco-limits of the regularizers, and thereby provides an explicit formula of the extension $\overline{\mathcal{I}}$ of $\mathcal{I}$ as introduced in (7.6).

Proposition 7.3.3 (Mosco-convergence of the regularizer). *Let* $\mathcal{R} : L^2(\Omega) \to [0,\infty]$ *satisfy* (H1$_\alpha$) *and* (H2$_\alpha$), *and let* $\{\mathcal{R}_\alpha\}_{\alpha\in(0,\infty)}$ *be as in* (7.13). *Then,*

$$\overline{\mathcal{R}}_\alpha := \mathrm{Mosc}(L^2)\text{-}\lim_{\alpha'\to\alpha} \mathcal{R}_{\alpha'} = \begin{cases} \mathcal{R}_\alpha & \text{for } \alpha \in (0,\infty), \\ 0 & \text{for } \alpha = 0, \\ \chi_C & \text{for } \alpha = \infty, \end{cases} \tag{7.15}$$

for $\alpha \in [0,\infty]$, *where* χ_C *is the indicator function of* $C := \{u \in L^2(\Omega) : u \text{ is constant}\}$.

Proof. Using standard arguments, we show that the Mosco-limit of $(\mathcal{R}_{\alpha_k})_k$ exists for every sequence $(\alpha_k)_k$ of positive real numbers with $\alpha_k \to \alpha \in [0,\infty]$, and corresponds to the right hand side of (7.15).

Case 1: $\alpha \in (0,\infty)$. Using (H2$_\alpha$) for the liminf inequality and a constant recovery sequence for the upper bound, we conclude that the Mosco-limit of $(\mathcal{R}_{\alpha_k})_k$ coincides with $\mathcal{R}_\alpha$.

Case 2: $\alpha = 0$. The liminf inequality is trivial. For the recovery sequence, take $u \in L^2(\Omega)$ and let $(u_k)_k \subset \mathrm{Dom}\,\mathcal{R}$ converge strongly to u in $L^2(\Omega)$, which is feasible due to (H1$_\alpha$). By possibly repeating certain entries of the sequence $(u_k)_k$ (not relabeled), one can slowdown the speed at which $\mathcal{R}(u_k)$ potentially blows up and assume that $\alpha_k\mathcal{R}(u_k) \to 0$ as $k \to \infty$. Thus,

$$\lim_{k\to\infty} \mathcal{R}_{\alpha_k}(u_k) = \lim_{k\to\infty} \alpha_k\mathcal{R}(u_k) = 0.$$

Case 3: $\alpha = \infty$. The limsup inequality follows by choosing constant recovery sequences. For the proof of the lower bound, consider $u_k \rightharpoonup u$ in $L^2(\Omega)$ with $r := \sup_{k\in\mathbb{N}} \alpha_k\mathcal{R}(u_k) = \sup_{k\in\mathbb{N}} \mathcal{R}_{\alpha_k}(u_k) < \infty$. Then, along with the weak lower semicontinuity of $\mathcal{R}$ (see Remark 7.3.1 a)),

$$\mathcal{R}(u) \le \liminf_{k\to\infty} \mathcal{R}(u_k) \le \lim_{k\to\infty} \frac{r}{\alpha_k} = 0.$$

This shows that $\mathcal{R}(u) = 0$, which implies by the assumption on the zero level set of $\mathcal{R}$ in (H1$_\alpha$) that u is constant, i.e., $u \in C$. □

As a consequence of the previous proposition, we deduce that the extension $\overline{\mathcal{I}} : \overline{\Lambda} \to [0,\infty]$ of $\mathcal{I}$ in the sense of (7.6) can be explicitly determined as

$$\overline{\mathcal{I}}(\alpha) = \begin{cases} \mathcal{I}(\alpha) & \text{for } \alpha \in (0,\infty), \\ \|u^\eta - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} & \text{for } \alpha = 0, \\ \left\| \fint_\Omega u^\eta \, dx - u^c \right\|^2_{L^2(\Omega;\mathbb{R}^N)} & \text{for } \alpha = \infty. \end{cases} \tag{7.16}$$

Indeed, a straight-forward calculation of the unique componentwise minimizer of the extended reconstruction functionals $\overline{\mathcal{J}}_\alpha$ at the boundary points $\alpha = 0$ and $\alpha = \infty$ leads to

$$\overline{K}_0 = \{u^\eta\} \quad \text{and} \quad \overline{K}_\infty = \left\{ \fint_\Omega u^\eta \, dx \right\}.$$

Since the assumptions (i) and (ii) of Theorem 7.2.5 are satisfied, $\overline{\mathcal{I}}$ coincides with the relaxation $\mathcal{I}^{\mathrm{rlx}}$. By Corollary 7.2.8 ($i$), $\overline{\mathcal{I}}$ attains its minimum at some $\bar\alpha \in [0,\infty]$. The degenerate cases $\bar\alpha \in \{0,\infty\}$ cannot be excluded a priori, but the next lemma shows that the minimum is attained in the interior $(0,\infty)$ under suitable assumptions on the training data.

Lemma 7.3.4. *Suppose that* (H1$_\alpha$) *and* (H2$_\alpha$) *hold, and let* $K_\alpha = \{w^{(\alpha)}\}$ *with* $w^{(\alpha)} = (w_1^{(\alpha)}, \dots, w_N^{\alpha}) \in L^2(\Omega;\mathbb{R}^N)$ *for* $\alpha \in (0,\infty)$, *cf. Remark 7.3.1 a).*

(*i*) *Under the additional assumption* (H3$_\alpha$)*, there exists* $\alpha \in (0,\infty)$ *such that*

$$\|w^{(\alpha)} - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} < \|u^\eta - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)}.$$

(*ii*) *Under the additional assumption* (H4$_\alpha$)*, there exists* $\alpha_0 \in (0,\infty)$ *such that, for all* $\alpha \in (0,\alpha_0)$*,*

$$\|w^{(\alpha)} - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} < \left\| \fint_\Omega u^\eta \, dx - u^c \right\|^2_{L^2(\Omega;\mathbb{R}^N)}. \tag{7.17}$$

Proof. We start by providing two useful auxiliary results about the asymptotic behavior of the reconstruction vector $w^{(\alpha)}$ as α tends to zero; precisely,

$$\lim_{\alpha\to 0}\|w^{(\alpha)} - u^\eta\|_{L^2(\Omega;\mathbb{R}^N)} = 0 \qquad \text{and} \qquad \lim_{\alpha\to 0}\mathcal{R}(w_j^{(\alpha)}) = \mathcal{R}(u_j^\eta) \text{ for every } j \in \{1,\dots,N\}. \tag{7.18}$$

Fix $j \in \{1,\dots,N\}$ and let $(\alpha_k)_k \subset (0,\infty)$ be such that $\alpha_k \to 0$ as $k \to \infty$. Take $u \in \operatorname{Dom}\mathcal{R}$ with $\|u - u_j^\eta\|^2_{L^2(\Omega)} \leq \varepsilon$ for some $\varepsilon > 0$, which is possible by (H1$_\alpha$). Then, the minimality of $w_j^{(\alpha_k)}$ for $\mathcal{J}_{\alpha_k,j}$ yields

$$\|w_j^{(\alpha_k)} - u_j^\eta\|^2_{L^2(\Omega)} \leq \mathcal{J}_{\alpha_k,j}(w_j^{(\alpha_k)}) \leq \mathcal{J}_{\alpha_k,j}(u) = \|u - u_j^\eta\|^2_{L^2(\Omega)} + \alpha_k\mathcal{R}(u) \leq \varepsilon + \alpha_k\mathcal{R}(u).$$

Since $\mathcal{R}(u) < \infty$, we find

$$\limsup_{k\to\infty}\|w_j^{(\alpha_k)} - u_j^\eta\|^2_{L^2(\Omega)} \leq \varepsilon,$$

which proves the first part of (7.18) due to the arbitrariness of ε. Exploiting the minimality of $w_j^{(\alpha)}$ for $\mathcal{J}_{\alpha,j}$ again with $\alpha \in (0,\infty)$ entails

$$\alpha\mathcal{R}(w_j^{(\alpha)}) = \mathcal{R}_\alpha(w_j^{(\alpha)}) \leq \mathcal{J}_{\alpha,j}(w_j^{(\alpha)}) \leq \mathcal{J}_{\alpha,j}(u_j^\eta) = \mathcal{R}_\alpha(u_j^\eta) = \alpha\mathcal{R}(u_j^\eta);$$

hence, $\mathcal{R}(w_j^{(\alpha)}) \leq \mathcal{R}(u_j^\eta)$ and, together with the first part of (7.18) and the lower semicontinuity of $\mathcal{R}$ by (H2$_\alpha$), it follows then that

$$\mathcal{R}(u_j^\eta) \geq \limsup_{k\to\infty}\mathcal{R}(w_j^{(\alpha_k)}) \geq \liminf_{k\to\infty}\mathcal{R}(w_j^{(\alpha_k)}) \geq \mathcal{R}(u_j^\eta).$$

Thus, $\lim_{k\to\infty}\mathcal{R}(w_j^{(\alpha_k)}) = \mathcal{R}(u_j^\eta)$, showing the second part of (7.18).

Regarding (*i*), we observe that the minimality of $w_j^{(\alpha)}$ for $\mathcal{J}_{\alpha,j}$ for any $\alpha \in (0,\infty)$ and $j \in \{1,\dots,N\}$ imposes the necessary condition $0 \in \partial\mathcal{J}_{\alpha,j}(w_j^{(\alpha)})$ or, equivalently,

$$2(u_j^\eta - w_j^{(\alpha)}) \in \partial\mathcal{R}_\alpha(w_j^{(\alpha)}) = \alpha\partial\mathcal{R}(w_j^{(\alpha)}),$$

where $\partial\mathcal{C}(u) \in L^2(\Omega)' \cong L^2(\Omega)$ is the subdifferential of a convex function $\mathcal{C} : L^2(\Omega) \to [0,\infty]$ at $u \in L^2(\Omega)$. Then,

$$\begin{aligned}\|u_j^\eta - u_j^c\|^2_{L^2(\Omega)} - \|w_j^{(\alpha)} - u_j^c\|^2_{L^2(\Omega)} &= 2\langle u_j^\eta - w_j^{(\alpha)}, w_j^{(\alpha)} - u_j^c\rangle_{L^2(\Omega)} + \|w_j^{(\alpha)} - u_j^\eta\|^2_{L^2(\Omega)} \\ &\geq \mathcal{R}_\alpha(w_j^{(\alpha)}) - \mathcal{R}_\alpha(u_j^c) = \alpha\big(\mathcal{R}(w_j^{(\alpha)}) - \mathcal{R}(u_j^c)\big),\end{aligned}$$

where $\langle\cdot,\cdot\rangle_{L^2(\Omega)}$ denotes the standard $L^2(\Omega)$-inner product. Summing both sides over $j \in \{1,\dots,N\}$ results in

$$\|u^\eta - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} - \|w^{(\alpha)} - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} \geq \alpha\sum_{j=1}^N\big(\mathcal{R}(w_j^{(\alpha)}) - \mathcal{R}(u_j^c)\big).$$

By (H3$_\alpha$) in combination with the second part of (7.18), there exists $\alpha_0 > 0$ such that

$$\textstyle\sum_{j=1}^N \mathcal{R}(w_j^{(\alpha)}) > \sum_{j=1}^N \mathcal{R}(u_j^c)$$

for all $\alpha \in (0, \alpha_0)$, so that choosing $\bar{\alpha} \in (0, \alpha_0)$ concludes the proof of (i).

To show (ii), we exploit the first limit in (7.18). Due to (H4$_\alpha$), it follows then for any $(\alpha_k)_k$ of positive real numbers with $\alpha_k \to 0$ as $k \to \infty$ that

$$\begin{aligned}\limsup_{k\to\infty} \|w^{(\alpha_k)} - u^c\|_{L^2(\Omega;\mathbb{R}^N)} &\le \limsup_{k\to\infty} \|w^{(\alpha_k)} - u^\eta\|_{L^2(\Omega;\mathbb{R}^N)} + \|u^\eta - u^c\|_{L^2(\Omega;\mathbb{R}^N)} \\ &< \Big\| \fint_\Omega u^\eta \, dx - u^c \Big\|_{L^2(\Omega;\mathbb{R}^N)},\end{aligned}$$

which gives rise to (7.17) for all k sufficiently large. □

Proof of Theorem 7.3.2. Since $\overline{\mathcal{I}}$ in (7.16) attains its infimum at a point $\bar{\alpha} \in (0, \infty)$ by Lemma 7.3.4, we conclude from Corollary 7.2.8 (iii) that $\bar{\alpha}$ is also a minimizer of $\mathcal{I}$. □

Let us finally remark that the assumptions (H3$_\alpha$) and (H4$_\alpha$) on the training data are necessary to obtain structure preservation in the sense of Theorem 7.3.2.

Remark 7.3.5. To see that (H3$_\alpha$) and (H4$_\alpha$) can generally not be dropped, consider, for example, a regularizer $\mathcal{R} : L^2(\Omega) \to [0, \infty]$ that satisfies (H1$_\alpha$) and (H2$_\alpha$) and is 2-homogeneous, i.e., $\mathcal{R}(\mu u) = \mu^2 \mathcal{R}(u)$ for all $u \in L^2(\Omega)$ and $\mu \in \mathbb{R}$. With a single, non-constant noisy image $u^\eta \in L^2(\Omega)$, so that $\mathcal{R}(u^\eta) \neq 0$, one has for any $\alpha \in (0, \infty)$ that the quadratic polynomial

$$\mu \mapsto \mathcal{J}_\alpha(\mu u^\eta) = (1 - \mu)^2 \|u^\eta\|^2_{L^2(\Omega)} + \mu^2 \alpha \mathcal{R}(u^\eta),$$

is not minimized at $\mu = 0$ or $\mu = 1$ because the derivative with respect to μ does not vanish there. Hence,

$$\mathcal{J}_\alpha(w^{(\alpha)}) < \mathcal{J}_\alpha(0) \quad \text{and} \quad \mathcal{J}_\alpha(w^{(\alpha)}) < \mathcal{J}_\alpha(u_\eta).$$

As a result, it follows that

$$w^{(\alpha)} \notin \{0, u^\eta\}.$$

If we now take $u^c = 0$ and suppose additionally that u^η has zero mean value, then $\mathcal{I}(\alpha) > 0$ for all $\alpha \in (0, \infty)$, while clearly $\overline{\mathcal{I}}(\infty) = 0$, that is, the minimum of $\overline{\mathcal{I}}$ is only attained at the boundary point $\alpha = \infty$. Similarly, for $u^c = u^\eta$, the unique minimizer of $\overline{\mathcal{I}}$ is $\alpha = 0$. △

7.4 Optimal integrability exponents

Here, we study the optimization of an integrability parameter, p, for a fixed nonlocal regularizer. Our motivation comes from the appearance of different L^p-norms in image processing, such as in quadratic, TV, and Lipschitz regularization [176, Section 4]. We focus on the parameter range $\Lambda = [1, \infty)$ with closure $\overline{\Lambda} = [1, \infty]$, paying particular attention to the structural change occurring at $p = \infty$.

Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain and consider a function $f : \Omega \times \Omega \times \mathbb{R} \times \mathbb{R} \to [0, \infty)$ that is Carathéodory, i.e., measurable in the first two and continuous with respect to the last two variables, and that satisfies the following bounds and convexity condition:

($H1_p$) There exist $M, \delta > 0$ and $\beta \in [0,1]$ such that for all $\xi, \zeta \in \mathbb{R}$, we have

$$f(x,y,\xi,\zeta) \le M\left(\frac{|\xi-\zeta|}{|x-y|^\beta} + |\xi| + |\zeta| + 1\right) \quad \text{for a.e. } x,y \in \Omega,$$

and

$$M^{-1}\frac{|\xi-\zeta|}{|x-y|^\beta} - M \le f(x,y,\xi,\zeta) \quad \text{for a.e. } x,y \in \Omega \text{ with } |x-y| < \delta.$$

($H2_p$) f is separately convex in the second two variables, i.e., $f(x,y,\cdot,\zeta)$ and $f(x,y,\xi,\cdot)$ are convex for a.e. $x,y \in \Omega$ and every $\xi, \zeta \in \mathbb{R}^n$.

In this setting, we take $p \in [1,\infty)$ and consider the regularization term $\mathcal{R}_p : L^2(\Omega) \to [0,\infty]$ defined by

$$\mathcal{R}_p(u) := \left(\frac{1}{|\Omega\times\Omega|}\int_\Omega\int_\Omega f^p(x,y,u(x),u(y))\ dx\ dy\right)^{1/p}. \tag{7.19}$$

Remark 7.4.1. a) Since the regularizer $\mathcal{R}_p$ is invariant under symmetrization, one can assume without loss of generality that f is symmetric in the sense that $f(x,y,\xi,\zeta) = f(y,x,\zeta,\xi)$ for all $x,y \in \Omega$ and $\xi,\zeta \in \mathbb{R}$.

b) Let $p, q \in [1,\infty)$ with $p > q$. Hölder's inequality then yields for every $u \in \operatorname{Dom}\mathcal{R}_p = \{u \in L^2(\Omega) : \mathcal{R}_p(u) < \infty\}$ that

$$\left(\int_\Omega\int_\Omega f^p(x,y,u(x),u(y))\ dx\ dy\right)^{1/p} \ge |\Omega\times\Omega|^{\frac{q-p}{pq}}\left(\int_\Omega\int_\Omega f^q(x,y,u(x),u(y))\ dx\ dy\right)^{1/q},$$

which translates into $\mathcal{R}_p(u) \ge \mathcal{R}_q(u)$; in particular, $\operatorname{Dom}\mathcal{R}_p \subset \operatorname{Dom}\mathcal{R}_q$. △

A basic example of a symmetric Carathéodory function f satisfying ($H1_p$) with $\beta = 0$ and ($H2_p$) is

$$f(x,y,\xi,\zeta) = a(x-y)|\xi-\zeta| \quad \text{for } x,y \in \Omega \text{ and } \xi,\zeta \in \mathbb{R},$$

where $a \in L^\infty(\mathbb{R}^n)$ is an even function such that $\operatorname{ess\,inf}_{\mathbb{R}^n} a > 0$. Another example of such a function f with $\beta = 1$ in ($H1_p$) is

$$f(x,y,\xi,\zeta) = b\frac{|\xi-\zeta|}{|x-y|} \quad \text{for } x,y \in \Omega \text{ and } \xi,\zeta \in \mathbb{R},$$

with $b > 0$; note that for the $p > n$ case, the corresponding regularizer $\mathcal{R}_p$ is, up to a multiplicative constant, the Gagliardo semi-norm of the fractional Sobolev space $W^{1-\frac{n}{p},p}(\Omega)$.

Before showing how the framework of Section 7.2 can be applied here, let us first collect and discuss a few properties of the regularizers $\mathcal{R}_p$ with $p \in [1,\infty)$. We introduce the notation

$$[u]_{p,\beta} := \left(\int_\Omega\int_\Omega \frac{|u(x)-u(y)|^p}{|x-y|^{\beta p}}\ \mathrm{d}x\ \mathrm{d}y\right)^{1/p}$$

to indicate a suitable (p,β)-nonlocal seminorm. Our first lemma shows that the boundedness of the regularizer $\mathcal{R}_p$ is equivalent to the simultaneous boundedness of the L^p-norm and of the (p,β)-seminorm.

Lemma 7.4.2. *There exists a constant $C > 0$, depending on n, p, Ω, M, δ, and β, such that*

$$\|u\|_{L^p(\Omega)} \le C\big(\mathcal{R}_p(u) + \|u\|_{L^2(\Omega)} + 1\big), \tag{7.20}$$

$$[u]_{p,\beta} \le C\big(\mathcal{R}_p(u) + \|u\|_{L^p(\Omega)} + 1\big), \tag{7.21}$$

and

$$\mathcal{R}_p(u) \le C\big([u]_{p,\beta} + \|u\|_{L^p(\Omega)} + 1\big) \tag{7.22}$$

for all $u \in L^2(\Omega)$, and for all $p \in [1, \infty)$.

Proof. Properties (7.20) and (7.21) are direct consequences of the coercivity bound on the double-integrand f in (H1$_p$). In fact, for (7.20), we use the nonlocal Poincaré inequality in [33, Proposition 4.2], which also holds for $u \in L^2(\Omega)$ via a truncation argument. From the upper bound in (H1$_p$), we infer (7.22). □

The next result provides a characterization of the domain of $\mathcal{R}_p$.

Lemma 7.4.3. *For any $p \in [1, \infty)$ there holds*

$$Dom\,\mathcal{R}_p = \big\{u \in L^p(\Omega) \cap L^2(\Omega) : [u]_{p,\beta} < \infty\big\}. \tag{7.23}$$

If, additionally, $\beta p < n$, then

$$Dom\,\mathcal{R}_p = L^p(\Omega) \cap L^2(\Omega).$$

If, instead, $\beta p > n$, then

$$Dom\,\mathcal{R}_p = W^{\beta - \frac{n}{p}, p}(\Omega) \cap L^2(\Omega). \tag{7.24}$$

Proof. By combining (7.20) and (7.21) with (7.22), we deduce (7.23). In the case $\beta p < n$, a direct computation shows that $[u]_{p,\beta} < \infty$ for all $u \in L^p(\Omega)$, hence we infer the statement. Property (7.24) follows by observing that for $\beta p > n$, the quantity $[u]_{p,\beta}$ corresponds to the Gagliardo seminorm of the fractional Sobolev space $W^{\beta - \frac{n}{p}, p}(\Omega)$ (cf. e.g. [96]). □

As a consequence of Lemma 7.4.3, we deduce, in particular, that $C_c^\infty(\mathbb{R}^n) \subset \mathrm{Dom}\,\mathcal{R}_p$, where the functions in $C_c^\infty(\mathbb{R}^n)$ are implicitly restricted to Ω.

The next lemma shows that any element of the domain of $\mathcal{R}_p$ can be extended to a function having compact support and finite (p, β)-seminorm.

Lemma 7.4.4. *Let $p \in [1, \infty)$. For any $u \in \mathrm{Dom}\,\mathcal{R}_p$, there is a $\bar{u} \in L^p(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ with compact support inside some bounded open set Ω' with $\Omega \subset \Omega' \subset \mathbb{R}^n$ satisfying $\bar{u} = u$ on Ω and*

$$\int_{\Omega'} \int_{\Omega'} \frac{|\bar{u}(x) - \bar{u}(y)|^p}{|x - y|^{\beta p}} \,\mathrm{d}x \,\mathrm{d}y < \infty. \tag{7.25}$$

Proof. If $\beta > \frac{n}{p}$, this follows directly from well-established extension results for fractional Sobolev spaces on Ω to those on $\mathbb{R}^n$ (cf. [96, Theorem 5.4]), considering (7.24). If $1 \le \beta p \le n$, the map $x \mapsto |x - y|^{-\beta p}$ is no longer integrable at infinity. Property (7.25) follows by minor modifications to the arguments in [96, Section 5]. □

Elements of the domain of $\mathcal{R}_p$ can be approximated by sequences of smooth maps with compact support.

Lemma 7.4.5. *Let $p \in [1,\infty)$. For every $u \in Dom\,\mathcal{R}_p$, there exists a sequence $(u_l)_l \subset C_c^\infty(\mathbb{R}^n)$ such that $u_l \to u$ in $L^p(\Omega)$ and $\lim_{l\to\infty} \mathcal{R}_p(u_l) = \mathcal{R}_p(u)$ as $l \to \infty$.*

Proof. Let $\bar{u}$ be an extension of u as in Lemma 7.4.3. We define $u_l = \varphi_{1/l} * \bar{u} \in C_c^\infty(\mathbb{R}^n)$ for $l \in \mathbb{N}$ with $(\varphi_\varepsilon)_{\varepsilon>0}$ a family of smooth standard mollifiers satisfying $0 \leq \varphi_\varepsilon \leq 1$ and $\int_{\mathbb{R}^n} \varphi_\varepsilon \, dx = 1$, and whose support lies in the ball centered at the origin and with radius $\varepsilon > 0$, $\operatorname{supp} \varphi_\varepsilon \subset B_\varepsilon(0) \subset \mathbb{R}^n$. Then, $u_l \to u$ in $L^p(\Omega)$ and $u_l \to u$ pointwise a.e. in Ω as $l \to \infty$. To show that Lebesgue's dominated convergence theorem can be applied, we use the upper bound in (H1$_p$) to derive the following estimate for any $l \in \mathbb{N}$:

$$f^p(x,y,u_l(x),u_l(y)) \leq 4^{p-1} M^p \left(\frac{|u_l(x)-u_l(y)|^p}{|x-y|^{\beta p}} + |u_l(x)|^p + |u_l(y)|^p + 1 \right) \tag{7.26}$$

for a.e. $(x,y) \in \Omega \times \Omega$. By Jensen's inequality and Fubini's theorem,

$$\begin{aligned} [u_l]_{p,\beta}^p &\leq \int_{B_{1/l}(0)} \varphi_{1/l}(z) \int_\Omega \int_\Omega \frac{|\bar{u}(x-z)-\bar{u}(y-z)|^p}{|x-y|^{\beta p}} \, dx \, dy \, dz \\ &\leq \int_{\Omega_{1/l}} \int_{\Omega_{1/l}} \frac{|\bar{u}(x)-\bar{u}(y)|^p}{|x-y|^{\beta p}} \, dx \, dy < \infty, \end{aligned}$$

with $\Omega_{1/l} = \{x \in \mathbb{R}^n : d(x,\Omega) < 1/l\}$; thus, $\limsup_{l\to\infty} [u_l]_{p,\beta}^p \leq [u]_{p,\beta}^p$. Conversely, the a.e. pointwise convergence of the mollified sequence gives $\liminf_{l\to\infty} [u_l]_{p,\beta}^p \geq [u]_{p,\beta}^p$ by Fatou's lemma. Along with the L^p-convergence of $(u_l)_l$, the upper bound in (7.26) is thus a converging sequence in $L^1(\Omega \times \Omega)$. This concludes the proof of the lemma. □

Finally, we characterize the weak lower-semicontinuity of the regularizers. We refer to [34, 107, 173] for a discussion on sufficient (and necessary) conditions for the weak lower semicontinuity of inhomogeneous double-integral functionals.

Lemma 7.4.6. *For every $p \in [1,\infty)$, the regularizer $\mathcal{R}_p$ is L^2-weak lower semicontinuous.*

Proof. The statement is an immediate consequence of the nonnegativity of f and (H2$_p$), see e.g. [174, Theorem 2.5] or [166]. □

Remark 7.4.7. Observe that Lemmas 7.4.3 and 7.4.6 imply in particular that the hypothesis (H) from Section 7.2 is fulfilled. △

Given a collection of noisy images $u^\eta \in L^2(\Omega;\mathbb{R}^N)$ and $p \in [1,\infty)$, we set, for each $j \in \{1,\cdots,N\}$,

$$\mathcal{J}_{p,j}(u) := \|u - u_j^\eta\|_{L^2(\Omega)}^2 + \mathcal{R}_p(u) \qquad \text{for } u \in L^2(\Omega),$$

with $K_{p,j} := \arg\min \mathcal{J}_{p,j} \neq \emptyset$ since (H) is satisfied. As in ($\mathcal{T}$), we define $\mathcal{I} : [1,\infty) \to [0,\infty)$ by

$$\mathcal{I}(p) = \inf_{w \in K_p} \|w - u^c\|_{L^2(\Omega;\mathbb{R}^N)}^2 \quad \text{for } p \in [1,\infty),$$

where $K_p = K_{p,1} \times K_{p,2} \times \cdots \times K_{p,N}$. Next, we prove the Mosco-convergence result that will provide us with an extension of $\mathcal{I}$ to $\overline{\Lambda} = [1,\infty]$. It is an L^p-approximation statement in the present nonlocal setting, which can be obtained from a modification of the arguments by Champion, De Pascale, & Prinari [61] in the local case, and those by Kreisbeck, Ritorto, & Zappale [139, Theorem 1.3], where the case of homogeneous double-integrands is studied.

Proposition 7.4.8 (Mosco-convergence of the regularizers). *Let* $\Lambda = [1,\infty)$*,* $\mathcal{R}_p$ *for* $p \in [1,\infty)$ *as in* (7.19)*, and suppose that* (H1$_p$) *and* (H2$_p$) *are satisfied. Then, for* $p \in \overline{\Lambda} = [1,\infty]$*,*

$$\overline{\mathcal{R}}_p := \mathrm{Mosc}(L^2)\text{-}\lim_{p'\to p} \mathcal{R}_{p'} = \begin{cases} \mathcal{R}_p & \text{if } p \in [1,\infty), \\ \mathcal{R}_\infty & \text{if } p = \infty, \end{cases} \tag{7.27}$$

with $\mathcal{R}_\infty : L^2(\Omega) \to [0,\infty]$ *given by*

$$\mathcal{R}_\infty(u) := \operatorname{ess\,sup}_{(x,y)\in\Omega\times\Omega} f(x,y,u(x),u(y)).$$

Proof. To show (7.27), it suffices to show that for every sequence $(p_k)_k \subset [1,\infty)$ converging to $p \in [1,\infty]$, (7.27) holds with p' replaced by p_k. We divide the proof into two cases.

Case 1: $p \in [1,\infty)$. For the recovery sequence, consider $u \in \operatorname{Dom}\mathcal{R}_p$ and take $(u_l)_l \subset C_c^\infty(\mathbb{R})$ as in Lemma 7.4.5, satisfying $u_l \to u$ in $L^p(\Omega)$ and $\mathcal{R}_p(u_l) \to \mathcal{R}_p(u)$ as $l \to \infty$. In view of Lemma 7.4.3, we know that $(u_l)_l$ is contained in $\operatorname{Dom}\mathcal{R}_p$ and $\operatorname{Dom}\mathcal{R}_{p_k}$ for all $k \in \mathbb{N}$, and we conclude via Lebesgue's dominated convergence theorem that

$$\lim_{k\to\infty} \mathcal{R}_{p_k}(u_l) = \mathcal{R}_p(u_l)$$

for every $l \in \mathbb{N}$. Hence,

$$\lim_{l\to\infty}\lim_{k\to\infty} \mathcal{R}_{p_k}(u_l) = \lim_{l\to\infty} \mathcal{R}_p(u_l) = \mathcal{R}_p(u),$$

so that one can find a recovery sequence by extracting an appropriate diagonal sequence.

To prove the lower bound, let $u_k \rightharpoonup u$ in $L^2(\Omega)$ be such that $\lim_{k\to\infty}\mathcal{R}_{p_k}(u_k) = \liminf_{k\to\infty}\mathcal{R}_{p_k}(u_k) < \infty$, and fix $s \in (1,p)$ (or $s = 1$ if $p = 1$). Observe that $p_k \geq s$ for all k sufficiently large because $p_k \to p$ for $k \to \infty$. Then, Remark 7.4.1 b) and the weak lower semicontinuity of $\mathcal{R}_s$ according to Lemma 7.4.6 imply that

$$\lim_{k\to\infty} \mathcal{R}_{p_k}(u_k) \geq \liminf_{k\to\infty} \mathcal{R}_s(u_k) \geq \mathcal{R}_s(u).$$

If $s = p = 1$ the argument is complete, whereas in the case $p > 1$, an additional application of Fatou's lemma shows $\liminf_{s\nearrow p}\mathcal{R}_s(u) \geq \mathcal{R}_p(u)$, giving rise to the desired liminf inequality.

Case 2: $p = \infty$. That constant sequences serve as recovery sequences results from the observation that $\mathcal{R}_{p_k}(u) \to \mathcal{R}_\infty(u)$ as $k \to \infty$ for all $u \in \operatorname{Dom}\mathcal{R}_\infty$. The latter is an immediate consequence of classical L^p-approximation, i.e., the well-known fact that $\lim_{p\to\infty}\|v\|_{L^p(V)} = \|v\|_{L^\infty(V)} = \operatorname{ess\,sup}_{x\in V}|v(x)|$ for all $v \in L^\infty(V)$ with $V \subset \mathbb{R}^m$ open and bounded.

To prove the lower bound, we argue via Young measure theory (see, e.g., [112,173] for a general introduction). Let $u_k \rightharpoonup u$ in $L^2(\Omega)$, and denote by $\nu = \{\nu_x\}_{x\in\Omega}$ the Young measure generated by a (non-relabeled) subsequence of $(u_k)_k$. The barycenter of $[\nu_x] := \int_{\mathbb{R}} \xi \,\mathrm{d}\nu_x(\xi)$ then coincides with $u(x)$ for a.e. $x \in \Omega$. Without loss of generality, one can suppose that $\infty > \liminf_{k\to\infty}\mathcal{R}_{p_k}(u_k) = \lim_{k\to\infty}\mathcal{R}_{p_k}(u_k)$. Recalling Remark 7.4.1 b), we have that

$$\lim_{k\to\infty} \mathcal{R}_{p_k}(u_k) \geq \liminf_{q\to\infty}\liminf_{k\to\infty} \mathcal{R}_q(u_k). \tag{7.28}$$

On the other hand, with the nonlocal field v_u associated with some $u : \Omega \to \mathbb{R}$ defined by

$$v_u(x,y) := (u(x),u(y)) \qquad \text{for } (x,y) \in \Omega\times\Omega,$$

the statement of [173, Proposition 2.3] allows us to extract a subsequence $(v_{u_k})_k$ that generates the Young measure $\{\nu_x \otimes \nu_y\}_{(x,y)\in\Omega\times\Omega}$. Hence, a standard result on Young measure lower semicontinuity (see e.g. [112, Section 8.1]) yields

$$\liminf_{k\to\infty} \mathcal{R}_q(u_k) \geq \Big(\frac{1}{|\Omega\times\Omega|}\int_\Omega\int_\Omega\int_{\mathbb{R}}\int_{\mathbb{R}} f^q(x,y,\xi,\zeta)\ \mathrm{d}\nu_x(\xi)\ \mathrm{d}\nu_y(\zeta)\ \mathrm{d}x\ \mathrm{d}y\Big)^{1/q}.$$

Letting $q \to \infty$, we use classical L^p-approximation results and the Jensen's type inequality for separately convex functions in [142, Lemma 3.5 (iv)] to conclude that

$$\begin{aligned}\liminf_{q\to\infty}\liminf_{k\to\infty} \mathcal{R}_q(u_k) &\geq \operatorname{ess\,sup}_{(x,y)\in\Omega\times\Omega}\big[(\nu_x\otimes\nu_y)\text{-}\operatorname{ess\,sup}_{(\xi,\zeta)\in\mathbb{R}\times\mathbb{R}} f(x,y,\xi,\zeta)\big]\\ &\geq \operatorname{ess\,sup}_{(x,y)\in\Omega\times\Omega} f(x,y,[\nu_x],[\nu_y])\\ &= \operatorname{ess\,sup}_{(x,y)\in\Omega\times\Omega} f(x,y,u(x),u(y)) = \mathcal{R}_\infty(u);\end{aligned}$$

note that $(\nu_x\otimes\nu_y)$-ess $\sup_{(\xi,\zeta)\in\mathbb{R}\times\mathbb{R}} f(x,y,\xi,\zeta) = \inf\{c\in\mathbb{R} : f(x,y,\cdot,\cdot)\leq c\ (\nu_x\otimes\nu_y)\text{-a.e. in } \mathbb{R}\times\mathbb{R}\}$. Finally, the lower bound follows from the previous estimate and (7.28). □

The above result implies that the reconstruction functional for $p = \infty$ and $j \in \{1,\cdots,N\}$ is given by

$$\overline{\mathcal{J}}_{\infty,j}(u) := \|u - u_j^\eta\|^2_{L^2(\Omega)} + \mathcal{R}_\infty(u) \qquad \text{for } u \in L^2(\Omega).$$

Under the additional convexity condition on the given function $f : \Omega\times\Omega\times\mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$ that

(H3$_p$) f is (jointly) level convex in its last two variables,

where level convexity means convexity of the sub-level sets of the function, the supremal functional $\mathcal{R}_\infty$ also becomes level convex. In combination with the strict convexity of the fidelity term, the reconstruction functional $\overline{\mathcal{J}}_{\infty,j}$ then admits a unique minimizer. Since level convexity is weaker than convexity, we do not necessarily have that $\mathcal{J}_{p,j}$ for $p \in [1,\infty)$ is (level) convex, and it may have multiple minimizers.

If we suppose that f fulfills (H1$_p$)–(H3$_p$), then Theorem 7.2.5 and Proposition 7.4.8 imply that the extension $\overline{\mathcal{I}} : [1,\infty] \to [0,\infty]$ is given by

$$\overline{\mathcal{I}}(p) = \begin{cases}\mathcal{I}(p) & \text{for } p\in[1,\infty),\\ \|w^{(\infty)} - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} & \text{for } p = \infty,\end{cases}$$

for $p \in [1,\infty]$, where $w^{(\infty)}$ denotes the unique componentwise minimizer of $\overline{\mathcal{J}}_\infty$. In particular, the hypothesis (ii) of Theorem 7.2.5 is satisfied, which shows that $\overline{\mathcal{I}}$ is the relaxation of $\mathcal{I}$ and, thus, admits a minimizer $\bar{p} \in \overline{\Lambda} = [1,\infty]$.

We conclude this section with a discussion of examples when optimal values of the integrability exponents are obtained in the interior of the original interval Λ or at its boundary, respectively. In one case, the presence of noise causes $\mathcal{R}_\infty$ to penalize u^c more than u^η, while $\mathcal{R}_q$ for some $q \in [1,\infty)$ prefers the clean image. This entails that the optimal parameter is attained in $\Lambda = [1,\infty)$. In the second case instead, the reconstruction functional for $p = \infty$ gives back the exact clean image and outperforms the reconstruction functionals for other parameter values.

Example 7.4.9. a) Let $f = \alpha\widehat{f} : \Omega\times\Omega\times\mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$, for some $\alpha > 0$ to be specified later, be a double-integrand satisfying (H1$_p$), (jointly) convex in the last two variables, and vanishing exactly on $\{(x,y,\xi,\xi) : x,y\in\Omega,\ \xi\in\mathbb{R}\}$. Following (7.19), we set

$$\mathcal{R}_p(u) = \alpha\bigg(\frac{1}{|\Omega\times\Omega|}\int_\Omega\int_\Omega \widehat{f}^p(x,y,u(x),u(y))\ \mathrm{d}x\ \mathrm{d}y\bigg)^{1/p} =: \alpha\widehat{\mathcal{R}}_p(u)$$

for $u \in L^2(\Omega)$ and $p \in [1, \infty)$.

We further introduce the following two conditions on the given data $u^\eta, u^c \in L^2(\Omega; \mathbb{R}^N)$:

(H4$_p$) $\sum_{j=1}^N \mathcal{R}_q(u_j^c) < \sum_{j=1}^N \mathcal{R}_q(u_j^\eta)$ for some $q \in [1, \infty)$;

(H5$_p$) $\sum_{j=1}^N \mathcal{R}_\infty(2u_j^\eta - u_j^c) < \sum_{j=1}^N \mathcal{R}_\infty(u_j^\eta)$.

By applying Lemma 7.3.4 (i) from the previous section with $\mathcal{R} = \widehat{\mathcal{R}}_q$ — the conditions (H1$_\alpha$), (H2$_\alpha$), and (H3$_\alpha$) are immediate to verify in view of Lemma 7.4.3, Lemma 7.4.6, and (H4$_p$) — we can then deduce for small enough α that $\overline{\mathcal{I}}(q) < \|u^\eta - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)}$. On the other hand, due to (H5$_p$), the same lemma can be applied to $\mathcal{R} = \widehat{\mathcal{R}}_\infty$ with $\widehat{\mathcal{R}}_\infty(u) = \operatorname{ess\,sup}_{(x,y)\in\Omega\times\Omega} \widehat{f}(x, y, u(x), u(y))$ for $u \in L^2(\Omega)$ to find

$$\|w^{(\infty)} - (2u^\eta - u^c)\|^2_{L^2(\Omega;\mathbb{R}^N)} < \|u^\eta - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)}, \tag{7.29}$$

provided α is sufficiently small. The reverse triangle inequality then yields

$$\begin{aligned}\overline{\mathcal{I}}(\infty) &\geq \left(\|w^{(\infty)} - (2u^\eta - u^c)\|^2_{L^2(\Omega;\mathbb{R}^N)} - 2\|u^\eta - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)}\right)^2 \\ &> \|w^{(\infty)} - (2u^\eta - u^c)\|^2_{L^2(\Omega;\mathbb{R}^N)} > \|u^\eta - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} > \overline{\mathcal{I}}(q),\end{aligned}$$

where in the second and third inequality we have used (7.29). This proves that the optimal parameter is attained inside $[1, \infty)$, and, therefore, is also a minimizer of $\mathcal{I}$.

b) We illustrate a) with a specific example. Consider $\Omega = (0, 1)$ and let $\widehat{f}(x, y, \xi, \zeta) = |\xi - \zeta|/|x - y|$ for $x, y \in \Omega$ and $\xi, \zeta \in \mathbb{R}^n$. This leads then to the difference quotient regularizers

$$\mathcal{R}_p(u) = \alpha \left(\int_0^1 \int_0^1 \frac{|u(x) - u(y)|^p}{|x - y|^p} \, dx \, dy \right)^{1/p} =: \alpha \widehat{\mathcal{R}}_p(u) \tag{7.30}$$

and

$$\mathcal{R}_\infty(u) = \alpha \operatorname{ess\,sup}_{(x,y)\in(0,1)^2} \frac{|u(x) - u(y)|}{|x - y|} = \alpha \mathrm{Lip}(u), \tag{7.31}$$

with $\mathrm{Lip}(u)$ denoting the Lipschitz constant of (a representative of) u, which could be infinite.

With the sawtooth function $v : [0, 1] \to \mathbb{R}$ defined by

$$v(x) = \begin{cases} x & \text{for } 0 \leq x \leq 1/4, \\ -x + 1/2 & \text{for } 1/4 < x \leq 3/4, \\ x - 1 & \text{for } 3/4 < x \leq 1, \end{cases}$$

we take a single clean and noisy image given by

$$u^c(x) = \begin{cases} 0 & \text{for } 0 < x \leq 1/3, \\ 10v(3x - 1) & \text{for } 1/3 < x \leq 2/3 \\ 0 & \text{for } 2/3 < x < 1. \end{cases} \quad \text{and} \quad u^\eta(x) = \begin{cases} v(3x) & \text{for } 0 < x \leq 1/3, \\ (10 - \varepsilon)v(3x - 1) & \text{for } 1/3 < x \leq 2/3, \\ v(3x - 2) & \text{for } 2/3 < x < 1, \end{cases}$$

respectively, where $\varepsilon > 0$ is small; see Figure 7.1. We observe that u^c is constant near the boundaries and only slightly steeper than u^η in the middle of the domain. Numerical calculations show that for small ε, such as $\varepsilon = 0.1$, the estimate $\mathcal{R}_2(u^c) < \mathcal{R}_2(u^\eta)$, and hence (H4$_p$) with $q = 2$, holds; moreover, (H5$_p$) holds since the clean image has a higher Lipschitz constant than the noisy image in the sense that

$$\mathrm{Lip}(2u^\eta - u^c) = 30 - 6\varepsilon < 30 - 3\varepsilon = \mathrm{Lip}(u^\eta).$$

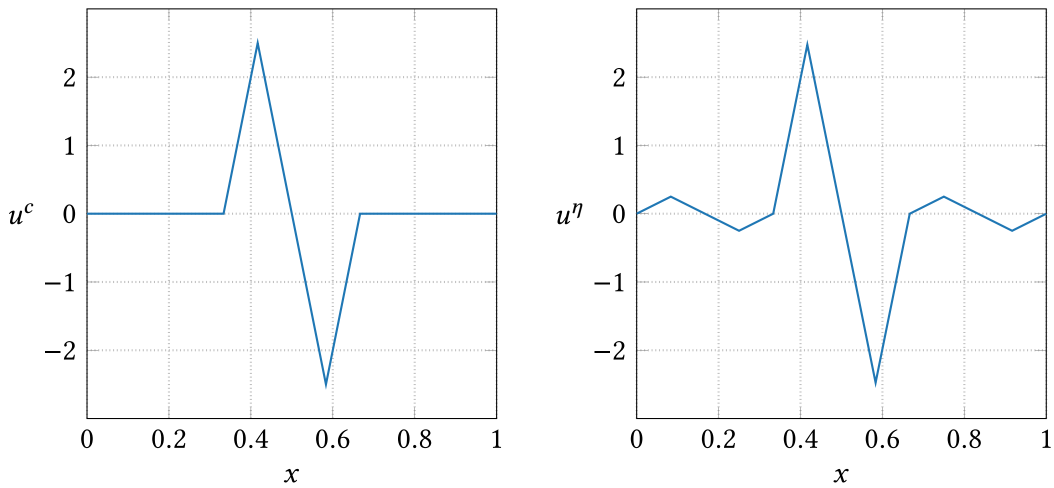

Figure 7.1: The graphs of the functions u^c and u^η from Example 7.4.9 a) with $\varepsilon = 0.1$.

Therefore, we find that for $\alpha > 0$ small enough, the optimal parameter lies inside $\Lambda = [1, \infty)$.

c) If we work with the same regularizers as in b), there are reasonable images for which the Lipschitz regularizer in (7.31) performs better than the other regularizers in (7.30). Let us consider with $\alpha > 0$ chosen as in b), the images

$$u^c(x) = x - 1/2 \quad \text{and} \quad u^\eta = (1 + 6\alpha)u^c.$$

Since u^η is affine, we can show that the reconstruction with the Lipschitz regularizer is also an affine function. Indeed, for every other function, one can find an affine function with at most the same Lipschitz constant without increasing the distance to u^η anywhere. This, in combination with the fact that the images are odd functions with respect to $x = 1/2$, shows that $w^{(\infty)}$ is of the form $w^{(\infty)}(x) = \gamma(x - 1/2) = \gamma u^c$ with $\gamma \geq 0$. Due to the optimality of $w^{(\infty)}$, the constant γ has to minimize the quantity

$$\|\gamma u^c - u^\eta\|^2_{L^2((0,1))} + \alpha \mathrm{Lip}(\gamma u^c) = \frac{1}{12}(\gamma - (1 + 6\alpha))^2 + \alpha\gamma,$$

which yields $\gamma = 1$. Hence, $w^{(\infty)}$ coincides with the clean image and therefore $\overline{\mathcal{I}}(\infty) = 0$, which implies that $p = \infty$ is the optimal parameter in this case.

7.5 Varying the amount of nonlocality

Next, we study two classes of nonlocal regularizers, $\mathcal{R}_\delta$ with $\delta \in \Lambda := (0, \infty)$, considered by Brezis & Nguyen [53] and Aubert & Kornprobst [21], respectively, in the context of image processing. In both cases, we aim at optimizing the parameter δ that encodes the amount of nonlocality in the problem. We mention further that both families of functionals recover the classical TV-reconstruction model in the limit $\delta \to 0$, cf. [21, 53].

To set the stage for our analysis, consider training data $(u^c, u^\eta) \in L^2(\Omega; \mathbb{R}^N) \times L^2(\Omega; \mathbb{R}^N)$ and the reconstruction functionals $\mathcal{J}_{\delta,j} : L^2(\Omega) \to [0, \infty]$ with $\delta \in \Lambda$ and $j \in \{1, 2, \dots, N\}$ given by

$$\mathcal{J}_{\delta,j}(u) = \|u - u^\eta_j\|^2_{L^2(\Omega)} + \mathcal{R}_\delta(u).$$

After showing that the sets

$$K_{\delta,j} = \operatorname*{arg\,min}_{u \in L^2(\Omega)} \mathcal{J}_{\delta,j}(u). \tag{7.32}$$

are non-empty for each of the two choices of the regularizers $\mathcal{R}_\delta$, the upper-level functional from $(\mathcal{T})$ in Section 7.2 becomes

$$\mathcal{I} : (0, \infty) \to [0, \infty), \quad \mathcal{I}(\delta) = \inf_{w \in K_\delta} \|w - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} \tag{7.33}$$

with $K_\delta = K_{\delta,1} \times K_{\delta,2} \times \cdots \times K_{\delta,N}$. In order to find its extension $\overline{\mathcal{I}}$ defined on $\overline{\Lambda} = [0, \infty]$, we determine the Mosco-limits of the regularizers (cf. (7.6) and Theorem 7.2.5). This is the content of Propositions 7.5.3 and 7.5.5 below, which provide the main results of this section.

7.5.1 Brezis & Nguyen setting

For every $\delta \in (0, \infty)$ and $u \in L^1(\Omega)$, we consider the regularizers

$$\mathcal{R}_\delta(u) := \delta \int_\Omega \int_\Omega \frac{\varphi(|u(x) - u(y)|/\delta)}{|x - y|^{n+1}} \, dx \, dy,$$

where, following [53], the function $\varphi : [0, \infty) \to [0, \infty)$ is assumed to satisfy the following hypotheses:

(H1$_\delta$) φ is lower semicontinuous in $[0, \infty)$ and continuous in $[0, \infty)$ except at a finite number of points, where it admits left- and right-side limits;

(H2$_\delta$) there exists a constant $a > 0$ such that $\varphi(t) \le \min\{at^2, a\}$ for all $t \in [0, \infty)$;

(H3$_\delta$) φ is non-decreasing;

(H4$_\delta$) it holds that $\gamma_n \int_0^\infty \varphi(t) t^{-2} \, dt = 1$ with $\gamma_n := \int_{\mathbb{S}^{n-1}} |e \cdot \sigma| \, d\sigma$ for any $e \in \mathbb{S}^{n-1}$.

Note that the assumptions on φ imply that the functional $\mathcal{R}_\delta$ is never convex.

Example 7.5.1. Examples of functions φ with the properties (H1$_\delta$)–(H4$_\delta$) include suitable normalizations of

$$t \mapsto \begin{cases} 0 & \text{if } t \le 1 \\ 1 & \text{if } t > 1 \end{cases}, \qquad t \mapsto \begin{cases} t^2 & \text{if } t \le 1 \\ 1 & \text{if } t > 1 \end{cases}, \qquad t \mapsto 1 - e^{-t^2}$$

for $t \ge 0$, cf. [53].

To guarantee that the functionals $\mathcal{R}_\delta$ satisfy a suitable compactness property, see Theorem 7.5.2 b), we must additionally assume that

(H5$_\delta$) $\varphi(t) > 0$ for all $t > 0$.

Clearly, the last two functions from Example 7.5.1 satisfy the positivity condition, while the first one does not. In identifying the Mosco-limits $\overline{\mathcal{R}}_\delta$ in each of the three cases $\delta \in (0, \infty)$, $\delta = 0$, and $\delta = \infty$, we make repeated use of [53, Theorems 1, 2 and 3], which we recall here for the reader's convenience.

Theorem 7.5.2 (cf. [53, Theorems 1–3]). *Let $\Omega \subset \mathbb{R}^n$ be a bounded and smooth domain, and let φ satisfy* (H1$_\delta$)–(H4$_\delta$).
a) If $(\delta_k)_k \subset (0, \infty)$ is such that $\delta_k \to 0$, then the following statements hold:

(*i*) *There exists a constant* $K(\varphi) \in (0,1]$, *independent of* Ω, *such that* $(\mathcal{R}_{\delta_k})_k$ Γ*-converges as* $k \to \infty$, *with respect to the* $L^1(\Omega)$*-topology, to* $\mathcal{R}_0 : L^1(\Omega) \to [0,\infty]$ *defined for* $u \in L^1(\Omega)$ *by*

$$\mathcal{R}_0(u) := \begin{cases} K(\varphi)|Du|(\Omega) & \text{if } u \in BV(\Omega), \\ \infty & \text{if } u \in L^1(\Omega) \setminus BV(\Omega). \end{cases}$$

(*ii*) *If* $(u_k)_k$ *is a bounded sequence in* $L^1(\Omega)$ *with* $\sup_k \mathcal{R}_{\delta_k}(u_k) < \infty$, *then there exist a subsequence* $(u_{k_l})_l$ *of* $(u_k)_k$ *and a function* $u \in L^1(\Omega)$ *such that* $\lim_{l\to\infty} \|u_{k_l} - u\|_{L^1(\Omega)} = 0$.

b) Suppose that (H5$_\delta$) *holds in addition to the above conditions, and let* $(u_k)_k$ *be a bounded sequence in* $L^1(\Omega)$ *with* $\sup_k \mathcal{R}_\delta(u_k) < \infty$ *for some* $\delta > 0$. *Then, there exists a subsequence* $(u_{k_l})_l$ *of* $(u_k)_k$ *and a function* $u \in L^1(\Omega)$ *such that* $\lim_{l\to\infty} \|u_{k_l} - u\|_{L^1(\Omega)} = 0$.

We point out that if φ fulfills (H1$_\delta$)–(H5$_\delta$), then (H) in Section 7.2 holds and the sets $K_{\delta,j}$ defined in (7.32) are non-empty (cf. [53, Corollary 7]). We are now in a position to characterize the asymptotic behavior of the regularizers $\mathcal{R}_{\delta'}$ as $\delta' \to \delta \in \overline{\Lambda} = [0,\infty]$.

Proposition 7.5.3 (Mosco-convergence of regularizers). *Let* $\Lambda = (0,\infty)$ *and* $\Omega \subset \mathbb{R}^n$ *be a bounded and smooth domain. Under the assumptions* (H1$_\delta$)–(H5$_\delta$) *on* $\varphi : [0,\infty) \to [0,\infty)$, *it holds that*

$$\overline{\mathcal{R}}_\delta := \text{Mosc}(L^2)\text{-}\lim_{\delta' \to \delta} \mathcal{R}_{\delta'} = \begin{cases} \mathcal{R}_\delta & \text{if } \delta \in (0,\infty), \\ \mathcal{R}_0 & \text{if } \delta = 0, \\ 0 & \text{if } \delta = \infty, \end{cases} \qquad \text{for } \delta \in \overline{\Lambda} = [0,\infty]. \tag{7.34}$$

Proof. Considering a sequence $(\delta_k)_k \subset (0,\infty)$ with limit $\delta \in [0,\infty]$, one needs to verify that the Mosco-limit of $(\mathcal{R}_{\delta_k})_k$ exist and is given by the right-hand side of (7.34). We split the proof into three cases.

Case 1: $\delta = 0$. Let $(u_k)_k \subset L^2(\Omega)$ and $u \in L^2(\Omega)$ be such that $u_k \rightharpoonup u$ in $L^2(\Omega)$. We aim to show that

$$\mathcal{R}_0(u) \le \liminf_{k\to\infty} \mathcal{R}_{\delta_k}(u). \tag{7.35}$$

One may thus assume without loss of generality that the limit inferior on the right-hand side of (7.35) is finite, and, after extracting a subsequence if necessary, also

$$\sup_k \mathcal{R}_{\delta_k}(u_k) < \infty.$$

Hence, by Theorem 7.5.2 a) (*ii*), it follows that $u_k \to u$ in $L^1(\Omega)$, which together with Theorem 7.5.2 a) (*i*) yields (7.35).

To complement this lower bound, we need to obtain for each $u \in L^2(\Omega) \cap BV(\Omega)$ a sequence $(u_k)_k \subset L^2(\Omega)$ such that $u_k \to u$ in $L^2(\Omega)$ and

$$\mathcal{R}_0(u) \ge \limsup_{k\to\infty} \mathcal{R}_{\delta_k}(u_k). \tag{7.36}$$

The idea is to suitably truncate a recovery sequence of the Γ-limit $\Gamma(L^1)\text{-}\lim_{k\to\infty} \mathcal{R}_{\delta_k}$ from Theorem 7.5.2 (*i*). For the details, fix $l \in \mathbb{N}$ and consider the truncation function, $T^l : \mathbb{R} \to \mathbb{R}$,

$$T^l(t) := \begin{cases} l & \text{if } t \ge l, \\ t & \text{if } -l \le t \le l, \\ -l & \text{if } t \le -l. \end{cases}$$

By Theorem 7.5.2 (i), there exists a sequence $(v_k)_k \subset L^1(\Omega)$ such that $v_k \to u$ in $L^1(\Omega)$ and

$$\lim_{k\to\infty} \mathcal{R}_{\delta_k}(v_k) = K(\varphi)|Du|(\Omega) = \mathcal{R}_0(u). \tag{7.37}$$

Choosing a sequence $(l_k)_k \subset \mathbb{R}$ such that $l_k \to \infty$ and $l_k\|v_k - u\|_{L^1(\Omega)} \to 0$ as $k \to \infty$, we define

$$u_k := T^{l_k} \circ v_k \in L^\infty(\Omega) \quad \text{for all } k \in \mathbb{N}.$$

Then, an application of Hölder's inequality shows that

$$\begin{aligned}\|u_k - u\|_{L^2(\Omega)} &\le \|u_k - T^{l_k} \circ u\|_{L^2(\Omega)} + \|T^{l_k} \circ u - u\|_{L^2(\Omega)}\\ &\le \left(2l_k\|v_k - u\|_{L^1(\Omega)}\right)^{1/2} + \|T^{l_k} \circ u - u\|_{L^2(\Omega)} \to 0,\end{aligned}$$

as $k \to \infty$. Therefore, $u_k \to u$ in $L^2(\Omega)$ and, in view of the monotonicity of φ in (H3$_\delta$), we conclude that

$$\begin{aligned}\limsup_{k\to\infty} \mathcal{R}_{\delta_k}(u_k) &= \limsup_{k\to\infty} \delta_k \int_\Omega\int_\Omega \frac{\varphi(\delta_k^{-1}|T^{l_k}(v_k(x)) - T^{l_k}(v_k(y))|)}{|x-y|^{n+1}}\,dx\,dy\\ &\le \lim_{k\to\infty} \delta_k \int_\Omega\int_\Omega \frac{\varphi(\delta_k^{-1}|v_k(x) - v_k(y)|)}{|x-y|^{n+1}}\,dx\,dy = \lim_{k\to\infty} \mathcal{R}_{\delta_k}(v_k),\end{aligned}$$

which implies (7.36) by (7.37).

Case 2: $\delta \in (0,\infty)$. Consider a sequence $(u_k)_k \subset L^2(\Omega)$ and $u \in L^2(\Omega)$ such that $u_k \rightharpoonup u$ in $L^2(\Omega)$ and

$$\sup_k \mathcal{R}_{\delta_k}(u_k) < \infty.$$

We start by observing that there exist $\bar\delta > 0$ and $K \in \mathbb{N}$ such that for all $k \ge K$, we have $\bar\delta/2 \le \delta_k \le \bar\delta$. Hence, the previous estimate and (H3$_\delta$) yield

$$\sup_{k\ge K} \mathcal{R}_{\bar\delta}(u_k) = \sup_{k\ge K}\left(\bar\delta \int_\Omega\int_\Omega \frac{\varphi(\bar\delta^{-1}|u_k(x) - u_k(y)|)}{|x-y|^{n+1}}\,dx\,dy\right) \le 2\sup_k \mathcal{R}_{\delta_k}(u_k) < \infty.$$

Consequently, in view of Theorem 7.5.2 b), we may further assume that

$$u_k \to u \text{ in } L^1(\Omega) \qquad \text{and} \qquad u_k(x) \to u(x) \text{ for a.e. } x \in \Omega. \tag{7.38}$$

Using Fatou's lemma first, and then (7.38) together with the lower semicontinuity of φ on $[0,\infty)$, we get

$$\begin{aligned}\liminf_{k\to\infty} \mathcal{R}_{\delta_k}(u_k) &= \liminf_{k\to\infty} \delta_k \int_\Omega\int_\Omega \frac{\varphi(\delta_k^{-1}|u_k(x) - u_k(y)|)}{|x-y|^{n+1}}\,dx\,dy\\ &\ge \delta \int_\Omega\int_\Omega \liminf_{k\to\infty} \frac{\varphi(\delta_k^{-1}|u_k(x) - u_k(y)|)}{|x-y|^{n+1}}\,dx\,dy\\ &\ge \delta \int_\Omega\int_\Omega \frac{\varphi(\delta^{-1}|u(x) - u(y)|)}{|x-y|^{n+1}}\,dx\,dy = \mathcal{R}_\delta(u),\end{aligned}$$

which proves the liminf inequality.

For the recovery sequence, fix $u \in L^2(\Omega)$ and take $u_k = \frac{\delta_k}{\delta}u$ for $k \in \mathbb{N}$. Then, $u_k \to u$ in $L^2(\Omega)$ as $k \to \infty$ and

$$\lim_{k\to\infty} \mathcal{R}_{\delta_k}(u_k) = \lim_{k\to\infty} \frac{\delta_k}{\delta}\mathcal{R}_\delta(u) = \mathcal{R}_\delta(u),$$

as desired.

Case 3: $\delta = \infty$. The lower bound follows immediately by the non-negativity of $\mathcal{R}_{\delta_k}$ for $k \in \mathbb{N}$. As a recovery sequence for $u \in L^2(\Omega)$, take a sequence $(u_k)_k \subset L^2(\Omega)$ such that $u_k \to u$ in $L^2(\Omega)$ and $\mathrm{Lip}(u_k) \leq \delta_k^{1/4}$, which is possible since $\delta_k \to \infty$ as $k \to \infty$. Then, using (H2$_\delta$),

$$\begin{aligned}\mathcal{R}_{\delta_k}(u_k) &= \delta_k \int_\Omega \int_\Omega \frac{\varphi(\delta_k^{-1}|u_k(x) - u_k(y)|)}{|x-y|^{n+1}}\, dx\, dy \\ &\leq a\frac{\mathrm{Lip}(u_k)^2}{\delta_k} \int_\Omega \int_\Omega \frac{1}{|x-y|^{n-1}}\, dx\, dy \leq a\delta_k^{-1/2} \int_\Omega \int_\Omega \frac{1}{|x-y|^{n-1}}\, dx\, dy.\end{aligned}$$

Hence, $\mathcal{R}_{\delta_k}(u_k) \to 0$ as $k \to \infty$, which concludes the proof. □

7.5.2 Aubert & Kornprobst setting

Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain. We fix a nonnegative function $\rho : [0,\infty) \to [0,\infty)$ satisfying

(H6$_\delta$) ρ is non-increasing and $\displaystyle\int_{\mathbb{R}^n} \rho(|x|)\ \mathrm{d}x = 1$,

and consider the regularizers given for $\delta \in \Lambda = (0,\infty)$ and $u \in L^2(\Omega)$ by

$$\mathcal{R}_\delta(u) = \frac{1}{\delta^n} \int_\Omega \int_\Omega \frac{|u(x) - u(y)|}{|x-y|} \rho\left(\frac{|x-y|}{\delta}\right)\ \mathrm{d}x\ \mathrm{d}y. \tag{7.39}$$

Remark 7.5.4. a) As ρ is non-increasing, we have for all $0 < \delta < \bar{\delta}$ and $x, y \in \Omega$ that $\rho(|x-y|/\delta) \leq \rho(|x-y|/\bar{\delta})$; consequently,

$$\mathcal{R}_\delta(u) \leq \frac{\bar{\delta}^n}{\delta^n}\mathcal{R}_{\bar{\delta}}(u)$$

for all $u \in L^2(\Omega)$.

b) Note that the assumption (H) from Section 7.2 is satisfied here; in particular, $\mathcal{R}_\delta$ is L^2-weakly lower semicontinuous. Indeed, as the dependence of the integrand on u is convex, it is enough to prove strong lower semicontinuity in $L^2(\Omega)$. This is in turn a simple consequence of Fatou's lemma.

c) In this set-up, the sets $K_{\delta,j}$ in (7.32) consist of a single element $w_j^{(\delta)} \in L^2(\Omega)$ in light of the strict convexity of the fidelity term and convexity of $\mathcal{R}_\delta$. The upper-level functional from (7.33) then becomes

$$\mathcal{I} : (0,\infty) \to [0,\infty), \quad \mathcal{I}(\delta) = \|w^{(\delta)} - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)}.$$

△

The nonlocal functionals in (7.39) have been applied to problems in imaging in [21], providing a derivative-free alternative to popular local models. The localization behavior of these functionals as $\delta \to 0$ is well-studied, originally by Bourgain, Brezis, & Mironescu [47] and later extended to the BV-case in [81, 178]. Using these results, we show that, as $\delta \to 0$, the reconstruction functional in our bi-level scheme turns into the TV-reconstruction functional, see Proposition 7.5.5 below. Moreover, in order to get structural stability inside the domain Λ, we exploit the monotonicity properties of the functional $\mathcal{R}_\delta$, cf. Remark 7.5.4 a). Lastly, as $\delta \to \infty$, we observe that the regularization term vanishes.

Proposition 7.5.5 (Mosco-convergence of the regularizers). *Let $\Lambda = (0,\infty)$, $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain and assume that* (H6$_\delta$) *holds. Then,*

$$\overline{\mathcal{R}_\delta} := \mathrm{Mosc}(L^2)\text{-}\lim_{\delta'\to\delta} \mathcal{R}_{\delta'} = \begin{cases} \mathcal{R}_\delta & \text{if } \delta \in (0,\infty), \\ \mathcal{R}_0 & \text{if } \delta = 0, \\ 0 & \text{if } \delta = \infty, \end{cases} \qquad \text{for } \delta \in \overline{\Lambda} = [0,\infty], \tag{7.40}$$

where

$$\mathcal{R}_0 : L^2(\Omega) \to [0,\infty], \quad \mathcal{R}_0(u) = \begin{cases} \kappa_n |Du|(\Omega), & \text{if } u \in BV(\Omega), \\ \infty & \text{if } u \in L^2(\Omega) \setminus BV(\Omega), \end{cases} \tag{7.41}$$

with $\kappa_n = \fint_{\mathbb{S}^{n-1}} |e\cdot\sigma| \, \mathrm{d}\sigma$ *for any* $e \in \mathbb{S}^{n-1}$.

Proof. Given $(\delta_k)_k \subset (0,\infty)$ with limit $\delta \in [0,\infty]$, the arguments below, subdivided into three different regimes, show that the Mosco-limit of $(\mathcal{R}_{\delta_k})_k$ exists and is equal to the right-hand side of (7.40).

Case 1: $\delta = 0$. For the lower bound, take a sequence $u_k \rightharpoonup u$ in $L^2(\Omega)$ and assume without loss of generality that

$$\sup_k \mathcal{R}_{\delta_k}(u_k) < \infty.$$

By [47, Theorem 4], $(u_k)_k$ is relatively compact in $L^1(\Omega)$, so that $u_k \to u$ in $L^1(\Omega)$. We now use the Γ-liminf result with respect to the $L^1(\Omega)$-convergence in [178, Corollary 8], to deduce that

$$\mathcal{R}_0(u) \le \liminf_{k\to\infty} \mathcal{R}_{\delta_k}(u_k),$$

as desired. For the recovery sequence, we may suppose that $u \in L^2(\Omega) \cap BV(\Omega)$. Then, it follows from [178, Corollary 1] that

$$\lim_{k\to\infty} \frac{1}{\delta_k^n} \int_\Omega \int_\Omega \frac{|u(x)-u(y)|}{|x-y|} \rho\left(\frac{|x-y|}{\delta_k}\right) \, \mathrm{d}x \, \mathrm{d}y = \kappa_n |Du|(\Omega),$$

showing that the constant sequence $u_k = u$ for all $k \in \mathbb{N}$ provides a recovery sequence.

Case 2: $\delta \in (0,\infty)$. For the liminf inequality, take a sequence $(u_k)_k$ converging weakly to u in $L^2(\Omega)$. If $\bar\delta \in (0,\delta)$, then $\delta_k > \bar\delta$ for all $k \in \mathbb{N}$ large enough. Hence, it follows from Remark 7.5.4 a) that

$$\liminf_{k\to\infty} \mathcal{R}_{\delta_k}(u_k) \ge \liminf_{k\to\infty} \frac{\bar\delta^n}{\delta_k^n} \mathcal{R}_{\bar\delta}(u_k) \ge \frac{\bar\delta^n}{\delta^n} \mathcal{R}_{\bar\delta}(u),$$

where the last inequality uses the weak lower semicontinuity of $\mathcal{R}_{\bar\delta}$, cf. Remark 7.5.4 b). Letting $\bar\delta \nearrow \delta$ and using the monotone convergence theorem gives

$$\liminf_{k\to\infty} \mathcal{R}_{\delta_k}(u_k) \ge \mathcal{R}_\delta(u).$$

For the limsup inequality, consider $u \in L^2(\Omega)$ with $\mathcal{R}_\delta(u) < \infty$. Since ρ is non-increasing by (H6$_\delta$), we may extend u to a function $\bar u \in L^2(\mathbb{R}^n)$ by reflection across the boundary of the Lipschitz domain Ω such that

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|\bar u(x)-\bar u(y)|}{|x-y|} \rho\left(\frac{|x-y|}{\delta}\right) \, \mathrm{d}x \, \mathrm{d}y < \infty,$$

cf. [47, Proof of Theorem 4]. With $(\varphi_\varepsilon)_\varepsilon$ a family of smooth standard mollifiers, the sequence $u_l := \varphi_{1/l} * \bar{u}$ for $l \in \mathbb{N}$ converges to u in $L^2(\Omega)$ as $l \to \infty$, and we may argue similarly to the proof of the smooth approximation property (iv) in Section 7.4 to conclude that

$$\lim_{l\to\infty} \mathcal{R}_\delta(u_l) = \mathcal{R}_\delta(u).$$

With $\rho_\delta := \delta^{-n}\rho(|\cdot|/\delta)$ and for a fixed $l \in \mathbb{N}$, we find that

$$\begin{aligned} |\mathcal{R}_\delta(u_l) - \mathcal{R}_{\delta_k}(u_l)| &\leq \int_\Omega\int_\Omega \frac{|u_l(x) - u_l(y)|}{|x-y|}|\rho_\delta(x-y) - \rho_{\delta_k}(x-y)| \ \mathrm{d}x \ \mathrm{d}y \\ &\leq \mathrm{Lip}(u_l)|\Omega|\|\rho_\delta - \rho_{\delta_k}\|_{L^1(\mathbb{R}^n)}, \end{aligned}$$

where $\mathrm{Lip}(u_l)$ is the Lipschitz constant of u_l. We have $\rho_{\delta_k} \to \rho_\delta$ in $L^1(\mathbb{R}^n)$ as $k \to \infty$ by a standard argument approximating ρ with smooth functions. Hence, we obtain

$$\lim_{k\to\infty} \mathcal{R}_{\delta_k}(u_l) = \mathcal{R}_\delta(u_l),$$

and, letting $l \to \infty$, results in

$$\lim_{l\to\infty}\lim_{k\to\infty} \mathcal{R}_{\delta_k}(u_l) = \mathcal{R}_\delta(u).$$

The limsup inequality now follows by extracting an appropriate diagonal sequence.

Case 3: $\delta = \infty$. The only nontrivial case is the limsup inequality, for which we take a sequence $(u_l)_l \subset C_c^\infty(\mathbb{R}^n)$ that converges to u in $L^2(\Omega)$. Then, with R larger than the diameter of Ω, one obtains for every $l \in \mathbb{N}$ that

$$\begin{aligned} \mathcal{R}_{\delta_k}(u_l) &= \frac{1}{\delta_k^n}\int_\Omega\int_\Omega \frac{|u_l(x) - u_l(y)|}{|x-y|}\rho\left(\frac{|x-y|}{\delta_k}\right) \ \mathrm{d}x \ \mathrm{d}y \\ &\leq \mathrm{Lip}(u_l)\int_\Omega\int_{\Omega/\delta_k} \rho\left(\left|z - \frac{y}{\delta_k}\right|\right) \ \mathrm{d}z \ \mathrm{d}y \leq \mathrm{Lip}(u_l)\int_\Omega\int_{B_{\frac{R}{\delta_k}}(0)} \rho(|w|) \ \mathrm{d}w \ \mathrm{d}y. \end{aligned}$$

As $k \to \infty$, the last quantity goes to zero since $\rho(|\cdot|) \in L^1(\mathbb{R}^n)$. Therefore, we deduce that

$$\lim_{k\to\infty} \mathcal{R}_{\delta_k}(u_l) = 0,$$

and conclude again with a diagonal argument. □

7.5.3 Conclusions and examples

In both the Brezis & Nguyen and the Aubert & Kornprobst settings, we now find that the extension $\overline{\mathcal{I}} : [0,\infty] \to [0,\infty]$ is given by

$$\overline{\mathcal{I}}(\delta) = \begin{cases} \mathcal{I}(\delta) & \text{if } \delta \in (0,\infty), \\ \|w^{(0)} - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} & \text{if } \delta = 0, \\ \|u^\eta - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} & \text{if } \delta = \infty, \end{cases}$$

where $w_j^{(0)}$ for $j \in \{1,\dots,N\}$ is the unique minimizer of the TV-reconstruction functional $\overline{\mathcal{J}}_{0,j}$ (with different weight factors in the two cases). In particular, we deduce from Theorem 7.2.5 and Corollary 7.2.8 that $\overline{\mathcal{I}}$ is the relaxation of $\mathcal{I}$ and that these extended upper-level functionals $\overline{\mathcal{I}}$ admit minimizers $\bar{\delta} \in [0,\infty]$. To get an intuition about when this optimal parameter is attained at the boundary or in the interior of Λ, we present the following examples.

Example 7.5.6. a) For both settings analyzed in this section, it is clear that if the noisy and clean image coincide, $u^c \equiv u^\eta$, then the reconstruction model with parameter $\delta = \infty$ gives the exact clean image back. Hence, in this case the optimal parameter is attained at the boundary point $\delta = \infty$.

b) Next, we illustrate the case when the optimal parameter is attained at the boundary point $\delta = 0$. Consider the Aubert & Kornprobst setting in Subsection 7.5.2 and let $\Omega = (-1, 1)$, $N = 1$, $u^c = 0$, and $u^\eta(x) = \kappa_n x$ for $x \in (-1, 1)$. The reconstruction of u^η with the total variation regularizer $\mathcal{R}_0$ in (7.41) is of the form

$$w^{(0)} = \max\{\theta_1, \min\{\theta_2, u^\eta\}\} \quad \text{for some } \theta_1, \theta_2 \in \mathbb{R}.$$

To see this, we observe that $\overline{\mathcal{J}}_0(\tilde{u}) \leq \overline{\mathcal{J}}_0(u)$ for any $u \in BV(-1, 1)$ with

$$\tilde{u} = \max\{u^-, \min\{u^+, u^\eta\}\},$$

where $u^- := \operatorname{ess\,inf}_{x\in(-1,1)} u(x)$ and $u^+ := \operatorname{ess\,sup}_{x\in(-1,1)} u(x)$. Indeed, the map $\tilde{u}$ has at most the same total variation as u and does not increase the distance to u^η anywhere. Next, since u^η is an odd function, the same should hold for the minimizer, meaning that $-\theta_1 = \theta_2 =: \theta \in [0, \kappa_n]$. We can now determine the value of θ by optimizing the quantity $\overline{\mathcal{J}}_0(w^{(0)})$ in θ. This boils down to minimizing

$$\frac{2}{3}\kappa_n^2 \left(1 - \frac{\theta}{\kappa_n}\right)^3 + 2\kappa_n\theta,$$

and yields $\theta = 0$. Hence, the reconstruction model for $\delta = 0$ yields the exact clean image, so that $\overline{\mathcal{I}}(0) = 0$. The same conclusions can be drawn for the Brezis & Nguyen setting by replacing κ_n in the example above with $K(\varphi)$.

c) Let us finally address the case when $\overline{\mathcal{I}}$ becomes minimal inside $\Lambda = (0, \infty)$. We work once again with the Aubert & Kornprobst model from Subsection 7.5.2, and assume in addition to (H6$_\delta$) that the function ρ is equal to 1 in a neighborhood of zero. We consider the following conditions on the pair of data points $(u^c, u^\eta) \in L^2(\Omega; \mathbb{R}^N) \times L^2(\Omega; \mathbb{R}^N)$:

(H7$_\delta$) $\|u^\eta - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)} < \|w^{(0)} - u^c\|^2_{L^2(\Omega;\mathbb{R}^N)}$;

(H8$_\delta$) $\sum_{j=1}^N \widetilde{\mathcal{R}}(u^c_j) < \sum_{j=1}^N \widetilde{\mathcal{R}}(u^\eta_j)$;

here, $w^{(0)}$ is the componentwise minimizer of the TV-reconstruction functional $\overline{\mathcal{J}}_0$ and we set

$$\widetilde{\mathcal{R}}(u) := \int_\Omega \int_\Omega \frac{|u(x) - u(y)|}{|x - y|} \, dx \, dy \quad \text{for } u \in L^2(\Omega). \tag{7.42}$$

The two hypotheses above can be realized, for example, by taking $u^\eta = (1 + \varepsilon)u^c$ for some small $\varepsilon > 0$ and $w^{(0)} \neq u^c$.

Notice that (H7$_\delta$) immediately rules out $\delta = 0$ as an optimal candidate, since the reconstruction at $\delta = \infty$ is better. On the other hand, ρ is supposed to be equal to 1 near the zero, so that we infer for large enough δ that

$$\mathcal{R}_\delta(u) = \frac{1}{\delta^n} \int_\Omega \int_\Omega \frac{|u(x) - u(y)|}{|x - y|} \, dx \, dy = \frac{1}{\delta^n} \widetilde{\mathcal{R}}(u) \tag{7.43}$$

for all $u \in L^2(\Omega)$. Since, for large δ, the dependence of the regularizer on δ is of the same type as the weight case from Section 7.3, we may apply Lemma 7.3.4 (i) in view of (H8$_\delta$). This yields, for all δ large enough, that

$$\|u^c - w^{(\delta)}\|^2_{L^2(\Omega;\mathbb{R}^N)} < \|u^c - u^\eta\|^2_{L^2(\Omega;\mathbb{R}^N)},$$

with $w^{(\delta)}$ the minimizer of $\mathcal{J}_\delta$. This shows that the optimal parameter is not attained at $\delta = \infty$ either and, as a result, needs to be attained inside $\Lambda = (0, \infty)$. Hence, the optimal regularizer lies within the class we started with.

The same conclusions can be drawn for the Brezis & Nguyen case described in Subsection 7.5.1 if we assume that $\varphi(t) = ct^r$ for small t with $c > 0$ and $r \geq 2$. One may take, for instance, the normalized version of the second function in Example 7.5.1. We then suppose that the pair of data points (u^c, u^η) satisfies (H7$_\delta$) and (H8$_\delta$), but now instead of (7.42), take

$$\widetilde{\mathcal{R}}(u) := c \int_\Omega \int_\Omega \frac{|u(x) - u(y)|^r}{|x - y|^{n+1}} \, \mathrm{d}x \, \mathrm{d}y \quad \text{for } u \in L^2(\Omega).$$

We observe with $l = \|u^\eta\|_{L^\infty(\Omega;\mathbb{R}^N)}$ (which we assume to be finite) and T^l the truncation as in the proof of Proposition 7.5.3 that

$$\mathcal{J}_\delta(T^l \circ u) \leq \mathcal{J}_\delta(u)$$

for all $u \in L^2(\Omega)$ and $\delta \in (0, \infty)$. Therefore, we may restrict our analysis to functions $u \in L^2(\Omega)$ with $|u(x) - u(y)| \leq 2l$ for all $x, y \in \Omega$. By additionally considering δ large enough, we now find

$$\varphi\left(\frac{|u(x) - u(y)|}{\delta}\right) = c\,\frac{|u(x) - u(y)|^r}{\delta^r};$$

hence,

$$\mathcal{R}_\delta(u) = \frac{c}{\delta^{r-1}} \int_\Omega \int_\Omega \frac{|u(x) - u(y)|^r}{|x - y|^{n+1}} \, \mathrm{d}x \, \mathrm{d}y = \frac{1}{\delta^{r-1}} \widetilde{\mathcal{R}}(u)$$

in analogy to (7.43).

7.6 Tuning the fractional parameter

This final section revolves around regularization via the L^2-norm of the spectral fractional Laplacian of order $s/2$, with s in the parameter range $\Lambda = (0, 1)$. Our aim here is twofold. First, we determine the Mosco-limits of the regularizers, which allows us to conclude in view of the general theory in Section 7.2 that the extended bi-level problem recovers local models at the boundary points of $\overline{\Lambda} = [0, 1]$. Second, we provide analytic conditions ensuring that the optimal parameter lies in the interior of $(0, 1)$, and illustrate them with an explicit example.

The motivation behind the fractional Laplacian as a regularizer comes from [12], where the authors show that replacing the total variation in the classical ROF model [186] with a spectral fractional Laplacian can lead to comparable reconstruction results with a much smaller computational cost, if the order is chosen correctly. An abstract optimization of the fractional parameter for the spectral fractional Laplacian has already been undertaken in [25], although we remark that a convex penalization term is added there to the model to ensure that the optimal fractional parameter lies inside $(0, 1)$.

We begin with the problem set-up. Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain and let $(\psi_m)_{m\in\mathbb{N}} \subset H^1_0(\Omega)$ be a sequence of eigenfunctions associated with the Laplace operator $(-\Delta)$ forming an orthonormal basis of $L^2(\Omega)$. With the corresponding eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots \nearrow \infty$, it holds for every $m \in \mathbb{N}$ that

$$\begin{cases} (-\Delta)\psi_m = \lambda_m \psi_m & \text{in } \Omega, \\ \psi_m = 0 & \text{on } \partial\Omega. \end{cases} \tag{7.44}$$

Denoting the projection of any $u \in L^2(\Omega)$ onto the mth eigenfunction ψ_m by

$$\hat{u}_m := \langle u, \psi_m \rangle_{L^2(\Omega)},$$

we have the representation $u = \sum_{m=1}^{\infty} \hat{u}_m \psi_m$.

With this at hand, one can define for $s \in (0, 1)$ the fractional Sobolev spaces

$$\mathbb{H}^s(\Omega) := \Big\{ u = \sum_{m=1}^{\infty} \hat{u}_m \psi_m \in L^2(\Omega) \ : \ \sum_{m=1}^{\infty} \lambda_m^s \hat{u}_m^2 < \infty \Big\},$$

endowed with the inner product

$$\langle u, v \rangle_{\mathbb{H}^s(\Omega)} := \sum_{m=1}^{\infty} \lambda_m^s \hat{u}_m \hat{v}_m.$$

It holds that $\mathbb{H}^s(\Omega)$ is a Hilbert space for every $s \in (0, 1)$; for more details on these spaces, we refer, e.g., to [58, 171]. In view of (7.44), the so-called spectral fractional Laplacian of order $s/2$ (with Dirichlet boundary conditions) on these spaces is defined as

$$(-\Delta_D)^{s/2} : \mathbb{H}^s(\Omega) \to L^2(\Omega), \quad (-\Delta_D)^{s/2} u = \sum_{m=1}^{\infty} \lambda_m^{s/2} \hat{u}_m \psi_m.$$

For $s \in (0, 1)$, we consider the regularizer

$$\mathcal{R}_s : L^2(\Omega) \to [0, \infty], \quad \mathcal{R}_s(u) = \begin{cases} \mu \|(-\Delta_D)^{s/2} u\|^2_{L^2(\Omega)} & \text{for } u \in \mathbb{H}^s(\Omega), \\ \infty & \text{otherwise}, \end{cases} \tag{7.45}$$

with some $\mu > 0$. At the end of this section (see Remark 7.6.4), the weight parameter μ will be used to exhibit examples where structure preservation holds. The regularizers $\mathcal{R}_s$ coincide with $\mu\|\cdot\|^2_{\mathbb{H}^s(\Omega)}$ on $\mathbb{H}^s(\Omega)$, and are L^2-weakly lower semicontinuous because $u_k \rightharpoonup u$ in $L^2(\Omega)$ yields

$$\liminf_{k\to\infty} \mathcal{R}_s(u_k) = \liminf_{k\to\infty} \mu \sum_{m=1}^{\infty} \lambda_m^s \widehat{(u_k)}_m^2 \geq \mu \sum_{m=1}^{\infty} \lambda_m^s \widehat{u}_m^2 = \mathcal{R}_s(u)$$

by a discrete version of Fatou's lemma. Therefore, the hypotheses in (H) from Section 7.2 are satisfied.

Next, we determine the Mosco-limits of the regularizers, and thereby, provide the basis for extending the upper-level functional according to Section 7.2.

Proposition 7.6.1 (Mosco-convergence of the regularizers). *Let $\Lambda := (0, 1)$ and $\mathcal{R}_s$ for each $s \in \Lambda$ be given by (7.45). Then, for $u \in L^2(\Omega)$ and $s \in \overline{\Lambda} = [0, 1]$,*

$$\overline{\mathcal{R}}_s(u) = \mathrm{Mosc}(L^2)\text{-}\lim_{s'\to s} \mathcal{R}_{s'}(u) = \begin{cases} \mathcal{R}_s(u) & \text{if } s \in (0, 1), \\ \mu\|u\|^2_{L^2(\Omega)} & \text{if } s = 0, \\ \mu\|\nabla u\|^2_{L^2(\Omega)} + \chi_{H_0^1(\Omega)}(u) & \text{if } s = 1. \end{cases} \tag{7.46}$$

Proof. Let us observe up front that for all $u \in L^2(\Omega)$,

$$\|u\|^2_{L^2(\Omega)} = \sum_{m=1}^{\infty} \widehat{u}_m^2 \quad \text{and} \quad \|\nabla u\|^2_{L^2(\Omega)} + \chi_{H_0^1(\Omega)}(u) = \sum_{m=1}^{\infty} \lambda_m \widehat{u}_m^2; \tag{7.47}$$

indeed, the first formula is simply Parseval's identity, while the second one is a consequence of $\nabla u = \sum_{m=1}^{\infty} \hat{u}_m \nabla \psi_m$ for $u \in H_0^1(\Omega)$ and the orthogonality in $L^2(\Omega; \mathbb{R}^n)$ of the gradients $(\nabla \psi_m)_m$ with

$$\|\nabla \psi_m\|^2_{L^2(\Omega;\mathbb{R}^n)} = -\int_\Omega \psi_m \, \Delta \psi_m \, \mathrm{d}x = \int_\Omega \lambda_m \psi_m^2 \, \mathrm{d}x = \lambda_m.$$

Fixing a sequence $(s_k)_k \subset (0,1)$ with limit $s \in [0,1]$, we want to prove now that the Mosco-limit of $(\mathcal{R}_{s_k})_k$ exists and is given by the right-hand side of (7.46).

Step 1: The liminf-inequality. Let $u_k \rightharpoonup u$ in $L^2(\Omega)$, and assume without loss of generality that $\liminf_{k\to\infty} \mathcal{R}_{s_k}(u_k) < \infty$. Then, since $\widehat{(u_k)}_m \to \widehat{u}_m$ for each $m \in \mathbb{N}$ as $k \to \infty$, it follows from a discrete version of Fatou's lemma that

$$\infty > \liminf_{k\to\infty} \mathcal{R}_{s_k}(u_k) = \liminf_{k\to\infty} \mu \sum_{m=1}^{\infty} \lambda_m^{s_k} \widehat{(u_k)}_m^2 \geq \mu \sum_{m=1}^{\infty} \lambda_m^s \widehat{u}_m^2.$$

In light of (7.47) for the cases $s \in \{0,1\}$, the last quantity equals the regularizer on the right hand side of (7.46) in all the three regimes. This finishes the proof of the lower bound.

Step 2: Construction of a recovery sequence. We first consider the $u \in H_0^1(\Omega)$ case. By the regularity of u and Lebesgue's dominated converge theorem (applied to the counting measure) and by considering the constant recovery sequence $u_k = u$, we get

$$\lim_{k\to\infty} \mathcal{R}_{s_k}(u_k) = \lim_{k\to\infty} \mathcal{R}_{s_k}(u) = \lim_{k\to\infty} \mu \sum_{m=1}^{\infty} \lambda_m^{s_k} \widehat{u}_m^2 = \mu \sum_{m=1}^{\infty} \lambda_m^s \widehat{u}_m^2 = \mathcal{R}_s(u),$$

which concludes the proof for $u \in H_0^1(\Omega)$.

In the general case where $u \in \mathbb{H}^s(\Omega)$, we consider the sequence $(u_l)_l \subset H_0^1(\Omega)$ defined by $u_l := \sum_{m=1}^{l} \hat{u}_m \psi_m$ for every $l \in \mathbb{N}$. Then, by construction, $u_l \to u$ strongly in $L^2(\Omega)$ and

$$\lim_{l\to\infty} \sum_{m=1}^{\infty} \lambda_m^s \widehat{(u_l)}_m^2 = \lim_{l\to\infty} \sum_{m=1}^{l} \lambda_m^s \widehat{u}_m^2 = \sum_{m=1}^{\infty} \lambda_m^s \widehat{u}_m^2.$$

The existence of a recovery sequence follows then by classical diagonalization arguments, using the previous case. □

Given clean and noisy images, $u^c, u^\eta \in L^2(\Omega; \mathbb{R}^N)$, we work with the reconstruction functionals

$$\mathcal{J}_{s,j} : L^2(\mathbb{R}^n) \to [0,\infty], \quad \mathcal{J}_{s,j}(u) = \|u - u_j^\eta\|_{L^2(\Omega)}^2 + \mathcal{R}_s(u)$$

for $s \in (0,1)$ and $j \in \{1,\dots,N\}$. Recalling $(\mathcal{T})$ and $(\overline{\mathcal{T}})$, we obtain as a consequence of Proposition 7.6.1 that the extension of the upper-level functional $\mathcal{I}$ to $\overline{\Lambda}$ is given by

$$\overline{\mathcal{I}} : [0,1] \to [0,\infty], \qquad \overline{\mathcal{I}}(s) = \|w^{(s)} - u^c\|_{L^2(\Omega;\mathbb{R}^N)}^2;$$

here, $w^{(s)} = (w_1^{(s)}, \dots, w_N^{(s)})$ with $w_j^{(s)}$ the unique minimizer of the strictly convex functional

$$\overline{\mathcal{J}}_{s,j}(u) = \|u - u_j^\eta\|_{L^2(\Omega)}^2 + \overline{\mathcal{R}}_s(u) = \sum_{m=1}^{\infty} (\hat{u}_m - \widehat{(u_j^\eta)}_m)^2 + \mu \lambda_m^s \hat{u}_m^2. \tag{7.48}$$

By Theorem 7.2.5, $\overline{\mathcal{I}}$ is then the relaxation of $\mathcal{I}$ and has a minimizer in $\overline{\Lambda} = [0,1]$.

We now continue by exhibiting conditions under which the minimum of $\overline{\mathcal{I}}$ is attained inside $(0,1)$. This is based on a direct approach, observing that the components of $w^{(s)}$ can be determined explicitly by minimizing the entries of the sum in (7.48) individually. This gives the representation

$$w_j^{(s)} = \sum_{m=1}^{\infty} \frac{1}{1 + \mu \lambda_m^s} \widehat{(u_j^\eta)}_m \psi_m \qquad \text{for } j \in \{1,\dots,N\}. \tag{7.49}$$

The following lemma investigates how $w^{(s)}$ varies with s. In the $s > 0$ case, this lemma is essentially contained in [25, Theorem 2] (i.e., in a slightly different setting with periodic instead of Dirichlet boundary conditions). The proof below contains some additional details for the reader's convenience.

Lemma 7.6.2. *Assume that $u^\eta \in \mathbb{H}^\varepsilon(\Omega;\mathbb{R}^N)$ for some $\varepsilon \in (0,1)$. Then, the map $[0,1] \mapsto L^2(\Omega;\mathbb{R}^N)$, $s \mapsto w^{(s)}$ is Fréchet-differentiable with derivative*

$$\partial_s w^{(s)} = -\sum_{m=1}^{\infty} \frac{\mu \log(\lambda_m)\lambda_m^s}{(1+\mu\lambda_m^s)^2} \widehat{u^\eta}_m \psi_m. \tag{7.50}$$

Proof. For $j \in \{1,\dots,N\}$, we set

$$v_j := -\sum_{m=1}^{\infty} \frac{\mu \log(\lambda_m)\lambda_m^s}{(1+\mu\lambda_m^s)^2} \widehat{(u_j^\eta)}_m \psi_m,$$

which is a well-defined element of $L^2(\Omega)$ for all $s \in [0,1]$ because $u_j^\eta \in \mathbb{H}^\varepsilon(\Omega)$. Since

$$\frac{w_j^{(\beta)} - w_j^{(s)}}{t-s} = \sum_{m=1}^{\infty} \frac{1}{t-s}\left(\frac{1}{1+\mu\lambda_m^t} - \frac{1}{1+\mu\lambda_m^s}\right)\widehat{(u_j^\eta)}_m \psi_m, \quad s,t \in [0,1],$$

in view of (7.49), we can apply the mean value theorem to obtain, for each $m \in \mathbb{N}$, a value γ in between s and t such that

$$\left|\frac{1}{t-s}\left(\frac{1}{1+\mu\lambda_m^t} - \frac{1}{1+\mu\lambda_m^s}\right)\right| \le \left|\frac{\mu\log(\lambda_m)\lambda_m^\gamma}{(1+\mu\lambda_m^\gamma)^2}\right| \le |\log(\lambda_m)|.$$

Exploiting once again that $u_j^\eta \in \mathbb{H}^\varepsilon(\Omega)$ gives

$$\begin{aligned}\left\|\frac{w_j^{(\beta)} - w_j^{(s)}}{t-s} - v_j\right\|_{L^2(\Omega)}^2 &= \sum_{m=1}^{\infty}\left|\frac{1}{t-s}\left(\frac{1}{1+\mu\lambda_m^t} - \frac{1}{1+\mu\lambda_m^s}\right) + \frac{\mu\log(\lambda_m)\lambda_m^s}{(1+\mu\lambda_m^s)^2}\right|^2 \widehat{(u_j^\eta)}_m^2 \\ &\le \sum_{m=1}^{\infty} |\log(\lambda_m)|^2 \widehat{(u_j^\eta)}_m^2 < \infty.\end{aligned}$$

In particular, we may take the limit $t \to s$ on the left-hand side of the preceding estimate and interchange with the sum to show the claim. □

It follows as a consequence of Lemma 7.6.2 that the upper level function $\overline{\mathcal{I}} : [0,1] \to [0,\infty]$ is differentiable with derivative

$$\overline{\mathcal{I}}'(s) = 2\left\langle \partial_s w^{(s)}, w^{(s)} - u^c \right\rangle_{L^2(\Omega;\mathbb{R}^N)}$$

for $s \in [0,1]$; at the boundary points $s = 0$ and $s = 1$, $\overline{\mathcal{I}}'(s)$ stands for the one-sided derivative. Plugging in the identities (7.50) and (7.49) in the inner product and using that the family $(\psi_m)_m$ is orthonormal yields

$$\overline{\mathcal{I}}'(s) = -2\sum_{j=1}^{N}\sum_{m=1}^{\infty} \frac{\mu\log(\lambda_m)\lambda_m^s}{(1+\mu\lambda_m^s)^2}\widehat{(u_j^\eta)}_m \left(\frac{1}{1+\mu\lambda_m^s}\widehat{(u_j^\eta)}_m - \widehat{(u_j^c)}_m\right), \tag{7.51}$$

for $s \in [0,1]$. Observe that the simple conditions

$$\overline{\mathcal{I}}'(0) < 0 \quad \text{and} \quad \overline{\mathcal{I}}'(1) > 0,$$

imply that $\overline{\mathcal{I}}$ does not attain its minimizer at $s = 0$ or at $s = 1$, respectively. After taking $s = 0$ and $s = 1$ in (7.51) and simplifying, these requirements can be written as follows:

$$(\mathrm{H1}_s)\quad \sum_{j=1}^{N}\sum_{m=1}^{\infty} \log(\lambda_m)\widehat{(u_j^\eta)}_m \left(\widehat{(u_j^\eta)}_m - (1+\mu)\widehat{(u_j^c)}_m\right) > 0;$$

$$(\mathrm{H2}_s)\quad \sum_{j=1}^{N}\sum_{m=1}^{\infty} \frac{\log(\lambda_m)\lambda_m}{(1+\mu\lambda_m)^3}\widehat{(u_j^\eta)}_m \left(\widehat{(u_j^\eta)}_m - (1+\mu\lambda_m)\widehat{(u_j^c)}_m\right) < 0.$$

Since $(\mathrm{H1}_s)$ guarantees that the minimizer of $\overline{\mathcal{I}}$ is not $s = 0$ and $(\mathrm{H2}_s)$ ensures the minimizer to be different from $s = 1$, Corollary 7.2.8 (*iii*) yields the following result.

Corollary 7.6.3. *Suppose that $u^\eta \in \mathbb{H}^\varepsilon(\Omega;\mathbb{R}^N)$ for some $\varepsilon \in (0,1)$, and that assumptions* $(\mathrm{H1}_s)$ *and* $(\mathrm{H2}_s)$ *are satisfied. Then, $\mathcal{I}$ admits a minimizer $\bar{s} \in (0,1)$.*

We close this section with an interpretation of the conditions $(\mathrm{H1}_s)$ and $(\mathrm{H2}_s)$, and a specific example in which they are both satisfied.

Remark 7.6.4. a) Suppose that $N = 1$. Decomposing the noisy image into the sum of the clean image and the noise, i.e., $u^\eta = u^c + \eta$, turns $(\mathrm{H1}_s)$ and $(\mathrm{H2}_s)$ into

$$\begin{cases} \displaystyle\sum_{m=1}^{\infty} \log(\lambda_m)\left(-\mu\widehat{u^c}_m^2 + (1-\mu)\widehat{u^c}_m\widehat{\eta}_m + \widehat{\eta}_m^2\right) > 0, \\ \displaystyle\sum_{m=1}^{\infty} \frac{\log(\lambda_m)\lambda_m}{(1+\mu\lambda_m)^3}\left(-\mu\lambda_m\widehat{u^c}_m^2 + (1-\mu\lambda_m)\widehat{u^c}_m\widehat{\eta}_m + \widehat{\eta}_m^2\right) < 0. \end{cases} \tag{7.52}$$

If we assume that the noise has mostly high frequencies and that the clean image has mostly moderate frequencies, then the mixed terms in (7.52) will be small. The first condition is then close to

$$-\mu\sum_{m=1}^{\infty}\log(\lambda_m)\widehat{u^c}_m^2 + \sum_{m=1}^{\infty}\log(\lambda_m)\widehat{\eta}_m^2 > 0,$$

which holds for sufficiently small μ. Similarly, for sufficiently large μ, the second condition is satisfied. As we analyse in b) below, there are instances where we can find a range for μ that implies both conditions.

b) In the case where $\Omega = (0,\pi)^2$, by indexing the eigenfunctions via $m = (m_1, m_2) \in \mathbb{N}^2$, we find

$$\psi_m(x) = \sin(m_1x_1)\sin(m_2x_2)$$

with corresponding eigenvalues $\lambda_m = m_1^2 + m_2^2$. By choosing $u^c = \psi_{(1,1)}$ as the clean image and $\eta = \frac{1}{10}\psi_{(10,10)}$ as the noise, the condition (7.52) turns into

$$\begin{cases} -100\,\mu\log(2) + \log(200) > 0, \\ -\mu\dfrac{4\log(2)}{(1+2\mu)^3} + \dfrac{2\log(200)}{(1+200\mu)^3} < 0, \end{cases}$$

which is satisfied for

$$0.0236 \approx \mu_- < \mu < \mu_+ \approx 0.0764.$$

On the other hand, when $\mu = 0.023$, then $s = 1$ is optimal, while the optimal solution for $\mu = 0.11$ is $s = 0$. This can be seen numerically as for these values of μ, the derivative $\overline{\mathcal{I}}'$ is either negative or positive on $[0,1]$, respectively. △

Bibliography

[1] N. H. Abel. Solution de quelques problèmes à l'aide d'intégrales définies. In L. Sylow and S. Lie, editors, *Oeuvres complètes de Niels Henrik Abel: Nouvelle édition*, volume 1 of *Cambridge Library Collection - Mathematics*, pages 11–27. Cambridge University Press, 2012.

[2] H. Abels and G. Grubb. Fractional-order operators on nonsmooth domains. *J. Lond. Math. Soc. (2)*, 107(4):1297–1350, 2023.

[3] E. Acerbi and N. Fusco. Semicontinuity problems in the calculus of variations. *Arch. Rational Mech. Anal.*, 86(2):125–145, 1984.

[4] J. Aczél and J. Dhombres. *Functional equations in several variables*, volume 31 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989. With applications to mathematics, information theory and to the natural and social sciences.

[5] R. A. Adams. *Sobolev spaces*. Academic Press, New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[6] R. A. Adams. On the Orlicz-Sobolev imbedding theorem. *J. Functional Analysis*, 24(3):241–257, 1977.

[7] J. J. Alibert and G. Bouchitté. Non-uniform integrability and generalized Young measures. *J. Convex Anal.*, 4(1):129–147, 1997.

[8] R. Alicandro, N. Ansini, A. Braides, A. Piatnitski, and A. Tribuzio. *A Variational Theory of Convolution-Type Functionals*. SpringerBriefs on PDEs and Data Science. Springer Singapore, 2023.

[9] G. Allaire. Homogenization and two-scale convergence. *SIAM J. Math. Anal.*, 23(6):1482–1518, 1992.

[10] L. Ambrosio and G. Dal Maso. On the relaxation in $\mathrm{BV}(\Omega; \mathbf{R}^m)$ of quasi-convex integrals. *J. Funct. Anal.*, 109(1):76–97, 1992.

[11] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[12] H. Antil and S. Bartels. Spectral approximation of fractional PDEs in image processing and phase field modeling. *Comput. Methods Appl. Math.*, 17(4):661–678, 2017.

[13] H. Antil, Z. W. Di, and R. Khatri. Bilevel optimization, deep learning and fractional Laplacian regularization with applications in tomography. *Inverse Problems*, 36(6):064001, 22, 2020.

[14] H. Antil, H. Díaz, T. Jing, and A. Schikorra. Nonlocal bounded variations with applications. *SIAM J. Math. Anal.*, 56(2):1903–1935, 2024.

[15] H. Antil and C. N. Rautenberg. Sobolev spaces with non-Muckenhoupt weights, fractional elliptic operators, and applications. *SIAM J. Math. Anal.*, 51(3):2479–2503, 2019.

[16] A. Arroyo-Rabasa. Functional and variational aspects of nonlocal operators associated with linear PDEs. Preprint, arXiv:2210.13161, 2023.

[17] A. Arroyo-Rabasa, G. De Philippis, and F. Rindler. Lower semicontinuity and relaxation of linear-growth integral functionals under PDE constraints. *Adv. Calc. Var.*, 13(3):219–255, 2020.

[18] H. Attouch. *Variational convergence for functions and operators.* Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[19] H. Attouch, G. Buttazzo, and G. Michaille. *Variational analysis in Sobolev and BV spaces*, volume 17 of *MOS-SIAM Series on Optimization.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, second edition, 2014. Applications to PDEs and optimization.

[20] G. Aubert and P. Kornprobst. *Mathematical problems in image processing*, volume 147 of *Applied Mathematical Sciences.* Springer, New York, second edition, 2006. Partial differential equations and the calculus of variations, With a foreword by Olivier Faugeras.

[21] G. Aubert and P. Kornprobst. Can the nonlocal characterization of Sobolev spaces by Bourgain et al. be useful for solving variational problems? *SIAM J. Numer. Anal.*, 47(2):844–860, 2009.

[22] A. Audrito, J.-C. Felipe-Navarro, and X. Ros-Oton. The Neumann problem for the fractional Laplacian: regularity up to the boundary. *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (5)*, 24(2):1155–1222, 2023.

[23] J. M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rational Mech. Anal.*, 63(4):337–403, 1977.

[24] J. M. Ball. Discontinuous equilibrium solutions and cavitation in nonlinear elasticity. *Philos. Trans. R. Soc. Lond. Ser. A*, 306:557–611, 1982.

[25] S. Bartels and N. Weber. Parameter learning and fractional differential operators: applications in regularized image denoising and decomposition problems. *Math. Control Relat. Fields*, 13(1):35–62, 2023.

[26] J. C. Bellido, J. Cueto, M. Foss, and P. Radu. Nonlocal Green theorems and Helmholtz decompositions for truncated fractional gradients. Preprint, arXiv:2311.05465, 2023.

[27] J. C. Bellido, J. Cueto, and C. Mora-Corral. Bond-based peridynamics does not converge to hyperelasticity as the horizon goes to zero. *J. Elasticity*, 141(2):273–289, 2020.

[28] J. C. Bellido, J. Cueto, and C. Mora-Corral. Fractional Piola identity and polyconvexity in fractional spaces. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 37(4):955–981, 2020.

[29] J. C. Bellido, J. Cueto, and C. Mora-Corral. Γ-convergence of polyconvex functionals involving s-fractional gradients to their local counterparts. *Calc. Var. Partial Differential Equations*, 60(1):Paper No. 7, 29, 2021.

[30] J. C. Bellido, J. Cueto, and C. Mora-Corral. Minimizers of nonlocal polyconvex energies in nonlocal hyperelasticity. *Adv. Calc. Var.*, 2023.

[31] J. C. Bellido, J. Cueto, and C. Mora-Corral. Non-local gradients in bounded domains motivated by continuum mechanics: fundamental theorem of calculus and embeddings. *Adv. Nonlinear Anal.*, 12(1):Paper No. 20220316, 48, 2023.

[32] J. C. Bellido, J. Cueto, and C. Mora-Corral. Eringen's model via linearization of nonlocal hyperelasticity. *Math. Mech. Solids*, 29(4):686–703, 2024.

[33] J. C. Bellido and C. Mora-Corral. Existence for nonlocal variational problems in peridynamics. *SIAM J. Math. Anal.*, 46(1):890–916, 2014.

[34] J. C. Bellido and C. Mora-Corral. Lower semicontinuity and relaxation via Young measures for nonlocal variational problems and applications to peridynamics. *SIAM J. Math. Anal.*, 50(1):779–809, 2018.

[35] J. C. Bellido, C. Mora-Corral, and P. Pedregal. Hyperelasticity as a Γ-limit of peridynamics when the horizon goes to zero. *Calc. Var. Partial Differential Equations*, 54(2):1643–1670, 2015.

[36] J. C. Bellido, C. Mora-Corral, and H. Schönberger. Nonlocal gradients: Fundamental theorem of calculus, Poincaré inequalities and embeddings. Preprint, arXiv:2402.16487, 2024.

[37] J. C. Bellido and A. Ortega. A restricted nonlocal operator bridging together the Laplacian and the fractional Laplacian. *Calc. Var. Partial Differential Equations*, 60(2):Paper No. 71, 29, 2021.

[38] J. C. Bellido and A. Ortega. Spectral stability for the peridynamic fractional p-Laplacian. *Appl. Math. Optim.*, 84:S253–S276, 2021.

[39] M. Bellieud and G. Bouchitté. Homogenization of elliptic problems in a fiber reinforced structure. Non local effects. *Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser.*, 26(3):407–436, 1998.

[40] M. Belloni, G. Buttazzo, and L. Freddi. Completion by Gamma-convergence for optimal control problems. *Ann. Fac. Sci. Toulouse Math. (6)*, 2(2):149–162, 1993.

[41] C. Bennett and R. Sharpley. *Interpolation of operators*, volume 129 of *Pure and Applied Mathematics*. Academic Press, Inc., Boston, MA, 1988.

[42] M. Benning and M. Burger. Modern regularization methods for inverse problems. *Acta Numer.*, 27:1–111, 2018.

[43] J. Bernoulli. Problema novum ad cujus solutionem mathematici invitantur. *Acta Eruditorum*, 15:264–269, 1696.

[44] F. Bobaru, J. T. Foster, P. H. Geubelle, and S. A. Silling, editors. *Handbook of peridynamic modeling*. Advances in Applied Mathematics. CRC Press, Boca Raton, FL, 2017.

[45] K. Bogdan, K. Burdzy, and Z.-Q. Chen. Censored stable processes. *Probab. Theory Related Fields*, 127(1):89–152, 2003.

[46] J. Boulanger, P. Elbau, C. Pontow, and O. Scherzer. Non-local functionals for imaging. In *Fixed-point algorithms for inverse problems in science and engineering*, volume 49 of *Springer Optim. Appl.*, pages 131–154. Springer, New York, 2011.

[47] J. Bourgain, H. Brezis, and P. Mironescu. Another look at Sobolev spaces. In *Optimal control and partial differential equations*, pages 439–455. IOS, Amsterdam, 2001.

[48] A. Braides. Homogenization of some almost periodic coercive functional. *Rend. Accad. Naz. Sci. XL Mem. Mat. (5)*, 9(1):313–321, 1985.

[49] A. Braides. *Γ-convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[50] A. Braides, M. Maslennikov, and L. Sigalotti. Homogenization by blow-up. *Appl. Anal.*, 87(12):1341–1356, 2008.

[51] K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM J. Imaging Sci.*, 3(3):492–526, 2010.

[52] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[53] H. Brezis and H.-M. Nguyen. Non-local functionals related to the total variation and connections with image processing. *Ann. PDE*, 4(1):Art. 9, 77, 2018.

[54] E. Bruè, M. Calzi, G. E. Comi, and G. Stefani. A distributional approach to fractional Sobolev spaces and fractional variation: asymptotics II. *C. R. Math. Acad. Sci. Paris*, 360:589–626, 2022.

[55] G. Buttazzo. Some relaxation problems in optimal control theory. *J. Math. Anal. Appl.*, 125(1):272–287, 1987.

[56] G. Buttazzo and G. Dal Maso. Γ-convergence and optimal control problems. *J. Optim. Theory Appl.*, 38(3):385–407, 1982.

[57] L. Caffarelli and L. Silvestre. An extension problem related to the fractional Laplacian. *Comm. Partial Differential Equations*, 32(7-9):1245–1260, 2007.

[58] L. A. Caffarelli and P. R. Stinga. Fractional elliptic equations, caccioppoli estimates and regularity. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, 33:767–807, 2016.

[59] P. M. Campos and J. F. Rodrigues. Unilateral problems for quasilinear operators with fractional Riesz gradients. Preprint, arXiv:2311.18428, 2023.

[60] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.*, 76(2):167–188, 1997.

[61] T. Champion, L. De Pascale, and F. Prinari. Γ-convergence and absolute minimizers for supremal functionals. *ESAIM Control Optim. Calc. Var.*, 10(1):14–27, 2004.

[62] Y. Chen, T. Pock, R. Ranftl, and H. Bischof. Revisiting loss-specific training of filter-based mrfs for image restoration. In J. Weickert, M. Hein, and B. Schiele, editors, *Pattern Recognition*, pages 271–281, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[63] Y. Chen, R. Ranftl, and T. Pock. Insights into analysis operator learning: from patch-based sparse models to higher order MRFs. *IEEE Trans. Image Process.*, 23(3):1060–1072, 2014.

[64] C. V. Chung, J. C. De los Reyes, and C. B. Schönlieb. Learning optimal spatially-dependent regularization parameters in total variation image denoising. *Inverse Problems*, 33(7):074005, 31, 2017.

[65] G. E. Comi, D. Spector, and G. Stefani. The fractional variation and the precise representative of $BV^{\alpha,p}$ functions. *Fract. Calc. Appl. Anal.*, 25(2):520–558, 2022.

[66] G. E. Comi and G. Stefani. A distributional approach to fractional Sobolev spaces and fractional variation: existence of blow-up. *J. Funct. Anal.*, 277(10):3373–3435, 2019.

[67] G. E. Comi and G. Stefani. Leibniz rules and Gauss-Green formulas in distributional fractional spaces. *J. Math. Anal. Appl.*, 514(2):Paper No. 126312, 41, 2022.

[68] G. E. Comi and G. Stefani. A distributional approach to fractional Sobolev spaces and fractional variation: asymptotics I. *Rev. Mat. Complut.*, 36(2):491–569, 2023.

[69] G. E. Comi and G. Stefani. Failure of the local chain rule for the fractional variation. *Port. Math.*, 80(1-2):1–25, 2023.

[70] J. B. Conway. *A course in functional analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

[71] J. Cueto. *Mathematical analysis of fractional and nonlocal models from nonlinear Solid Mechanics*. PhD thesis, Universidad de Castilla-La Mancha, 2021.

[72] J. Cueto, C. Kreisbeck, and H. Schönberger. A variational theory for integral functionals involving finite-horizon fractional gradients. *Fract. Calc. Appl. Anal.*, 26(5):2001–2056, 2023.

[73] J. Cueto, C. Kreisbeck, and H. Schönberger. Γ-convergence involving nonlocal gradients with varying horizon: Recovery of fractional and local models. Preprint, arXiv:2404.18509, 2024.

[74] B. Dacorogna. Quasiconvexity and relaxation of nonconvex problems in the calculus of variations. *J. Functional Analysis*, 46(1):102–118, 1982.

[75] B. Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2008.

[76] B. Dacorogna and P. Marcellini. *Implicit partial differential equations*, volume 37 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Boston, Inc., Boston, MA, 1999.

[77] B. Dacorogna and P. Marcellini. On the solvability of implicit nonlinear systems in the vectorial case. In *Nonlinear partial differential equations (Evanston, IL, 1998)*, volume 238 of *Contemp. Math.*, pages 89–113. Amer. Math. Soc., Providence, RI, 1999.

[78] B. Dacorogna and G. Pisante. A general existence theorem for differential inclusions in the vector valued case. *Port. Math. (N.S.)*, 62(4):421–436, 2005.

[79] G. Dal Maso. Integral representation on $BV(\Omega)$ of Γ-limits of variational integrals. *Manuscripta Math.*, 30(4):387–416, 1979/80.

[80] G. Dal Maso. *An introduction to Γ-convergence*, volume 8 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Boston, Inc., Boston, MA, 1993.

[81] J. Dávila. On an open question about functions of bounded variation. *Calc. Var. Partial Differential Equations*, 15(4):519–527, 2002.

[82] E. Davoli, R. Ferreira, C. Kreisbeck, and H. Schönberger. Structural changes in nonlocal denoising models arising through bi-level parameter learning. *Appl. Math. Optim.*, 88(1):Paper No. 9, 47, 2023.

[83] E. Davoli, I. Fonseca, and P. Liu. Adaptive image processing: first order PDE constraint regularizers and a bilevel training scheme. *J. Nonlinear Sci.*, 33(3):Paper No. 41, 38, 2023.

[84] E. Davoli and P. Liu. One dimensional fractional order TGV: gamma-convergence and bilevel training scheme. *Commun. Math. Sci.*, 16(1):213–237, 2018.

[85] E. De Giorgi and T. Franzoni. Su un tipo di convergenza variazionale. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat. (8)*, 58(6):842–850, 1975.

[86] J. C. De los Reyes and C.-B. Schönlieb. Image denoising: learning the noise model via nonsmooth PDE-constrained optimization. *Inverse Probl. Imaging*, 7(4):1183–1214, 2013.

[87] J. C. De Los Reyes, C.-B. Schönlieb, and T. Valkonen. The structure of optimal parameters for image restoration problems. *J. Math. Anal. Appl.*, 434(1):464–500, 2016.

[88] J. C. De los Reyes, C.-B. Schönlieb, and T. Valkonen. Bilevel parameter learning for higher-order total variation regularisation models. *J. Math. Imaging Vision*, 57(1):1–25, 2017.

[89] L. M. Del Pezzo and A. M. Salort. The first non-zero Neumann p-fractional eigenvalue. *Nonlinear Anal.*, 118:130–143, 2015.

[90] M. D'Elia, J. C. De Los Reyes, and A. Miniguano-Trujillo. Bilevel parameter learning for nonlocal image denoising models. *J. Math. Imaging Vision*, 63(6):753–775, 2021.

[91] M. D'Elia, Q. Du, and M. D. Gunzburger. Recent progress in mathematical and computational aspects of peridynamics. In G. Z. Voyiadjis, editor, *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*. Springer, 2019.

[92] M. D'Elia, M. Gulian, T. Mengesha, and J. M. Scott. Connections between nonlocal operators: from vector calculus identities to a fractional Helmholtz decomposition. *Fract. Calc. Appl. Anal.*, 25(6):2488–2531, 2022.

[93] M. D'Elia, M. Gulian, H. Olson, and G. E. Karniadakis. Towards a unified theory of fractional and nonlocal vector calculus. *Fract. Calc. Appl. Anal.*, 24(5):1301–1355, 2021.

[94] F. Demengel and G. Demengel. *Functional spaces for the theory of elliptic partial differential equations*. Universitext. Springer, London; EDP Sciences, Les Ulis, 2012. Translated from the 2007 French original by Reinie Erné.

[95] S. Dempe and A. Zemkoho, editors. *Bilevel optimization—advances and next challenges*, volume 161 of *Springer Optimization and Its Applications*. Springer, Cham, [2020] ©2020.

[96] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573, 2012.

[97] R. J. DiPerna and A. J. Majda. Oscillations and concentrations in weak solutions of the incompressible fluid equations. *Comm. Math. Phys.*, 108(4):667–689, 1987.

[98] S. Dipierro, E. P. Lippi, and E. Valdinoci. (Non)local logistic equations with Neumann conditions. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, 40(5):1093–1166, 2023.

[99] S. Dipierro, L. E. Proietti, and E. Valdinoci. Linear theory for a mixed operator with Neumann conditions. *Asymptotic Anal.*, 128(4):571–594, 2022.

[100] S. Dipierro, X. Ros-Oton, and E. Valdinoci. Nonlocal problems with Neumann boundary conditions. *Rev. Mat. Iberoam.*, 33(2):377–416, 2017.

[101] J. Domke. Generic methods for optimization-based modeling. In N. D. Lawrence and M. Girolami, editors, *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*, volume 22 of *Proceedings of Machine Learning Research*, pages 318–326, La Palma, Canary Islands, 21–23 Apr 2012. PMLR.

[102] Q. Du, M. Gunzburger, R. B. Lehoucq, and K. Zhou. A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws. *Math. Models Methods Appl. Sci.*, 23(3):493–540, 2013.

[103] Q. Du, T. Mengesha, and X. Tian. Fractional Hardy-type and trace theorems for nonlocal function spaces with heterogeneous localization. *Anal. Appl. (Singap.)*, 20(3):579–614, 2022.

[104] Q. Du and X. Tian. Stability of nonlocal Dirichlet integrals and implications for peridynamic correspondence material modeling. *SIAM J. Appl. Math.*, 78(3):1536–1552, 2018.

[105] Q. Du, X. Tian, C. Wright, and Y. Yu. Nonlocal trace spaces and extension results for nonlocal calculus. *J. Funct. Anal.*, 282(12):63, 2022. Id/No 109453.

[106] J. Duoandikoetxea. *Fourier analysis*, volume 29 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2001.

[107] P. Elbau. Sequential lower semi-continuity of non-local functionals, 2011.

[108] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, second edition, 2010.

[109] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions.* Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.

[110] H. Federer. *Geometric measure theory*, volume Band 153 of *Die Grundlehren der mathematischen Wissenschaften.* Springer-Verlag New York, Inc., New York, 1969.

[111] G. Foghem and M. Kassmann. A general framework for nonlocal Neumann problems. *Commun. Math. Sci.*, 22(1):15–66, 2024.

[112] I. Fonseca and G. Leoni. *Modern methods in the calculus of variations: L^p spaces.* Springer Monographs in Mathematics. Springer, New York, 2007.

[113] I. Fonseca and S. Müller. Relaxation of quasiconvex functionals in $\mathrm{BV}(\Omega, \mathbf{R}^p)$ for integrands $f(x, u, \nabla u)$. *Arch. Rational Mech. Anal.*, 123(1):1–49, 1993.

[114] M. Foss. Nonlocal poincaré inequalities for integral operators with integrable nonhomogeneous kernels. Preprint, arXiv:1911.10292, 2019.

[115] M. Foss. Traces on general sets in $\mathbb{R}^n$ for functions with no differentiability requirements. *SIAM J. Math. Anal.*, 53(4):4212–4251, 2021.

[116] W. H. Gerstle. *Introduction to practical peridynamics*, volume 1 of *Frontier Research in Computation and Mechanics of Materials and Biology.* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2016. Computational solid mechanics without stress and strain.

[117] G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. *Multiscale Model. Simul.*, 7(3):1005–1028, 2008.

[118] E. Giusti. *Minimal surfaces and functions of bounded variation*, volume 80 of *Monographs in Mathematics.* Birkhäuser Verlag, Basel, 1984.

[119] K. Goebel and S. Reich. *Uniform convexity, hyperbolic geometry, and nonexpansive mappings*, volume 83 of *Monographs and Textbooks in Pure and Applied Mathematics.* Marcel Dekker, Inc., New York, 1984.

[120] C. Goffman and J. Serrin. Sublinear functions of measures and variational integrals. *Duke Math. J.*, 31:159–178, 1964.

[121] H. H. Goldstine. *A history of the calculus of variations from the 17th through the 19th century*, volume 5 of *Studies in the History of Mathematics and Physical Sciences.* Springer-Verlag, New York-Berlin, 1980.

[122] L. Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics.* Springer, New York, third edition, 2014.

[123] L. Grafakos. *Modern Fourier analysis*, volume 250 of *Graduate Texts in Mathematics.* Springer, New York, third edition, 2014.

[124] G. Grubb. Fractional Laplacians on domains, a development of Hörmander's theory of μ-transmission pseudodifferential operators. *Adv. Math.*, 268:478–528, 2015.

[125] G. Grubb. Resolvents for fractional-order operators with nonhomogeneous local boundary conditions. *J. Funct. Anal.*, 284(7):Paper No. 109815, 55, 2023.

[126] F. Grube and T. Hensiek. Robust nonlocal trace spaces and Neumann problems. *Nonlinear Anal.*, 241:113481, 2024.

[127] F. Grube and M. Kassmann. Robust nonlocal trace and extension theorems. Preprint, arXiv:2305.05735, 2023.

[128] Z. Han, T. Mengesha, and X. Tian. Compactness results for a Dirichlet energy of nonlocal gradient with applications. Preprint, arXiv:2404.09079, 2024.

[129] Z. Han and X. Tian. Nonlocal half-ball vector operators on bounded domains: Poincaré inequality and its applications. *Math. Models Methods Appl. Sci.*, 33(12):2507–2556, 2023.

[130] D. Harutyunyan, T. Mengesha, H. Mikayelyan, and J. M. Scott. Fractional Korn's inequalities without boundary conditions. *Math. Mech. Complex Syst.*, 11(4):497–524, 2023.

[131] D. Hilbert. Mathematical problems. *Bull. Amer. Math. Soc.*, 8(10):437–479, 1902.

[132] R. Hilfer, editor. *Applications of fractional calculus in physics.* World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[133] M. Hintermüller and C. N. Rautenberg. Optimal selection of the regularization function in a weighted total variation model. Part I: Modelling and theory. *J. Math. Imaging Vision*, 59(3):498–514, 2017.

[134] G. Holler and K. Kunisch. Learning nonlocal regularization operators. *Math. Control Relat. Fields*, 12(1):81–114, 2022.

[135] G. Holler, K. Kunisch, and R. C. Barnard. A bilevel approach for parameter learning in inverse problems. *Inverse Problems*, 34(11):115012, 28, 2018.

[136] L. Hörmander. *The analysis of linear partial differential operators. III*, volume 274 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 1985. Pseudodifferential operators.

[137] J. Horváth. On some composition formulas. *Proc. Amer. Math. Soc.*, 10:433–437, 1959.

[138] R. Kawasumi, E. Nakai, and M. Shi. Characterization of the boundedness of generalized fractional integral and maximal operators on Orlicz-Morrey and weak Orlicz-Morrey spaces. *Math. Nachr.*, 296(4):1483–1503, 2023.

[139] C. Kreisbeck, A. Ritorto, and E. Zappale. Cartesian convexity as the key notion in the variational existence theory for nonlocal supremal functionals. *Nonlinear Anal.*, 225:Paper No. 113111, 33, 2022.

[140] C. Kreisbeck and H. Schönberger. Quasiconvexity in the fractional calculus of variations: characterization of lower semicontinuity and relaxation. *Nonlinear Anal.*, 215:Paper No. 112625, 26, 2022.

[141] C. Kreisbeck and H. Schönberger. Non-constant functions with zero nonlocal gradient and their role in nonlocal Neumann-type problems. Preprint, arXiv:2402.11308, 2024.

[142] C. Kreisbeck and E. Zappale. Lower semicontinuity and relaxation of nonlocal L^∞-functionals. *Calc. Var. Partial Differential Equations*, 59(4):Paper No. 138, 36, 2020.

[143] C. Kreisbeck and E. Zappale. Loss of double-integral character during relaxation. *SIAM J. Math. Anal.*, 53(1):351–385, 2021.

[144] J. Kristensen and F. Rindler. Characterization of generalized gradient Young measures generated by sequences in $W^{1,1}$ and BV. *Arch. Ration. Mech. Anal.*, 197(2):539–598, 2010.

[145] J. Kristensen and F. Rindler. Relaxation of signed integral functionals in BV. *Calc. Var. Partial Differential Equations*, 37(1-2):29–62, 2010.

[146] M. Kwaśnicki. Ten equivalent definitions of the fractional Laplace operator. *Fract. Calc. Appl. Anal.*, 20(1):7–51, 2017.

[147] S. Lang. *Complex analysis*, volume 103 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, fourth edition, 1999.

[148] H. Lee and Q. Du. Nonlocal gradient operators with a nonspherical interaction neighborhood and their applications. *ESAIM Math. Model. Numer. Anal.*, 54(1):105–128, 2020.

[149] G. Leoni. *A first course in Sobolev spaces*, volume 181 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2017.

[150] D. Liberzon. *Calculus of variations and optimal control theory*. Princeton University Press, Princeton, NJ, 2012. A concise introduction.

[151] A. Lischke, G. Pang, M. Gulian, and et al. What is the fractional Laplacian? A comparative review with new results. *J. Comput. Phys.*, 404:109009, 62, 2020.

[152] P. Liu and C.-B. Schönlieb. Learning optimal orders of the underlying Euclidean norm in total variation image denoising. Preprint, arXiv:1903.11953, 2019.

[153] C. W. K. Lo and J. F. Rodrigues. On a class of nonlocal obstacle type problems related to the distributional Riesz fractional derivative. *Port. Math.*, 80(1-2):157–205, 2023.

[154] E. Madenci and E. Oterkus. *Peridynamic Theory and Its Applications*. Springer, New York, NY, 2014.

[155] P. Marcellini. Approximation of quasiconvex functions, and lower semicontinuity of multiple integrals. *Manuscripta Math.*, 51(1-3):1–28, 1985.

[156] M. M. Meerschaert, J. Mortensen, and S. W. Wheatcraft. Fractional vector calculus for fractional advection-dispersion. *Phys. A: Stat. Mech. and its Appl.*, 367:181–190, 2006.

[157] M. M. Meerschaert and A. Sikorskii. *Stochastic models for fractional calculus*, volume 43 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 2012.

[158] T. Mengesha and Q. Du. On the variational limit of a class of nonlocal functionals related to peridynamics. *Nonlinearity*, 28(11):3999–4035, 2015.

[159] T. Mengesha and Q. Du. Characterization of function spaces of vector fields and an application in nonlinear peridynamics. *Nonlinear Anal.*, 140:82–111, 2016.

[160] T. Mengesha and J. M. Scott. Linearization and localization of nonconvex functionals motivated by nonlinear peridynamic models. *Continuum Mech. Thermodyn.*, 2024.

[161] T. Mengesha and D. Spector. Localization of nonlocal gradients in various topologies. *Calc. Var. Partial Differential Equations*, 52(1-2):253–279, 2015.

[162] N. G. Meyers. Quasi-convexity and lower semi-continuity of multiple variational integrals of any order. *Trans. Amer. Math. Soc.*, 119:125–149, 1965.

[163] Y. Mizuta. *Potential theory in Euclidean spaces*, volume 6 of *GAKUTO International Series. Mathematical Sciences and Applications*. Gakkōtosho Co., Ltd., Tokyo, 1996.

[164] C. Mora-Corral and A. Tellini. Relaxation of a scalar nonlocal variational problem with a double-well potential. *Calc. Var. Partial Differential Equations*, 59(2), 2020.

[165] C. B. Morrey, Jr. Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific J. Math.*, 2:25–53, 1952.

[166] J. Muñoz. Characterisation of the weak lower semicontinuity for a type of nonlocal integral functional: the n-dimensional scalar case. *J. Math. Anal. Appl.*, 360(2):495–502, 2009.

[167] D. Mugnai and E. Proietti Lippi. Neumann fractional p-Laplacian: eigenvalues and existence results. *Nonlinear Anal., Theory Methods Appl., Ser. A, Theory Methods*, 188:455–474, 2019.

[168] S. Müller. Homogenization of nonconvex integral functionals and cellular elastic materials. *Arch. Rational Mech. Anal.*, 99(3):189–212, 1987.

[169] S. Müller. Variational models for microstructure and phase transitions. In *Calculus of variations and geometric evolution problems (Cetraro, 1996)*, volume 1713 of *Lecture Notes in Math.*, pages 85–210. Springer, Berlin, 1999.

[170] S. Müller and M. A. Sychev. Optimal existence theorems for nonhomogeneous differential inclusions. *J. Funct. Anal.*, 181(2):447–475, 2001.

[171] R. H. Nochetto, E. Otárola, and A. J. Salgado. A PDE approach to fractional diffusion in general domains: a priori error analysis. *Found. Comput. Math.*, 15(3):733–791, 2015.

[172] V. Pagliari, K. Papafitsoros, B. Raiţă, and A. Vikelis. Bilevel training schemes in imaging for total variation-type functionals with convex integrands. *SIAM J. Imaging Sci.*, 15(4):1690–1728, 2022.

[173] P. Pedregal. *Parametrized measures and variational principles*, volume 30 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Verlag, Basel, 1997.

[174] P. Pedregal. Weak lower semicontinuity and relaxation for a class of non-local functionals. *Rev. Mat. Complut.*, 29(3):485–495, 2016.

[175] R. L. Pego. Compactness in L^2 and the Fourier transform. *Proc. Amer. Math. Soc.*, 95(2):252–254, 1985.

[176] T. Pock, D. Cremers, H. Bischof, and A. Chambolle. Global solutions of variational models with convex regularization. *SIAM J. Imaging Sci.*, 3(4):1122–1145, 2010.

[177] I. Podlubny. *Fractional differential equations*, volume 198 of *Mathematics in Science and Engineering*. Academic Press, Inc., San Diego, CA, 1999. An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications.

[178] A. C. Ponce. A new approach to Sobolev spaces and connections to Γ-convergence. *Calc. Var. Partial Differential Equations*, 19(3):229–255, 2004.

[179] A. C. Ponce. *Elliptic PDEs, measures and capacities*, volume 23 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2016. From the Poisson equations to nonlinear Thomas-Fermi problems.

[180] M. M. Rao and Z. D. Ren. *Theory of Orlicz spaces*, volume 146 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1991.

[181] Y. G. Reshetnyak. Weak convergence of completely additive vector functions on a set. *Sib. Math. J.*, 9(6):1039–1045, 1968.

[182] F. Rindler. *Calculus of variations*. Universitext. Springer, Cham, 2018.

[183] F. Rindler and G. Shaw. Liftings, Young measures, and lower semicontinuity. *Arch. Ration. Mech. Anal.*, 232(3):1227–1328, 2019.

[184] X. Ros-Oton and J. Serra. The Dirichlet problem for the fractional Laplacian: regularity up to the boundary. *J. Math. Pures Appl. (9)*, 101(3):275–302, 2014.

[185] B. Ross. The development of fractional calculus 1695–1900. *Historia Math.*, 4:75–89, 1977.

[186] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992. Experimental mathematics: computational issues in nonlinear science (Los Alamos, NM, 1991).

[187] T. Runst and W. Sickel. *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, volume 3 of *De Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, 1996.

[188] S. G. Samko, A. A. Kilbas, and O. I. Marichev. *Fractional integrals and derivatives*. Gordon and Breach Science Publishers, Yverdon, 1993. Theory and applications, Edited and with a foreword by S. M. Nikolşkiĭ, Translated from the 1987 Russian original, Revised by the authors.

[189] H. Schönberger. Extending linear growth functionals to functions of bounded fractional variation. *Proc. Roy. Soc. Edinburgh Sect. A*, 154(1):304–327, 2024.

[190] J. Scott and T. Mengesha. A fractional Korn-type inequality. *Discrete Contin. Dyn. Syst.*, 39(6):3315–3343, 2019.

[191] J. M. Scott and Q. Du. Nonlocal problems with local boundary conditions I: function spaces and variational principles. Preprint, arXiv:2307.08855, 2023.

[192] J. M. Scott and Q. Du. Nonlocal problems with local boundary conditions II: Green's identities and regularity of solutions. Preprint, arXiv:2308.05180, 2023.

[193] T.-T. Shieh and D. E. Spector. On a new class of fractional partial differential equations. *Adv. Calc. Var.*, 8(4):321–336, 2015.

[194] T.-T. Shieh and D. E. Spector. On a new class of fractional partial differential equations II. *Adv. Calc. Var.*, 11(3):289–307, 2018.

[195] S. Silling and R. Lehoucq. Peridynamic theory of solid mechanics. In H. Aref and E. van der Giessen, editors, *Advances in Applied Mechanics*, volume 44 of *Advances in Applied Mechanics*, pages 73–168. Elsevier, 2010.

[196] S. A. Silling. Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids*, 48(1):175–209, 2000.

[197] S. A. Silling, M. Epton, O. Weckner, J. Xu, and E. Askari. Peridynamic states and constitutive modeling. *J. Elasticity*, 88(2):151–184, 2007.

[198] S. A. Silling, D. J. Littlewood, and P. Seleson. Variable horizon in a peridynamic medium. *J. Mech. Mater. Struct.*, 10(5):591–612, 2015.

[199] E. M. Stein. *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[200] M. A. Sychev. Comparing various methods of resolving differential inclusions. *J. Convex Anal.*, 18(4):1025–1045, 2011.

[201] Y. Tao, X. Tian, and Q. Du. Nonlocal models with heterogeneous localization and their application to seamless local-nonlocal coupling. *Multiscale Model. Simul.*, 17(3):1052–1075, 2019.

[202] M. F. Tappen, C. Liu, E. H. Adelson, and W. T. Freeman. Learning gaussian conditional random fields for low-level vision. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.

[203] X. Tian and Q. Du. Trace theorems for some nonlocal function spaces with heterogeneous localization. *SIAM J. Math. Anal.*, 49(2):1621–1644, 2017.

[204] H. Triebel. *Theory of function spaces*, volume 78 of *Monographs in Mathematics.* Birkhäuser Verlag, Basel, 1983.

[205] H. Triebel. *The structure of functions*, volume 97 of *Monographs in Mathematics.* Birkhäuser Verlag, Basel, 2001.

[206] B. van Brunt. *The calculus of variations.* Universitext. Springer-Verlag, New York, 2004.

[207] M. I. Višik and G. I. Èskin. Convolution equations in a bounded region. *Uspehi Mat. Nauk*, 20(3 (123)):89–152, 1965.

[208] M. Šilhavý. Fractional vector analysis based on invariance requirements (critique of coordinate approaches). *Contin. Mech. Thermodyn.*, 32(1):207–228, 2020.

[209] K. Weierstrass. Über das sogenannte Dirichlet'sche Princip. In *Ausgewählte Kapitel aus der Funktionenlehre*, pages 184–189. Vieweg+Teubner Verlag, Wiesbaden, 1988.

[210] W. P. Ziemer. *Weakly differentiable functions*, volume 120 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989. Sobolev spaces and functions of bounded variation.