

# Compactness results for a Dirichlet energy of nonlocal gradient with applications

Zhaolong Han<sup>1</sup>  | Tadele Mengesha<sup>2</sup> | Xiaochuan Tian<sup>1</sup> 

<sup>1</sup>Department of Mathematics, University of California, San Diego, California, USA

<sup>2</sup>Department of Mathematics, University of Tennessee, Knoxville, Tennessee, USA

## Correspondence

Zhaolong Han, Department of Mathematics, University of California, San Diego, CA 92093, USA.

Email: [zhhan@ucsd.edu](mailto:zhhan@ucsd.edu)

## Funding information

National Science Foundation, Grant/Award

Numbers: DMS-2111608, DMS-2240180,

DMS-2206252

## Abstract

We prove two compactness results for function spaces with finite Dirichlet energy of half-space nonlocal gradients. In each of these results, we provide sufficient conditions on a sequence of kernel functions that guarantee the asymptotic compact embedding of the associated nonlocal function spaces into the class of square-integrable functions. Moreover, we will demonstrate that the sequence of nonlocal function spaces converges in an appropriate sense to a limiting function space. As an application, we prove uniform Poincaré-type inequalities for sequence of half-space gradient operators. We also apply the compactness result to demonstrate the convergence of appropriately parameterized nonlocal heterogeneous anisotropic diffusion problems. We will construct asymptotically compatible schemes for these type of problems. Another application concerns the convergence and robust discretization of a nonlocal optimal control problem.

## KEYWORDS

asymptotically compatible schemes, Bourgain–Brezis–Mironeanu, compactness, half-space gradient, nonlocal Dirichlet energy, nonlocal gradient, optimal control problem, uniform Poincaré inequality

## 1 | INTRODUCTION

In this article, we revisit the half-space nonlocal gradient operator studied in [20] defined in the principal value sense as

$$\mathcal{G}_w^\nu u(\mathbf{x}) = \lim_{\epsilon \rightarrow 0} \int_{\mathcal{H}_\nu \setminus B_\epsilon(\mathbf{0})} \frac{z}{|z|} (u(\mathbf{x} + z) - u(\mathbf{x})) w(z) dz, \quad \mathbf{x} \in \mathbb{R}^d,$$

Dedicated to Professor Qiang Du on the occasion of his 60th birthday.

where for  $d \geq 1$ ,  $u : \mathbb{R}^d \rightarrow \mathbb{R}$  is a measurable function,  $\nu \in \mathbb{R}^d$  is a fixed unit vector,  $\mathcal{H}_\nu$  is the half-space  $\mathcal{H}_\nu = \{z \in \mathbb{R}^d : z \cdot \nu \geq 0\}$ , and  $w$  is a nonnegative kernel function that satisfies appropriate conditions specified later. Additionally, inspired by the integration by parts formula, a distributional nonlocal gradient  $\mathfrak{G}_w^\nu$  is introduced (see eq. (7)), and the related nonlocal Sobolev-type space  $S_w^\nu(\Omega)$ , defined by

$$S_w^\nu(\Omega) = \{u \in L^2(\mathbb{R}^d) : \mathfrak{G}_w^\nu u \in L^2(\mathbb{R}^d; \mathbb{R}^d), u = 0 \text{ a.e. in } \mathbb{R}^d \setminus \Omega\},$$

for an open domain  $\Omega \subset \mathbb{R}^d$  is studied in [20].

The study of nonlocal operators is largely motivated by peridynamics, introduced by Silling [32], which has been effective in modeling materials undergoing large strains and fractures. Within the peridynamics theory, a specific category of material models known as correspondence models employs nonlocal gradients and classical stress-strain relationships to calculate the forces between particles [19, 34, 38, 39]. Despite their effectiveness, correspondence models are challenged by material instability issues [33], which, however, can be addressed through the use of half-space gradient operators. The exploration of such operators began with the work of Lee and Du [21] in the context of periodic domains. Subsequent developments, carried out by Han and Tian [20], extended the study to general domains in  $\mathbb{R}^d$ , where they showed a nonlocal Poincaré inequality on  $S_w^\nu(\Omega)$  with  $\Omega$  being a bounded domain with continuous boundary. The well-posedness of nonlocal diffusion problems and peridynamics correspondence models using half-space nonlocal operators is therefore implied via the application Lax–Milgram theorem. For additional studies related to the nonlocal gradient, please refer to [4–6, 11, 13, 24, 26, 30, 31] and the references therein.

This work primarily aims to analyze the convergence of variational problems associated with the nonlocal gradient  $\mathfrak{G}_w^\nu$  and the energy space  $S_w^\nu(\Omega)$ . In particular, the focus is on understanding the convergence of solutions when the variational problems are associated with a sequence of kernels  $w_n$  that approaches a specific limit as  $n$  approaches infinity. The key for showing such convergence relies on establishing compactness results in the limit as  $n$  approaches infinity in the spirit of Bourgain, Brezis and Mironescu [7]. In this work, we prove two types of compactness results, each associated with a particular type of kernel sequence convergence. One type of kernel convergence, as detailed in Assumption 3.9, is fundamentally characterized by the convergence of  $\min(1, |z|)w_n(z)$  to a measure with atomic mass at the origin as  $n \rightarrow \infty$  (or as  $\delta \rightarrow 0$  in the context of Assumption 3.9). Such a result is crucial for establishing the convergence of solutions from nonlocal variational problems to their local counterparts as nonlocality diminishes. A different kind of kernel convergence presupposes  $\{w_n\}_n$  is a sequence of nonnegative radial kernels converges increasingly to a kernel  $w$  almost everywhere, that is,  $0 \leq w_n \nearrow w$  a.e., where  $w$  is nonintegrable near the origin (for detailed assumptions, refer to Assumption 3.2), resulting in nonlocal-to-local convergence. A key intermediate result for this form of compactness is the locally compact embedding of  $S_w^\nu(\mathbb{R}^d)$  into  $L^2(\mathbb{R}^d)$ , which extends the classical compact embedding of  $H^s$  ( $s \in (0, 1)$ ) into  $L^2$  by choosing  $w(z)$  as the Riesz fractional kernel  $C/|z|^{d+s}$  [5, 30, 31]. Furthermore, the compactness results yield Poincaré inequalities that remain uniform regardless of the size of  $n$ . The uniform Poincaré inequalities also serve as essential intermediary steps in showing the convergence of variational solutions.

To establish the compactness results, we rely on a variant of the Riesz–Kolmogorov–Fréchet compactness criterion [22]. Addressing nonlocal Dirichlet energies with nonlocal gradients presents more challenges compared to the “double-integral” type nonlocal energies studies in previous works, for example, [7, 23]. In recent studies [5, 10], compactness results for nonlocal energies with truncated Riesz fractional gradients are established where a nonlocal version of the fundamental theorem of calculus plays a pivotal role. However, none of the existing methodologies are suitable for addressing the challenges posed by nonlocal energies defined through a half-space gradient. In our study, to show

the Riesz–Kolmogorov–Fréchet compactness criterion, we employ Fourier analysis along with fine estimates (see Section 2) of the Fourier symbols for half-space nonlocal gradient operators.

The compactness results are applied to obtain convergence of parameterized nonlocal heterogeneous anisotropic diffusion problems, together with the asymptotic compatibility of their Galerkin approximations with appropriate finite dimensional subspaces. The concept of asymptotic compatibility is introduced by [35] and is aimed for robust numerical discretization of parameterized nonlocal variational problems as the nonlocal modeling parameter and the discretization parameter both go to an asymptotic limit. We summarize the main convergence theorem in [35] with a slight relaxation on the conditions, that is, the assumption that strong convergence of operators acting on a dense subset is replaced by the convergence of bilinear forms in a weak sense. It is worth noting that in the case of  $0 \leq w_n \nearrow w$  for a sufficiently singular kernel  $w$ , the Galerkin approximation to the nonlocal problem associated with  $w_n$  serves as a nonconforming discontinuous Galerkin (DG) scheme for the nonlocal problem tied to  $w$ , as discussed in [36]. Another application we discuss is an optimal control problem with nonlocal diffusion equation as the constraint. The approach we follow and the results we obtain for this part of the study parallel [25] an optimal control problem is analyzed where the constraint is the linear bond-based peridynamics model.

The article is organized as follows. In Section 2, we recall the notion and some basic properties of nonlocal half-space operators studied in [20]. Further, the nonlocal function spaces are reformulated using the Fourier symbol of the nonlocal gradient, with a crucial lower bound estimate for the symbol presented. The major compactness results are established in Section 3, utilizing the symbol estimates. Applications are demonstrated through the proof of convergence for nonlocal diffusion problems and the establishment of asymptotic compatibility in Section 4. Additionally, the study of the optimal control problem is undertaken in Section 5. Finally, we conclude in Section 6.

## 2 | REVISITING NONLOCAL HALF-SPACE OPERATORS AND ASSOCIATED FUNCTION SPACES

### 2.1 | Nonlocal half-space Operators

In this section, we revisit definitions of nonlocal half-space vector operators as introduced in [20] and present some new properties.

Let  $\mathbf{v} \in \mathbb{R}^d$  be a fixed unit vector. Denote by  $\chi_{\mathbf{v}}(z)$  the characteristic function of the half-space  $\mathcal{H}_{\mathbf{v}} := \{z \in \mathbb{R}^d : z \cdot \mathbf{v} \geq 0\}$  parameterized by the unit vector  $\mathbf{v}$ . Throughout this article, we assume that  $w$  satisfies the following conditions.

$$\begin{cases} w \in L^1_{\text{loc}}(\mathbb{R}^d \setminus \{\mathbf{0}\}), w \geq 0 \text{ a.e., } w \text{ is radial;} \\ M_w^1 := \int_{|z| \leq 1} |z| w(z) dz \in (0, \infty) \quad \text{and} \quad M_w^2 := \int_{|z| > 1} w(z) dz < \infty. \end{cases} \quad (1)$$

We notice that by assumption (1) there exists  $\epsilon_0 > 0$  such that

$$0 < \int_{\epsilon_0 < |z| < 1} w(z) dz < \infty. \quad (2)$$

We recall that the nonlocal half-space gradient, divergence, and curl operators are defined in the principal value sense (see [20, Definition 2.1]). For smooth functions  $u \in C_c^1(\mathbb{R}^d)$  and  $\mathbf{v} \in C_c^1(\mathbb{R}^d; \mathbb{R}^d)$ , the operators are meaningful pointwise and are given as ([20, Lemma 2.1])

$$\mathcal{G}_w^{\mathbf{v}} u(\mathbf{x}) = \int_{\mathbb{R}^d} \chi_{\mathbf{v}}(\mathbf{y} - \mathbf{x}) \frac{\mathbf{y} - \mathbf{x}}{|\mathbf{y} - \mathbf{x}|} (u(\mathbf{y}) - u(\mathbf{x})) w(\mathbf{y} - \mathbf{x}) d\mathbf{y}, \quad \mathbf{x} \in \mathbb{R}^d, \quad (3)$$

$$D_w^{\mathbf{v}} \mathbf{v}(\mathbf{x}) = \int_{\mathbb{R}^d} \chi_{\mathbf{v}}(\mathbf{y} - \mathbf{x}) \frac{\mathbf{y} - \mathbf{x}}{|\mathbf{y} - \mathbf{x}|} \cdot (\mathbf{v}(\mathbf{y}) - \mathbf{v}(\mathbf{x})) w(\mathbf{y} - \mathbf{x}) d\mathbf{y}, \quad \mathbf{x} \in \mathbb{R}^d, \quad (4)$$

and for  $p \in [1, \infty]$ ,  $\mathcal{G}_w^\nu u \in L^p(\mathbb{R}^d; \mathbb{R}^d)$ ,  $\mathcal{D}_w^\nu v \in L^p(\mathbb{R}^d)$  with the estimate that for a constant  $C$  depending on  $d$  and  $p$ ,

$$\|\mathcal{G}_w^\nu u\|_{L^p(\mathbb{R}^d; \mathbb{R}^d)} \leq C(M_w^1 \|\nabla u\|_{L^p(\mathbb{R}^d; \mathbb{R}^d)} + M_w^2 \|u\|_{L^p(\mathbb{R}^d)}), \text{ and} \quad (5)$$

$$\|\mathcal{D}_w^\nu v\|_{L^p(\mathbb{R}^d)} \leq C(M_w^1 \|\nabla v\|_{L^p(\mathbb{R}^d; \mathbb{R}^{d \times d})} + M_w^2 \|v\|_{L^p(\mathbb{R}^d; \mathbb{R}^d)}). \quad (6)$$

In addition, the nonlocal gradient and nonlocal divergence are related via a nonlocal integration by parts formula

$$\int_{\mathbb{R}^d} \mathcal{G}_w^\nu u(x) \cdot v(x) dx = - \int_{\mathbb{R}^d} u(x) \mathcal{D}_w^{-\nu} v(x) dx,$$

for any  $u \in C_c^1(\mathbb{R}^d)$  and  $v \in C_c^1(\mathbb{R}^d; \mathbb{R}^d)$ . In fact, as shown in [20, Proposition 2.4], the above formula is valid for  $u \in L^1(\mathbb{R}^d)$  such that  $w(x-y)|u(x) - u(y)| \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$ . This formula motivates the introduction of the distributional nonlocal gradient operator  $\mathfrak{G}_w^\nu u \in (C_c^\infty(\mathbb{R}^d; \mathbb{R}^d))'$  defined for  $u \in L^p(\mathbb{R}^d)$ ,  $1 \leq p \leq \infty$ , as

$$\langle \mathfrak{G}_w^\nu u, \phi \rangle := - \int_{\mathbb{R}^d} u(x) \cdot \mathcal{D}_w^{-\nu} \phi(x) dx, \quad \forall \phi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d), \quad (7)$$

(see [20, Definition 2.2]).

It has also been demonstrated in [21] that for  $u \in C_c^\infty(\mathbb{R}^d)$ ,  $v \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$ , and  $\xi \in \mathbb{R}^d$

$$\mathcal{F}(\mathcal{G}_w^\nu u)(\xi) = \lambda_w^\nu(\xi) \hat{u}(\xi), \text{ and} \quad (8)$$

$$\mathcal{F}(\mathcal{D}_w^\nu v)(\xi) = \lambda_w^\nu(\xi)^T \hat{v}(\xi), \quad (9)$$

where the Fourier symbol  $\lambda_w^\nu(\xi)$  is given by

$$\lambda_w^\nu(\xi) := \int_{\mathbb{R}^d} \chi_\nu(z) \frac{z}{|z|} w(z) (e^{2\pi i \xi \cdot z} - 1) dz. \quad (10)$$

In the above, and hereafter,  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  represent the Fourier transform and the inverse Fourier transform, respectively. We will also use the  $\hat{\cdot}$ -notation for Fourier transform interchangeably. That is,  $\mathcal{F}(u) = \hat{u}$ .

By [20, Lemma 2.4], the Fourier symbol  $\lambda_w^\nu(\xi)$  is bounded by a linear function, that is,

$$|\lambda_w^\nu(\xi)| \leq 2\sqrt{2\pi} M_w^1 |\xi| + \sqrt{2} M_w^2, \quad \xi \in \mathbb{R}^d. \quad (11)$$

In fact, in [20, Lemma 5.1], the following estimates on the behavior of  $\lambda_w^\nu(\xi)$  near the origin and at infinity are proved. It is worth noting that in [20], it is assumed that the kernel does not vanish in a neighborhood of the origin. Nevertheless, the arguments presented in that work remain valid with this assumption replaced by  $w \neq 0 \in L^1_{\text{loc}}(\mathbb{R}^d \setminus \{\mathbf{0}\})$ , in particular  $w$  satisfying (1).

**Lemma 2.1.** *Let  $w$  satisfy (1).*

1. *For any  $N > 0$  there exists a constant  $C = C(N, w, d) > 0$  such that*

$$|\lambda_w^\nu(\xi)| \geq C \int_{\mathbb{R}^d} \min(1, w(x)) dx, \quad \forall |\xi| \geq N.$$

2. *Suppose that in addition  $w$  has finite first moment, that is,  $\int_{\mathbb{R}^d} |z| w(z) dz < \infty$ . Then there exist constants  $N_1 = N_1(w, d) \in (0, 1)$  and  $C_1 = C_1(w, d) > 0$  such that*

$$|\lambda_w^\nu(\xi)| \geq C_1 |\xi|, \quad \forall |\xi| \leq N_1.$$

For the compactness result we obtain below, we will need an improved lower bound of  $\lambda_w^\nu(\xi)$  for large  $\xi$  than provided in the first part of the above lemma. The following lemma has this improved lower bound whose proof will be postponed to Appendix A.1 since it is long and technical.

**Lemma 2.2.** Let  $w$  satisfy (1). Assume in addition that  $w$  is nonincreasing if  $d = 1$ . Then for any  $N > 0$  and  $\epsilon > 0$ , there exists a constant  $C = C(N\epsilon, d) > 0$  such that

$$|\lambda_w^\nu(\xi)| \geq |\Re(\lambda_w^\nu)(\xi)| \geq C \int_{|z| > \frac{N\epsilon}{|\xi|}} w(z) dz, \quad \forall |\xi| > N.$$

Here  $\Re(\lambda_w^\nu)$  stands for the real part of  $\lambda_w^\nu$ .

## 2.2 | Nonlocal function spaces via Fourier symbols

In this section, we introduce a function space via the Fourier transform formula of the nonlocal gradient operator defined in the previous section. We will show that this function space is a Hilbert space and that it is the same space as the space of function with square integrable distributional nonlocal gradient that was defined and thoroughly studied in [20].

Define the function space  $\mathcal{H}_w^\nu(\mathbb{R}^d)$  as

$$\mathcal{H}_w^\nu(\mathbb{R}^d) := \left\{ u \in L^2(\mathbb{R}^d) : |u|_{\mathcal{H}_w^\nu(\mathbb{R}^d)}^2 := \int_{\mathbb{R}^d} |\lambda_w^\nu(\xi) \hat{u}(\xi)|^2 d\xi < \infty \right\} \quad (12)$$

equipped with norm

$$\|u\|_{\mathcal{H}_w^\nu(\mathbb{R}^d)} := \left( \|u\|_{L^2(\mathbb{R}^d)}^2 + |u|_{\mathcal{H}_w^\nu(\mathbb{R}^d)}^2 \right)^{\frac{1}{2}}$$

It is direct to check that  $\mathcal{H}_w^\nu(\mathbb{R}^d)$  is a normed vector space. Moreover, with the inner product given by  $(u, v)_{\mathcal{H}_w^\nu(\mathbb{R}^d)} = (u, v)_{L^2(\mathbb{R}^d)} + (\lambda_w^\nu \hat{u}, \lambda_w^\nu \hat{v})_{L^2(\mathbb{R}^d; \mathbb{C}^d)}$  for any  $u, v \in \mathcal{H}_w^\nu(\mathbb{R}^d)$ ,  $\mathcal{H}_w^\nu(\mathbb{R}^d)$  is, in fact, a Hilbert space. Before we prove this statement we make the following remark.

Let  $\mathcal{S}(\mathbb{R}^d; \mathbb{C})$  be the space of Schwartz functions and  $\mathcal{S}'(\mathbb{R}^d; \mathbb{C}^d)$  be the space of tempered distribution. Since  $u \in L^2(\mathbb{R}^d)$ ,  $\hat{u} \in L^2(\mathbb{R}^d; \mathbb{C})$ . Moreover, we may define the product  $\lambda_w^\nu \hat{u}$  as a tempered distribution in the following way:

$$\langle \lambda_w^\nu \hat{u}, \phi \rangle := (\lambda_w^\nu \hat{u}, \phi)_{L^2(\mathbb{R}^d; \mathbb{C}^d)} = \left( \hat{u}, (\lambda_w^\nu)^T \phi \right)_{L^2(\mathbb{R}^d; \mathbb{C})}, \quad \forall \phi \in \mathcal{S}(\mathbb{R}^d; \mathbb{C}^d).$$

and use eq. (11) to show that  $\lambda_w^\nu \hat{u} \in \mathcal{S}'(\mathbb{R}^d; \mathbb{C}^d)$ . We will use this observation to prove the completeness of  $\mathcal{H}_w^\nu(\mathbb{R}^d)$ .

**Proposition 2.3.**  $\mathcal{H}_w^\nu(\mathbb{R}^d)$  is a Hilbert space.

*Proof.* Let  $\{u_n\}_{n=1}^\infty$  be a Cauchy sequence in  $\mathcal{H}_w^\nu(\mathbb{R}^d)$ . Then there exist  $u \in L^2(\mathbb{R}^d)$  and  $v \in L^2(\mathbb{R}^d; \mathbb{C}^d)$  such that  $u_n \rightarrow u$  and  $\lambda_w^\nu \hat{u}_n \rightarrow v$  in  $L^2(\mathbb{R}^d; \mathbb{C}^d)$ . Then for any  $\phi \in \mathcal{S}(\mathbb{R}^d; \mathbb{C}^d)$ ,

$$\begin{aligned} \langle \lambda_w^\nu \hat{u}, \phi \rangle &= \left( \hat{u}, (\lambda_w^\nu)^T \phi \right)_{L^2(\mathbb{R}^d; \mathbb{C})} = \left( u, \mathcal{F}^{-1} \left( (\lambda_w^\nu)^T \phi \right) \right)_{L^2(\mathbb{R}^d; \mathbb{C})} \\ &= \lim_{n \rightarrow \infty} \left( u_n, \mathcal{F}^{-1} \left( (\lambda_w^\nu)^T \phi \right) \right)_{L^2(\mathbb{R}^d; \mathbb{C})} = \lim_{n \rightarrow \infty} \langle \lambda_w^\nu \hat{u}_n, \phi \rangle. \end{aligned}$$

Note that the above equalities hold true since  $(\lambda_w^\nu)^T \phi \in L^2(\mathbb{R}^d; \mathbb{C})$  as a result of Equation (11) and  $\phi \in \mathcal{S}(\mathbb{R}^d; \mathbb{C}^d)$ . Thus,  $\lambda_w^\nu \hat{u}_n \rightarrow \lambda_w^\nu \hat{u}$  in  $\mathcal{S}'(\mathbb{R}^d; \mathbb{C}^d)$ . Uniqueness of limits in the sense of tempered distributions, we have  $\lambda_w^\nu \hat{u} = v \in L^2(\mathbb{R}^d)$ . That is,  $u \in \mathcal{H}_w^\nu(\mathbb{R}^d)$  with  $u_n \rightarrow u$  in  $\mathcal{H}_w^\nu(\mathbb{R}^d)$ . Thus  $\mathcal{H}_w^\nu(\mathbb{R}^d)$  is complete. ■

In [20], the nonlocal function space  $\mathcal{S}_w^\nu(\mathbb{R}^d)$  is defined as follows:

$$\mathcal{S}_w^\nu(\mathbb{R}^d) := \{u \in L^2(\mathbb{R}^d) : \mathfrak{G}_w^\nu u \in L^2(\mathbb{R}^d; \mathbb{R}^d)\}$$

with norm

$$\|u\|_{S_w^v(\mathbb{R}^d)} := \left( \|u\|_{L^2(\mathbb{R}^d)}^2 + |u|_{S_w^v(\mathbb{R}^d)}^2 \right)^{\frac{1}{2}},$$

where  $|u|_{S_w^v(\mathbb{R}^d)} := \|\mathfrak{G}_w^v u\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)}$ . For a given open domain  $\Omega \subset \mathbb{R}^d$ , we define the closed subspace

$$S_w^v(\Omega) = \{u \in S_w^v(\mathbb{R}^d) : u = 0 \text{ a.e. in } \mathbb{R}^d \setminus \Omega\}$$

that collects functions in  $S_w^v(\mathbb{R}^d)$  that vanish outside of  $\Omega$ . It is shown in [20] that  $S_w^v(\Omega)$  is a Hilbert space and that if  $\Omega$  has a continuous boundary  $C_c^\infty(\Omega)$  is dense in  $S_w^v(\Omega)$ . We now establish the equivalence of the space  $S_w^v(\mathbb{R}^d)$  with  $\mathcal{H}_w^v(\mathbb{R}^d)$ .

**Proposition 2.4.**  $S_w^v(\mathbb{R}^d) = \mathcal{H}_w^v(\mathbb{R}^d)$  with equal norms. More precisely, the identity map is an isometrical isomorphism between  $S_w^v(\mathbb{R}^d)$  and  $\mathcal{H}_w^v(\mathbb{R}^d)$ .

*Proof.* First, we show  $S_w^v(\mathbb{R}^d) \subseteq \mathcal{H}_w^v(\mathbb{R}^d)$ . It suffices to show that for  $u \in S_w^v(\mathbb{R}^d)$ ,  $\|\lambda_w^v \hat{u}\|_{L^2(\mathbb{R}^d; \mathbb{C}^d)} = \|\mathfrak{G}_w^v u\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} < \infty$ . Since  $C_c^\infty(\mathbb{R}^d)$  is dense in  $S_w^v(\mathbb{R}^d)$ , there exists  $\{u_n\}_{n=1}^\infty \subset C_c^\infty(\mathbb{R}^d)$  such that  $u_n \rightarrow u$  in  $L^2(\mathbb{R}^d)$  and  $\mathfrak{G}_w^v u_n \rightarrow \mathfrak{G}_w^v u$  in  $L^2(\mathbb{R}^d)$ . Since Fourier transform is an isomorphism on  $L^2(\mathbb{R}^d)$ ,  $\lambda_w^v \hat{u}_n = \mathcal{F}(\mathfrak{G}_w^v u_n)$  converges to  $\widehat{\mathfrak{G}_w^v u}$  in  $L^2(\mathbb{R}^d; \mathbb{C}^d)$ . Similar to the proof of Proposition 2.3, one has  $\lambda_w^v \hat{u}_n \rightarrow \lambda_w^v \hat{u}$  in  $\mathcal{S}'(\mathbb{R}^d; \mathbb{C}^d)$ . Then we conclude that  $\lambda_w^v \hat{u} = \widehat{\mathfrak{G}_w^v u} \in L^2(\mathbb{R}^d; \mathbb{C}^d)$  and  $\|\lambda_w^v \hat{u}\|_{L^2(\mathbb{R}^d; \mathbb{C}^d)} = \|\mathfrak{G}_w^v u\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} < \infty$ .

Second, we show that  $\mathcal{H}_w^v(\mathbb{R}^d) \subseteq S_w^v(\mathbb{R}^d)$ . It suffices to show that for any  $u \in \mathcal{H}_w^v(\mathbb{R}^d)$ ,  $\mathfrak{G}_w^v u \in L^2(\mathbb{R}^d; \mathbb{R}^d)$ . By definition,  $\mathfrak{G}_w^v u \in (C_c^\infty(\mathbb{R}^d; \mathbb{R}^d))'$  and for any  $\phi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$ ,

$$\langle \mathfrak{G}_w^v u, \phi \rangle = -(u, \mathcal{D}_w^{-v} \phi)_{L^2(\mathbb{R}^d)} = \left( \hat{u}, \overline{\lambda_w^v}^T \hat{\phi} \right)_{L^2(\mathbb{R}^d; \mathbb{C})},$$

where we used  $\mathcal{F}(\mathcal{D}_w^{-v} \phi) = -\overline{\lambda_w^v}^T \hat{\phi}$  by Equations (9) and (10). On the other hand, since  $u \in \mathcal{H}_w^v(\mathbb{R}^d)$ ,  $\lambda_w^v \hat{u} \in L^2(\mathbb{R}^d; \mathbb{C}^d)$  and thus  $\mathcal{F}^{-1}(\lambda_w^v \hat{u}) \in L^2(\mathbb{R}^d; \mathbb{C}^d)$ . For any  $\phi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$ ,

$$(\mathcal{F}^{-1}(\lambda_w^v \hat{u}), \phi)_{L^2(\mathbb{R}^d; \mathbb{C}^d)} = (\lambda_w^v \hat{u}, \hat{\phi})_{L^2(\mathbb{R}^d; \mathbb{C}^d)} = \left( \hat{u}, \overline{\lambda_w^v}^T \hat{\phi} \right)_{L^2(\mathbb{R}^d; \mathbb{C})}.$$

Therefore,  $\mathfrak{G}_w^v u = \mathcal{F}^{-1}(\lambda_w^v \hat{u})$  in the sense of distribution. We note that  $\mathcal{F}^{-1}(\lambda_w^v \hat{u})$  is real-valued. This follows from the argument that for any  $\phi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$ ,

$$(\mathcal{F}^{-1}(\lambda_w^v \hat{u}), \phi)_{L^2(\mathbb{R}^d; \mathbb{C}^d)} = \left( \hat{u}, \overline{\lambda_w^v}^T \hat{\phi} \right)_{L^2(\mathbb{R}^d; \mathbb{C})} = -(u, \mathcal{D}_w^{-v} \phi)_{L^2(\mathbb{R}^d)} \in \mathbb{R},$$

and so  $\mathcal{F}^{-1}(\lambda_w^v \hat{u}) \in L^2(\mathbb{R}^d; \mathbb{R}^d)$ , thus  $\mathfrak{G}_w^v u \in L^2(\mathbb{R}^d; \mathbb{R}^d)$  and the proof is complete.  $\blacksquare$

We remark that, on the one hand, in the event that the kernel  $w$  satisfying (1) also belongs to  $L^1(\mathbb{R}^d)$ , we have that  $\|\lambda_w^v\|_{L^\infty(\mathbb{R}^d; \mathbb{R}^d)} \leq 2\|w\|_{L^1(\mathbb{R}^d)}$ . In this case, the function space  $S_w^v(\mathbb{R}^d)$  coincides with  $L^2(\mathbb{R}^d)$  with norm estimate

$$\|u\|_{S_w^v(\mathbb{R}^d)}^2 = \int_{\mathbb{R}^d} (1 + |\lambda_w^v(\xi)|^2) |\hat{u}(\xi)|^2 d\xi \leq (1 + 4\|w\|_{L^1(\mathbb{R}^d)}^2) \|u\|_{L^2(\mathbb{R}^d)}^2, \quad \forall u \in S_w^v(\mathbb{R}^d).$$

On the other hand, when the kernel  $w$  satisfying (1) is not integrable, that is,  $\int_{\mathbb{R}^d} w(z) dz = +\infty$ , or equivalently  $\int_{|z|<1} w(z) dz = +\infty$ , the space  $S_w^v(\mathbb{R}^d)$  is properly contained in  $L^2(\mathbb{R}^d)$ . In fact, as one of our main results of the next section, we will show that  $S_w^v(\mathbb{R}^d)$  is locally compactly contained in  $L^2(\mathbb{R}^d)$ . An example of such type of kernel is  $w_s(z) = |z|^{-d-s}$  for  $s \in (0, 1)$ . In this case,  $C_1(s)|\xi|^s \leq |\lambda_{w_s}^v(\xi)| \leq C_2(s)|\xi|^s$ , where  $C_1(s)(1-s) \rightarrow 1$  and  $C_2(s)(1-s) \rightarrow 2\pi$  as  $s \nearrow 1$ , leading to the conclusion that  $S_{w_s}^v(\mathbb{R}^d) = \mathcal{H}_{w_s}^v(\mathbb{R}^d) = H^s(\mathbb{R}^d)$ , the fractional Sobolev space.

As we will see in the sections below, working on function spaces with compactly supported kernels leads to simplifications. To that end, given a nonintegrable kernel  $w$ , we will show that we can have an equivalent characterization of the function space  $S_w^v(\mathbb{R}^d)$  using only the cut-off  $w^c(z) := w(z)\chi_{B_1(0)}(z)$  of the kernel  $w$ .

**Lemma 2.5.** *Let  $w(z)$  be a kernel satisfying (1). Denote  $w^c(z) := w(z)\chi_{B_1(0)}(z)$  as the compactly supported kernel obtained cutting of the tail of  $w$ . Then there exist positive constants  $c_1$  and  $c_2$  depending on  $M_w^2 = \int_{|z|>1} w(z)dz$  such that*

$$c_1 \|u\|_{S_w^v(\mathbb{R}^d)} \leq \|u\|_{S_{w^c}^v(\mathbb{R}^d)} \leq c_2 \|u\|_{S_w^v(\mathbb{R}^d)}, \quad \forall u \in S_w^v(\mathbb{R}^d). \quad (13)$$

*Proof.* We begin noting that  $w^c$  satisfies (1), and so from Proposition 2.4 we have that

$$\|u\|_{S_w^v(\mathbb{R}^d)}^2 = \int_{\mathbb{R}^d} (1 + |\lambda_w^v(\xi)|^2) |\hat{u}(\xi)|^2 d\xi, \quad \forall u \in S_w^v(\mathbb{R}^d).$$

A similar equality also holds for functions in  $S_{w^c}^v(\mathbb{R}^d)$  with  $w$  replaced by  $w^c$ . Now to compare  $\|u\|_{S_w^v(\mathbb{R}^d)}$  and  $\|u\|_{S_{w^c}^v(\mathbb{R}^d)}$ , it suffices to compare  $\lambda_w^v(\xi)$  and  $\lambda_{w^c}^v(\xi)$ . By (11),

$$|\lambda_w^v(\xi) - \lambda_{w^c}^v(\xi)| = |\lambda_{w-w^c}^v(\xi)| \leq 2M_w^2, \quad \forall \xi \in \mathbb{R}^d. \quad (14)$$

Then

$$|\lambda_{w^c}^v(\xi)|^2 \leq (|\lambda_w^v(\xi)| + 2M_w^2)^2 \leq 2(|\lambda_w^v(\xi)|^2 + 4(M_w^2)^2)$$

and

$$\|u\|_{S_{w^c}^v(\mathbb{R}^d)}^2 \leq \max\{2, 1 + 8(M_w^2)^2\} \|u\|_{S_w^v(\mathbb{R}^d)}^2, \quad \forall u \in S_w^v(\mathbb{R}^d).$$

Therefore,  $c_2 = (\max\{2, 1 + 8(M_w^2)^2\})^{\frac{1}{2}}$ . Changing the role of  $w$  and  $w^c$ , we get the other inequality with  $c_1 = c_2^{-1}$ .  $\blacksquare$

### 3 | COMPACTNESS RESULTS

In this section, we prove two compactness results. The first one states that for a fixed nonintegrable kernel  $w$  satisfying (1), the function space  $S_w^v(\mathbb{R}^d)$  is compactly contained in  $L^2(\mathbb{R}^d)$  with respect to the  $L_{\text{loc}}^2$ -topology. This, in particular, implies that for such  $w$ , and any bounded domain with continuous boundary  $S_w^v(\Omega)$  is compactly contained in  $L^2(\Omega)$ . Other variants of this compactness results will also be proved. The second result is related to the sequence of parameterized function spaces  $\{S_{w_n}^v(\mathbb{R}^d)\}_n$  associated with a sequence of kernel  $\{w_n\}_n$  where  $\{\min(1, |z|)w_n(z)\}_n$  is concentrating around 0 and behaving like a Dirac delta-sequence in the appropriate sense. We will give the precise statement later, but the result essentially states that a sequence  $\{u_n\}$  of  $L^2$  functions with the property that  $\sup_n \|u_n\|_{S_{w_n}^v(\mathbb{R}^d)} < \infty$ , has a compact closure in  $L^2(\Omega)$ , for any bounded domain  $\Omega$  of  $\mathbb{R}^d$ . These compactness results are proved with the aid of a variant of Riesz–Kolmogorov–Fréchet compactness criterion proved in [7], see also [22, Lemma 5.4].

To state the criterion, we observe that for  $\mathbf{P} \in L^1(\mathbb{R}^d; \mathbb{R}^d)$  and  $f \in L^p(\mathbb{R}^d)$ ,  $p \geq 1$ , we understand the convolution  $\mathbf{P} * f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  as  $(\mathbf{P} * f)_j(\mathbf{x}) := (\mathbf{P}_j * f)(\mathbf{x})$ . Then  $\mathbf{P} * f \in L^p(\mathbb{R}^d; \mathbb{R}^d)$ .

**Lemma 3.1.** *Suppose that  $\mathbf{P} \in L^1(\mathbb{R}^d; \mathbb{R}^d)$  and  $\mathbf{p} := \int_{\mathbb{R}^d} \mathbf{P}(\mathbf{x})d\mathbf{x} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ . Define  $\mathbf{P}_\tau(\mathbf{x}) := \tau^{-d} \mathbf{P}(\mathbf{x}/\tau)$  for  $\tau > 0$ . For any  $1 \leq p < \infty$ , if  $\{f_n\}_{n=1}^\infty$  is a bounded sequence of*

functions in  $L^p(\mathbb{R}^d)$  and

$$\lim_{\tau \rightarrow 0^+} \limsup_{n \rightarrow \infty} \|P_\tau * f_n - Pf_n\|_{L^p(\mathbb{R}^d; \mathbb{R}^d)} = 0, \quad (15)$$

then  $\{f_n\}$  has a compact closure in  $L^p(\Omega)$  for any bounded domain  $\Omega \subset \mathbb{R}^d$ .

### 3.1 | Compact embedding of function spaces associated with nonintegrable kernel

In this subsection we show a local compact embedding result of  $S_w^v(\mathbb{R}^d)$  into  $L^2(\mathbb{R}^d)$  when  $w$  is a nonintegrable kernel. To be precise, we state the assumption on  $w$  as follows which will remain in force throughout this subsection.

**Assumption 3.2.** Assume that  $w$  satisfies (1) and the following conditions:

1.  $\int_{\mathbb{R}^d} w(z) dz = +\infty$ , or equivalently  $\int_{|z|<1} w(z) dz = +\infty$ .
2. If  $d = 1$ , then we assume in addition that  $w$  is nonincreasing.

We use  $\bar{w}$  as the radial representation of  $w$ , that is,  $\bar{w} : [0, \infty) \rightarrow [0, \infty)$  such that  $\bar{w}(|x|) = w(x)$  for all  $x \in \mathbb{R}^d$ . Note that the above assumptions hold for  $w(z) = |z|^{-d-s} \chi_{B_1(0)}(z)$  with  $s \in (0, 1)$ , as well as for  $w(z) = -\ln(|z|)|z|^{-d} \chi_{B_1(0)}(z)$ . The main result we prove in this subsection is the following.

**Theorem 3.3.** Let  $w$  satisfy Assumption 3.2. Let  $\{u_n\}_{n=1}^\infty \subset S_w^v(\mathbb{R}^d)$  be a bounded sequence in  $L^2(\mathbb{R}^d)$  with bounded nonlocal gradient seminorm

$$\sup_n \|\mathfrak{G}_w^v u_n\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} = B < \infty, \quad (16)$$

then for any bounded domain  $\Omega \subset \mathbb{R}^d$ ,  $\{u_n|_\Omega\}_{n=1}^\infty$  is precompact in  $L^2(\Omega)$ . In other words,  $S_w^v(\mathbb{R}^d)$  is locally compactly embedded in  $L^2(\mathbb{R}^d)$ .

As we mentioned earlier, we prove the theorem using the criterion given in Lemma 3.1. To that end, we set  $P(x) := \chi_v(x) \chi_{B_1(0)}(x) x / |x|$ . By taking the Fourier transform, one can verify that for  $u \in L^2(\mathbb{R}^d)$  and  $\xi \in \mathbb{R}^d$ ,

$$\mathcal{F}(P_\tau * u - P u)(\xi) = \eta_\tau(\xi) \hat{u}(\xi),$$

where

$$\eta_\tau(\xi) := \int_{\mathbb{R}^d} \frac{1}{\tau^d} \chi_v(z) \chi_{B_\tau(0)}(z) \frac{z}{|z|} (e^{-2\pi i z \cdot \xi} - 1) dz, \quad \xi \in \mathbb{R}^d. \quad (17)$$

By a change of variable one implies that

$$\eta_\tau(\xi) = \int_{\mathbb{R}^d} \chi_v(z) \chi_{B_1(0)}(z) \frac{z}{|z|} (e^{-2\pi i z \cdot \xi} - 1) dz, \quad \xi \in \mathbb{R}^d.$$

It then follows from a calculation similar to [21, Theorem 2.4] that

$$|\eta_\tau(\xi)| \leq V_d \min\{\sqrt{2\pi\tau}|\xi|, 1\}, \quad (18)$$

where  $V_d$  is the volume of the  $d$ -dimensional unit ball  $B_1^d(\mathbf{0})$ .

After noticing that the boundedness condition (16) of Theorem 3.3 can be expressed using the Fourier symbols as  $\lambda_{w^c}^v(\xi)$ , our next goal is to estimate  $|\eta_\tau(\xi)|$  in terms of  $|\lambda_{w^c}^v(\xi)|$ . To that end, we have the following crucial estimating device.

**Remark 3.4.** We reiterate that for any  $w$  satisfying (1), its cut-off  $w^c = w \chi_{B_1(0)}$  also satisfies (1) and as a consequence of Lemma 2.1(2), there exist constants  $N_1 = N_1(w, d) > 0$

and  $C_1 = C_1(w, d) > 0$  such that

$$|\lambda_{w^c}^v(\xi)| \geq C_1 |\xi|, \quad \forall |\xi| \leq N_1.$$

We recall from (2) that there exists  $\epsilon_0 > 0$  such that

$$0 < \int_{\epsilon_0 < |z| < 1} w(z) dz < \infty,$$

and thus, taking  $N = N_1$  and  $\epsilon = \epsilon_0$  in Lemma 2.2, it follows that there exists a constant  $C_2 = C_2(N_1 \epsilon_0, d) > 0$  such that

$$|\lambda_{w^c}^v(\xi)| \geq C_2 \int_{\frac{N_1 \epsilon_0}{|\xi|} < |z| < 1} w(z) dz > 0, \quad \forall |\xi| > N_1.$$

**Lemma 3.5.** Suppose that  $\epsilon_0 > 0$  as in (2). There exists  $C = C(w, d)$  such that for  $0 < \tau < \epsilon_0$ ,

$$|\eta_\tau(\xi)| \leq Cg(\tau) |\lambda_{w^c}^v(\xi)|, \quad \forall \xi \in \mathbb{R}^d,$$

where  $g : (0, \epsilon_0) \rightarrow \mathbb{R}_+$  satisfies

$$\lim_{\tau \rightarrow 0^+} g(\tau) = 0.$$

*Proof.* Fix  $0 < \tau < \epsilon_0$ , we prove the inequality by discussing  $0 \leq |\xi| \leq N_1$ ,  $N_1 \leq |\xi| \leq N_1 \epsilon_0 / \tau$  and  $|\xi| > N_1 \epsilon_0 / \tau$ , where  $\epsilon_0$  and  $N_1$  are as in Remark 3.4.

We begin from the estimate given in (18) where for some  $c_\eta = c_\eta(d) > 0$  we have

$$|\eta_\tau(\xi)| \leq c_\eta(d) \min\{\tau |\xi|, 1\}, \quad \text{for all } \xi \in \mathbb{R}^d.$$

Then on the one hand for  $|\xi| \leq N_1$ , one has

$$|\eta_\tau(\xi)| \leq c_\eta \tau |\xi| \leq \frac{c_\eta}{C_1} \tau |\lambda_{w^c}^v(\xi)|.$$

On the other hand, for  $|\xi| > N_1 \epsilon_0 / \tau > N_1$  by Remark 3.4 one has

$$\begin{aligned} |\eta_\tau(\xi)| &\leq c_\eta \leq \frac{c_\eta}{C_2} \left( \int_{\frac{N_1 \epsilon_0}{|\xi|} < |z| < 1} w(z) dz \right)^{-1} |\lambda_{w^c}^v(\xi)| \\ &\leq \frac{c_\eta}{C_2} \left( \int_{\tau < |z| < 1} w(z) dz \right)^{-1} |\lambda_{w^c}^v(\xi)|, \end{aligned}$$

where we used nonnegativity of  $w$  in the last step. Finally, for  $N_1 \leq |\xi| \leq N_1 \epsilon_0 / \tau$ , define

$$f(\tau) := \sup_{|\xi| \in [N_1, \frac{N_1 \epsilon_0}{\tau}]} \frac{\tau |\xi|}{\int_{\frac{N_1 \epsilon_0}{|\xi|} < |z| < 1} w(z) dz} = \sup_{s \in [N_1 \tau, N_1 \epsilon_0]} \frac{s}{\int_{\frac{N_1 \epsilon_0 \tau}{s} < |z| < 1} w(z) dz},$$

then again by Remark 3.4 we know for  $N_1 \leq |\xi| \leq N_1 \epsilon_0 / \tau$ ,

$$|\eta_\tau(\xi)| \leq c_\eta \tau |\xi| \leq \frac{c_\eta}{C_2} f(\tau) |\lambda_{w^c}^v(\xi)|.$$

It remains to show  $f(\tau) \rightarrow 0$  as  $\tau \rightarrow 0^+$ . Suppose that there exist a constant  $\alpha > 0$  and a sequence  $\{\tau_n\}_{n=1}^\infty$  such that  $\tau_n \rightarrow 0$  as  $n \rightarrow \infty$  and  $f(\tau_n) > \alpha$  for any  $n \geq 1$ . Then for each  $n \geq 1$  there exists  $s_n \in [N_1 \tau_n, N_1 \epsilon_0]$  such that

$$\frac{s_n}{\int_{\frac{N_1 \epsilon_0 \tau_n}{s_n} < |z| < 1} w(z) dz} > \alpha.$$

We claim that  $\tau_n/s_n \rightarrow 0$  as  $n \rightarrow \infty$ . Suppose this does not hold. Then there exist  $0 < B \leq 1/N_1$  and subsequences  $\{\tau_{n_k}\}_{k=1}^\infty$  and  $\{s_{n_k}\}_{k=1}^\infty$  such that  $B \leq \tau_{n_k}/s_{n_k} \leq 1/N_1$ . Then since  $N_1 \tau_{n_k}/s_{n_k} \leq 1$ , we have

$$\alpha < \frac{s_{n_k}}{\int_{\frac{N_1 \epsilon_0 \tau_{n_k}}{s_{n_k}} < |z| < 1} w(z) dz} \leq \frac{\tau_{n_k}}{B \int_{\epsilon_0 < |z| < 1} w(z) dz} \rightarrow 0, \quad k \rightarrow +\infty.$$

This is a contradiction. Therefore, the claim holds. Pick a decreasing subsequence of  $\{\tau_n/s_n\}_{n=1}^\infty$  and still denote it by  $\{\tau_n/s_n\}_{n=1}^\infty$ . Then since  $\int_{|z| < 1} w(z) dz = \infty$ , we have that as  $n \rightarrow \infty$ ,

$$\alpha < \frac{s_n}{\int_{\frac{N_1 \epsilon_0 \tau_n}{s_n} < |z| < 1} w(z) dz} \leq \frac{N_1 \epsilon_0}{\int_{\frac{N_1 \epsilon_0 \tau_n}{s_n} < |z| < 1} w(z) dz} \rightarrow \frac{N_1 \epsilon_0}{\int_{|z| < 1} w(z) dz} = 0,$$

where we used monotone convergence theorem. This is a contradiction. Therefore,  $f(\tau) \rightarrow 0$  as  $\tau \rightarrow 0^+$  and

$$g(\tau) := \max \left\{ \tau, f(\tau), \left( \int_{\tau < |z| < 1} w(z) dz \right)^{-1} \right\} \rightarrow 0, \quad \text{as } \tau \rightarrow 0^+.$$

Hence, the proof is complete with  $C = C(w, d) = c_\eta / \min\{C_1, C_2\}$ . ■

We are now ready to give the proof of the main result.

*Proof of Theorem 3.3.* Note that by Proposition 2.4,

$$\sup_n \|\lambda_w^\nu \hat{u}_n\|_{L^2(\mathbb{R}^d; \mathbb{C}^d)} = \sup_n \|\mathfrak{G}_w^\nu u_n\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} = B < \infty.$$

By Lemma 2.5, we may use the cut-off kernel  $w^c$  and there exists a constant  $B'$  depending on  $B$ ,  $M_w^2$  and  $\sup_n \|u_n\|_{L^2(\mathbb{R}^d)}$  such that

$$\sup_n \|\lambda_{w^c}^\nu \hat{u}_n\|_{L^2(\mathbb{R}^d; \mathbb{C}^d)} \leq B' < \infty.$$

We use the compactness criterion Lemma 3.1 to show that  $\{u_n\}_{n=1}^\infty$  has a compact closure in  $L^2(\Omega)$  for bounded domain  $\Omega \subset \mathbb{R}^d$ . Choose  $P(x) := \chi_\nu(x) \chi_{B_1(0)}(x) x / |x|$ . To that end, it suffices to show that

$$\lim_{\tau \rightarrow 0^+} \limsup_{n \rightarrow \infty} \|P_\tau * u_n - p u_n\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} = 0 \quad (19)$$

But recall that

$$\|P_\tau * u_n - p u_n\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} = \|\mathcal{F}(P_\tau * u_n - p u_n)\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} = \|\eta_\tau \hat{u}_n\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)}$$

where  $\eta_\tau$  is given by (17). By Lemma 3.5, there exists  $C = C(w, d)$  such that for  $0 < \tau < \epsilon_0$ ,

$$|\eta_\tau(\xi)| \leq Cg(\tau) |\lambda_{w^c}^\nu(\xi)|, \quad \forall \xi \in \mathbb{R}^d,$$

where  $g : (0, \epsilon_0) \rightarrow \mathbb{R}_+$  satisfies  $\lim_{\tau \rightarrow 0^+} g(\tau) = 0$ . Then for any  $n \geq 1$ ,

$$\|\eta_\tau \hat{u}_n\|_{L^2(\mathbb{R}^d; \mathbb{C}^d)} \leq Cg(\tau) \|\lambda_{w^c}^\nu \hat{u}_n\|_{L^2(\mathbb{R}^d; \mathbb{C}^d)} \leq B' Cg(\tau) \rightarrow 0, \quad \tau \rightarrow 0^+.$$

Hence, (19) holds and that completes the proof of the theorem. ■

An important variant of the above compactness result holds true when applied to a sequence of function spaces  $S_{w_n}^\nu(\mathbb{R}^d)$  corresponding to the radial kernels  $w_n(z)$  satisfying  $0 \leq w_n(z) \nearrow w(z)$  a.e.  $z$ .

We will present two classes of examples that satisfy this convergence property. First, take  $w_n(z) := \min\{n, w(z)\}$  for  $n \in \mathbb{N}_+$ . For each  $n$ ,  $w_n \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$  and  $w_n(z) \nearrow w(z)$  for all  $z \in \mathbb{R}^d$ . Note that  $S_{w_n}^\nu(\mathbb{R}^d)$  coincides with  $L^2(\mathbb{R}^d)$  for each  $n$ , while  $S_w^\nu(\mathbb{R}^d)$  is locally compactly contained in  $L^2(\mathbb{R}^d)$ . For another example, we take  $s \in (0, 1)$  and consider  $w_s(z) = |z|^{-d-s} \chi_{B_1(0)}(z)$ . For a sequence of numbers  $0 \leq s_n \nearrow s$  we know  $w_{s_n}(z) := |z|^{-d-s_n} \chi_{B_1(0)}(z) \nearrow w_s(z)$  for all  $z \in \mathbb{R}^d$ , and that  $S_{w_{s_n}}^\nu(\mathbb{R}^d) = H^{s_n}(\mathbb{R}^d)$ , and  $S_{w_s}^\nu(\mathbb{R}^d) = H^s(\mathbb{R}^d)$ , which is compactly contained in  $L^2(\mathbb{R}^d)$ .

**Theorem 3.6.** *Let  $w$  satisfy Assumption 3.2. Suppose that  $\{w_n\}_{n=1}^\infty$  is a sequence of radial kernels satisfying  $0 \leq w_n(z) \nearrow w(z)$  for almost every  $z \in \mathbb{R}^d$  and  $\{u_n\}_{n=1}^\infty \subset L^2(\mathbb{R}^d)$  is a bounded sequence with*

$$\sup_n \|\mathfrak{G}_{w_n}^\nu u_n\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} = B < \infty,$$

*then for any bounded domain  $\Omega \subset \mathbb{R}^d$ ,  $\{u_n|_\Omega\}_{n=1}^\infty$  is precompact in  $L^2(\Omega)$ . Moreover, if in addition  $u_n \in S_{w_n}^\nu(\Omega)$  for every  $n$ , that is,  $u_n = 0$  in  $\Omega^c$ , then for any limit point  $u$ , the zero-extension  $\tilde{u}$  of  $u$  outside  $\Omega$  satisfies  $\tilde{u} \in S_w^\nu(\Omega)$  with*

$$\|\mathfrak{G}_w^\nu \tilde{u}\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} \leq B.$$

*Proof.* We begin by noting that for sufficiently large  $n$ ,  $w_n$  satisfies (1). Indeed, by monotone convergence theorem,  $M_{w_n}^1 \rightarrow M_w^1$  and  $M_{w_n}^2 \rightarrow M_w^2$  as  $n \rightarrow \infty$  which yields  $M_{w_n}^1 \in (0, \infty)$  and  $M_{w_n}^2 < \infty$  for sufficiently large  $n$ . Without loss of generality, we may assume  $w_n$  satisfies (1) for all  $n$ . As before, we denote the cut-off by  $w_n^c(z) := w_n(z) \chi_{B_1(0)}(z)$  for  $n \in \mathbb{N}_+$ , and again satisfies (1). Similar to the proof of Lemma 2.5, applying (11) it follows that

$$|\lambda_{w_n^c}^\nu(\xi) - \lambda_{w_n}^\nu(\xi)| = |\lambda_{w_n - w_n^c}^\nu(\xi)| \leq 2 \int_{|z|>1} w_n(z) dz \leq 2M_w^2 + 1, \quad (20)$$

for  $n$  large enough and thus there exists a constant  $\tilde{B}$  depending only on  $B$ ,  $M_w^2$  and  $\sup_n \|u_n\|_{L^2(\mathbb{R}^d)}$  such that

$$\sup_n \int_{\mathbb{R}^d} |\lambda_{w_n^c}^\nu(\xi) \hat{u}_n(\xi)|^2 d\xi \leq \tilde{B}^2.$$

For the first part of the proof, similar to the proof of Theorem 3.3, it suffices to show that

$$\lim_{\tau \rightarrow 0^+} \limsup_{n \rightarrow \infty} \|\boldsymbol{\eta}_\tau \hat{u}_n\|_{L^2(\mathbb{R}^d; \mathbb{C}^d)} = 0, \quad (21)$$

where  $\boldsymbol{\eta}_\tau$  is given by (17). To that end, after estimating as

$$\begin{aligned} \|\boldsymbol{\eta}_\tau \hat{u}_n\|_{L^2(\mathbb{R}^d; \mathbb{C}^d)}^2 &= \int_{\mathbb{R}^d} |\boldsymbol{\eta}_\tau(\xi) \hat{u}_n(\xi)|^2 d\xi \\ &\leq \left( \sup_{\xi \neq 0} \frac{|\boldsymbol{\eta}_\tau(\xi)|}{|\lambda_{w_n^c}^\nu(\xi)|} \right)^2 \int_{\mathbb{R}^d} |\lambda_{w_n^c}^\nu(\xi) \hat{u}_n(\xi)|^2 d\xi \\ &\leq \tilde{B}^2 \left( \sup_{\xi \neq 0} \frac{|\boldsymbol{\eta}_\tau(\xi)|}{|\lambda_{w_n^c}^\nu(\xi)|} \right)^2, \end{aligned}$$

it suffices to show that

$$\lim_{\tau \rightarrow 0^+} \limsup_{n \rightarrow \infty} \sup_{\xi \neq 0} \frac{|\boldsymbol{\eta}_\tau(\xi)|}{|\lambda_{w_n^c}^\nu(\xi)|} = 0. \quad (22)$$

Let  $\epsilon_0$  and  $N_1 = N_1(w, d) > 0$  be as in Remark 3.4 such that for some  $C_1 = C_1(w, d) > 0$

$$|\lambda_{w^c}^\nu(\xi)| \geq C_1 |\xi|, \quad \forall |\xi| \leq N_1.$$

Applying (11) to  $w^c - w_n^c$ , we know that for  $n$  sufficiently large,

$$|\lambda_{w^c}^\nu(\xi) - \lambda_{w_n^c}^\nu(\xi)| = |\lambda_{w^c - w_n^c}^\nu(\xi)| \leq 2\sqrt{2\pi} |\xi| M_{w^c - w_n^c}^1 \leq \frac{1}{2} C_1 |\xi|, \quad \forall \xi \in \mathbb{R}^d,$$

since  $M_{w^c - w_n^c}^1 = M_w^1 - M_{w_n}^1 \rightarrow 0$  as  $n \rightarrow \infty$ . Therefore, for  $n$  sufficiently large, one has

$$|\lambda_{w_n^c}^\nu(\xi)| \geq \frac{1}{2} C_1 |\xi|, \quad \forall |\xi| \leq N_1.$$

On the other hand, since  $w_n^c$  satisfies (1), applying Lemma 2.2 for  $w_n^c$  and  $N = N_1$  we know that there exists a constant  $C_2 = C_2(N_1 \epsilon_0, d) > 0$  independent of  $n$  such that

$$|\lambda_{w_n^c}^\nu(\xi)| \geq C_2 \int_{\substack{N_1 \epsilon_0 \\ |\xi|}}^{N_1 \epsilon_0} |w_n(z)| dz > 0, \quad \forall |\xi| > N_1.$$

The rest of the proof builds upon the proof of Lemma 3.5. We sketch the proof using the same notations from the proof of Lemma 3.5. Fix  $0 < \tau < \epsilon_0$ . To show (22), we discuss three cases:  $0 < |\xi| \leq N_1$ ,  $N_1 \leq |\xi| \leq N_1 \epsilon_0 / \tau$  and  $|\xi| > N_1 \epsilon_0 / \tau$ . By checking the proof of Lemma 3.5 and replacing  $w^c$  by  $w_n^c$ , one may show that for  $n$  sufficiently large,

$$\sup_{0 < |\xi| \leq N_1} \frac{|\eta_\tau(\xi)|}{|\lambda_{w_n^c}^\nu(\xi)|} \leq 2 \frac{c_\eta}{C_1} \tau$$

and

$$\sup_{|\xi| > \frac{N_1 \epsilon_0}{\tau}} \frac{|\eta_\tau(\xi)|}{|\lambda_{w_n^c}^\nu(\xi)|} \leq \frac{c_\eta}{C_2} \left( \int_{\tau < |z| < 1} |w_n(z)| dz \right)^{-1}.$$

Finally, notice that

$$\begin{aligned} \sup_{|\xi| \in \left[ N_1, \frac{N_1 \epsilon_0}{\tau} \right]} \frac{|\eta_\tau(\xi)|}{|\lambda_{w_n^c}^\nu(\xi)|} &\leq \frac{c_\eta}{C_2} \sup_{|\xi| \in \left[ N_1, \frac{N_1 \epsilon_0}{\tau} \right]} \frac{\tau |\xi|}{\int_{\frac{N_1 \epsilon_0}{|\xi|} < |z| < 1} |w_n(z)| dz} \\ &= \frac{c_\eta}{C_2} \sup_{s \in [N_1 \tau, N_1 \epsilon_0]} \frac{s}{\int_{\frac{N_1 \epsilon_0 \tau}{s} < |z| < 1} |w_n(z)| dz}, \end{aligned}$$

where we did a change of variables in the last equality. Introduce the notation

$$f(n, \tau) := \sup_{s \in [N_1 \tau, N_1 \epsilon_0]} \frac{s}{\int_{\frac{N_1 \epsilon_0 \tau}{s} < |z| < 1} |w_n(z)| dz},$$

we show next that

$$\lim_{\tau \rightarrow 0^+} \limsup_{n \rightarrow \infty} f(n, \tau) = 0, \quad (23)$$

Suppose that (23) is false. Then there exists  $\alpha_0 > 0$  and sequences  $\{\tau_j\}_{j=1}^\infty$ ,  $\{n_j\}_{j=1}^\infty$  and  $\{s_j\}_{j=1}^\infty$  with  $\tau_j \rightarrow 0$ ,  $n_j \rightarrow +\infty$  and  $s_j \in [N_1 \tau_j, N_1 \epsilon_0]$  such that

$$\frac{s_j}{\int_{\frac{N_1 \epsilon_0 \tau_j}{s_j} < |z| < 1} |w_{n_j}(z)| dz} > \alpha_0.$$

Arguing as in the proof of Lemma 3.5, one may show that  $\tau_j / s_j \rightarrow 0$  as  $j \rightarrow \infty$ , and without loss of generality, one may assume that  $\{\tau_j / s_j\}_{j=1}^\infty$  is decreasing and  $\{n_j\}_{j=1}^\infty$  is increasing. Since  $\{\chi_{B_{N_1 \epsilon_0 \tau_j / s_j}^c}(0)(z) w_{n_j}(z)\}_{j=1}^\infty$  is an increasing sequence of nonnegative functions, by

monotone convergence theorem,

$$\alpha_0 < \frac{s_j}{\int_{\frac{N_1 \epsilon_0 \tau_j}{s_j} < |z| < 1} w_{n_j}(z) dz} \leq \frac{N_1 \epsilon_0}{\int_{\frac{N_1 \epsilon_0 \tau_j}{s_j} < |z| < 1} w_{n_j}(z) dz} \rightarrow \frac{N_1 \epsilon_0}{\int_{|z| < 1} w(z) dz} = 0$$

as  $n \rightarrow \infty$ , a contradiction. Therefore (23) holds, and consequently (22) holds by combining the three cases. The precompactness follows from Lemma 3.1. Finally, assuming that  $u_n \in S_w^v(\Omega)$  for every  $n$ , we show that for any limit point  $u$ , the zero-extension  $\tilde{u}$  of  $u$  outside  $\Omega$  satisfies  $\tilde{u} \in S_w^v(\Omega)$  with

$$\|\mathfrak{G}_w^v \tilde{u}\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} \leq B.$$

Without loss of generality, assume that  $u_n \rightarrow u$  in  $L^2(\Omega)$ . Then  $u_n \rightarrow \tilde{u}$  in  $L^2(\mathbb{R}^d)$ . Thus  $\hat{u}_n \rightarrow \hat{\tilde{u}}$  in  $L^2(\mathbb{R}^d)$  and up to a subsequence we know  $\hat{u}_n \rightarrow \hat{\tilde{u}}$  a.e. in  $\mathbb{R}^d$ . On the other hand, by (11) one may show that  $\lambda_{w_n}^v(\xi) \rightarrow \lambda_w^v(\xi)$  as  $n \rightarrow \infty$  for any  $\xi \in \mathbb{R}^d$  as  $M_{w_n}^1 \rightarrow M_w^1$  and  $M_{w_n}^2 \rightarrow M_w^2$ . Then by Fatou's lemma, one obtains that

$$\int_{\mathbb{R}^d} |\lambda_w^v(\xi) \hat{u}(\xi)|^2 d\xi \leq \liminf_{n \rightarrow \infty} \int_{\mathbb{R}^d} |\lambda_{w_n}^v(\xi) \hat{u}_n(\xi)|^2 d\xi \leq B^2.$$

This completes the proof.  $\blacksquare$

As a consequence of the above compactness result Theorem 3.6, we establish the following uniform Poincaré inequality.

**Theorem 3.7.** *Let  $\{w_n : n \in \mathbb{N}_+\}$  be a family of kernels as in Theorem 3.6. Then there exist  $N_0 \in \mathbb{N}_+$  and  $C(N_0) > 0$  such that for any  $n \in \mathbb{N}_+$  with  $n > N_0$ ,*

$$\|u\|_{L^2(\Omega)} \leq C(N_0) \|\mathfrak{G}_{w_n}^v u\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)}, \quad \forall u \in S_w^v(\Omega). \quad (24)$$

*Proof.* Let

$$\frac{1}{A} := \inf \left\{ \|\mathfrak{G}_w^v u\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} : u \in S_w^v(\Omega), \|u\|_{L^2(\Omega)} = 1 \right\}.$$

By the nonlocal Poincaré inequality for the kernel  $w$ , that is, [20, Theorem 5.1], one has  $0 < A < \infty$ . To prove the theorem, we will prove that for any  $\epsilon > 0$ , there exist  $N_0(\epsilon) > 0$  and  $A + \epsilon > 0$  such that for any  $n > N_0$ ,

$$\|u\|_{L^2(\Omega)} \leq (A + \epsilon) \|\mathfrak{G}_{w_n}^v u\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)}, \quad \forall u \in S_w^v(\Omega).$$

We prove this statement by contradiction. Suppose that there exist  $C > A$ , sequences  $\{n_k\}_{k=1}^\infty$  and  $\{u_{n_k}\}_{k=1}^\infty$  where  $n_k > k$  for all  $k \in \mathbb{N}$  such that  $u_{n_k} \in S_{w_{n_k}}^v(\Omega)$ ,  $\|u_{n_k}\|_{L^2(\Omega)} = 1$  and

$$\|\mathfrak{G}_{w_{n_k}}^v u_{n_k}\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} < \frac{1}{C}.$$

Then by Theorem 3.6, there exists  $u \in S_w^v(\Omega)$  such that  $u_{n_k} \rightarrow u$  in  $L^2(\Omega)$  up to a subsequence and

$$\|\mathfrak{G}_w^v u\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} \leq \frac{1}{C} < \frac{1}{A}.$$

But since the strong limit  $\|u\|_{L^2(\Omega)} = 1$ , this gives a contradiction to the fact that  $A$  is chosen the best constant.  $\blacksquare$

### 3.2 | Compactness with vanishing nonlocality

In this subsection, we prove the second compactness criterion which is based on boundedness of the sequence of nonlocal gradient energies associated with kernels with vanishing nonlocality. To that end, in this section we will be working with the set of kernels  $\{w_\delta : \delta \in (0, 1)\}$  satisfying (1) and the additional assumptions that for any  $R > 0$ ,

$$\lim_{\delta \rightarrow 0} \int_{B_R^c(\mathbf{0})} w_\delta(z) dz = 0, \quad (25)$$

$$\lim_{\delta \rightarrow 0} \int_{B_R(\mathbf{0})} |z| w_\delta(z) dz = 2d. \quad (26)$$

Conditions (25) and (26) imply that for any  $R > 0$

$$\lim_{\delta \rightarrow 0} \int_{B_R(\mathbf{0})} |z|^2 w_\delta(z) dz = 0. \quad (27)$$

Indeed, given any  $\epsilon \in (0, R)$ ,

$$\begin{aligned} \int_{B_R(\mathbf{0})} |z|^2 w_\delta(z) dz &= \int_{B_\epsilon(\mathbf{0})} |z|^2 w_\delta(z) dz + \int_{\epsilon < |z| < R} |z|^2 w_\delta(z) dz \\ &\leq \epsilon \int_{B_\epsilon(\mathbf{0})} |z| w_\delta(z) dz + R^2 \int_{B_\epsilon(\mathbf{0})^c} w_\delta(z) dz. \end{aligned}$$

Letting  $\delta \rightarrow 0$  and using (25) and (26) we obtain that

$$\limsup_{\delta \rightarrow 0} \int_{B_R(\mathbf{0})} |z|^2 w_\delta(z) dz \leq 2d\epsilon.$$

Since  $\epsilon > 0$  can be arbitrarily small, (27) holds.

A consequence of this is that these conditions on the sequence of kernels are sufficient for the convergence of corresponding nonlocal operators to their local counterparts as  $\delta$  tends to zero as stated below. The proof is presented in Appendix A.1.

**Proposition 3.8.** *Let  $u \in C_c^1(\mathbb{R}^d)$  and  $v \in C_c^1(\mathbb{R}^d; \mathbb{R}^d)$ . Let  $\{w_\delta : \delta \in (0, 1)\}$  satisfy Equations (1), (25) and (26). Denote  $G_\delta^v := G_{w_\delta}^v$  and  $D_\delta^v := D_{w_\delta}^v$ . Then for any  $\mathbf{x} \in \mathbb{R}^d$ ,*

$$G_\delta^v u(\mathbf{x}) \rightarrow \nabla u(\mathbf{x}), \quad \delta \rightarrow 0, \quad (28)$$

$$D_\delta^v v(\mathbf{x}) \rightarrow \operatorname{div} v(\mathbf{x}), \quad \delta \rightarrow 0. \quad (29)$$

Moreover, for any  $p \in [1, \infty]$ , the above two convergences are strong in  $L^p(\mathbb{R}^d; \mathbb{R}^d)$  and  $L^p(\mathbb{R}^d)$  respectively.

This localization result leads to the natural question about the convergence property of bounded sequence  $\{u_\delta \in S_{w_\delta}^v(\mathbb{R}^d)\}$ ; that is, if  $\sup_\delta |u_\delta|_{S_{w_\delta}^v} < \infty$ , is the sequence compact in  $L_{\text{loc}}^2$ ? And if yes, what can we say about limit points  $u$ ? As the localization result Proposition 3.8 suggest, does  $u$  belong in  $H_{\text{loc}}^1(\mathbb{R}^d)$ ? We conjecture that these questions can be answered in the affirmative for the general set of kernels satisfying (1), (25), and (26). Although we are unable to answer these questions in full generality, we will address the question in this subsection for a selected set of sequence of kernels. The types of sequence of kernels we will be focused on are listed in the following assumption.

**Assumption 3.9.** *We assume that  $\{w_\delta : \delta \in (0, 1)\}$  is given by one of the three cases:*

(I)  $w_\delta(z) := \delta^{-d-1} w(z/\delta)$  for  $\delta \in (0, 1)$  where  $w$  satisfies (1) and

$$\int_{\mathbb{R}^d} |z| w(z) dz = 2d.$$

(II)  $w_\delta(z) := 2d\delta|z|^{\delta-d-1}$  for  $\delta \in (0, 1)$ .

(III)  $\{w_\delta : \delta \in (0, 1)\}$  is a set of kernels satisfying Equations (1), (25), and (26), and there exist constants  $R_1 > 0$ ,  $R_2 > 0$  and  $\bar{\delta}_0 \in (0, 1)$  and a function  $h : (0, \bar{\delta}_0) \rightarrow \mathbb{R}_+$  with  $\lim_{\delta \rightarrow 0} h(\delta) = 0$  such that

$$\lim_{\delta \rightarrow 0} \int_{|z| > R_1 h(\delta)} w_\delta(z) dz = +\infty \quad (30)$$

and

$$\sup_{\delta \in (0, \bar{\delta}_0)} \frac{1}{h(\delta)} \int_{B_{R_2}(0)} |z|^2 w_\delta(z) dz := M_0 < +\infty. \quad (31)$$

In addition, when  $d = 1$  assume  $w_\delta(z)$  is nonincreasing for each  $\delta \in (0, 1)$ .

**Remark 3.10.** A typical example satisfying Assumption 3.9(III) is given by

$$w_\delta(z) = \frac{1}{|\log \delta|} \frac{1}{|z|(|z| + \delta)^d} \quad (32)$$

with  $R_1 = R_2 = \bar{\delta}_0 = 1$  and  $h(\delta) = \delta$ . One may show that the first two types of kernels in Assumption 3.9 satisfy Equations (1), (25) and (26) but not Assumption 3.9(III). Specifically, Assumption 3.9(I) does not satisfy (31) and Assumption 3.9(II) does not satisfy (30).

The main theorem of this subsection is the following.

**Theorem 3.11.** Let  $\{w_{\delta_n} : \delta_n \in (0, 1)\}_{n=1}^\infty$  be a sequence of kernels that belongs to either (I), (II) or (III) of Assumption 3.9. Suppose  $\{u_n\}_{n=1}^\infty \subset L^2(\mathbb{R}^d)$  is a bounded sequence and

$$\sup_n \|u_n\|_{S_{w_{\delta_n}}^\nu(\mathbb{R}^d)} = \sup_n \left\| \mathfrak{G}_{w_{\delta_n}}^\nu u_n \right\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} = B < \infty,$$

then  $\{u_n\}_{n=1}^\infty$  has a compact closure in  $L^2(\Omega)$  for bounded domain  $\Omega \subset \mathbb{R}^d$ . Moreover, if in addition  $\Omega$  has a Lipschitz boundary and  $u_n \in S_{w_{\delta_n}}^\nu(\Omega)$  for every  $n$ , then any limit point  $u \in H_0^1(\Omega)$  with

$$\|\nabla u\|_{L^2(\Omega; \mathbb{R}^d)} \leq B.$$

The theorem will be shown for the three classes of kernels in Assumption 3.9 separately, each requiring subtly different techniques. From now on, we let  $\{\delta_n\}_{n=1}^\infty$  be a sequence of positive numbers such that  $\delta_n \in (0, 1)$  and  $\delta_n \rightarrow 0$  as  $n \rightarrow \infty$ .

*Proof of Theorem 3.11 for the class of kernels satisfying Assumption 3.9(I).* Taking the Fourier transform and by applying Proposition 2.4 we have,

$$\sup_n \left\| \lambda_{w_{\delta_n}}^\nu \hat{u}_n \right\|_{L^2(\mathbb{R}^d; \mathbb{C}^d)} = \sup_n \left\| \mathfrak{G}_{w_{\delta_n}}^\nu u_n \right\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} = B < \infty.$$

We prove the theorem in two step. We first show the sequence is compact in the  $L^2_{\text{loc}}$  topology, and then demonstrate that any limit point is in  $H_0^1(\Omega)$ .

**Step I.** To show that  $\{u_n\}_{n=1}^\infty$  has a compact closure in  $L^2(\Omega)$  for bounded domain  $\Omega \subset \mathbb{R}^d$ , we will apply the compactness criterion given in Lemma 3.1. For that we choose  $P(x) := \chi_\nu(x) \chi_{B_1(0)}(x) x/|x|$  and demonstrate that

$$\lim_{\tau \rightarrow 0^+} \limsup_{n \rightarrow \infty} \|P_\tau * u_n - P u_n\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} = 0. \quad (33)$$

As before, we consider  $\eta_\tau$  given by (17).

**Claim.** There exists  $C = C(w, d)$  such that for  $0 < \delta \leq \tau < 1$ ,

$$|\eta_\tau(\xi)| \leq C\tau|\lambda_{w_\delta}^\nu(\xi)|, \quad \forall \xi \in \mathbb{R}^d. \quad (34)$$

Once this claim is proved, (33) holds. Indeed, since for any  $\tau \in (0, 1)$ , there exists  $n_\tau > 0$  such that for any  $n > n_\tau$ ,  $\delta_n < \tau$  we have  $|\eta_\tau(\xi)| \leq C\tau|\lambda_n^\nu(\xi)|$  for all  $\xi \in \mathbb{R}^d$ . Thus for  $n > n_\tau$ ,

$$\begin{aligned} \|P_\tau * u_n - pu_n\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)}^2 &= \int_{\mathbb{R}^d} |\eta_\tau(\xi)\hat{u}_n(\xi)|^2 d\xi \\ &\leq C\tau^2 \int_{\mathbb{R}^d} \left| \lambda_{w_{\delta_n}}^\nu(\xi)\hat{u}_n(\xi) \right|^2 d\xi \leq CB^2\tau^2. \end{aligned}$$

What remains is to prove the above claim. By a change of variable, one may check that  $\lambda_{w_\delta}^\nu(\xi) = \delta^{-1}\lambda_w^\nu(\delta\xi)$  for all  $\xi \in \mathbb{R}^d$ . Therefore, for  $0 < \delta \leq \tau < 1$ , by Lemma 2.1, there exist constants  $N_1 = N_1(w, d)$  and  $C_1 = C_1(w, d)$  such that

$$|\lambda_{w_\delta}^\nu(\xi)| \geq C_1|\xi|, \quad \forall |\xi| < \frac{N_1}{\delta}$$

and

$$|\lambda_{w_\delta}^\nu(\xi)| \geq \frac{C_1}{\delta} \int_{\mathbb{R}^d} \min(1, w(x)) dx \geq \frac{C_2(w, d)}{\tau}, \quad \forall |\xi| \geq \frac{N_1}{\delta}.$$

Comparing these estimates with (18) proves the (34), completing the proof of the claim.

**Step II.** Suppose  $u_n \in S_{w_{\delta_n}}^\nu(\Omega)$  for every  $n$  and  $u$  is a limit point of  $\{u_n\}_{n=1}^\infty$  in  $L^2(\Omega)$ , that is,  $u_n \rightarrow u$  in  $L^2(\Omega)$  up to a subsequence. Then we show that  $u \in H_0^1(\Omega)$  and  $\|\nabla u\|_{L^2(\Omega; \mathbb{R}^d)} \leq B$ . Let  $\tilde{u}$  be the zero extension of  $u$  outside  $\Omega$ , that is,

$$\tilde{u}(x) = \begin{cases} u(x), & x \in \Omega, \\ 0, & x \in \Omega^c. \end{cases}$$

Then  $u_n \rightarrow \tilde{u}$  in  $L^2(\mathbb{R}^d)$ . By definition of distributional gradient, for any  $\phi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$ ,

$$\begin{aligned} \left| \int_{\mathbb{R}^d} u_n(x) D_{w_{\delta_n}}^{-\nu} \phi(x) dx \right| &= \left| - \int_{\mathbb{R}^d} \mathfrak{G}_{w_{\delta_n}}^\nu u_n(x) \cdot \phi(x) dx \right| \\ &\leq \|\mathfrak{G}_{w_{\delta_n}}^\nu u_n\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} \|\phi\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} \\ &\leq B \|\phi\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)}. \end{aligned}$$

Passing to the limit as  $n \rightarrow \infty$ , one obtains

$$\left| \int_{\mathbb{R}^d} \tilde{u}(x) \operatorname{div} \phi(x) dx \right| \leq B \|\phi\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)}, \quad \forall \phi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d), \quad (35)$$

where we used Proposition 3.8 to assert  $D_{w_{\delta_n}}^{-\nu} \phi \rightarrow \operatorname{div} \phi$  in  $L^2(\mathbb{R}^d)$ . Therefore,  $\tilde{u} \in H^1(\mathbb{R}^d)$  with

$$\|\nabla \tilde{u}\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} \leq B.$$

Then by [40, Theorem 3.7],  $u \in H_0^1(\Omega)$  and  $\|\nabla u\|_{L^2(\Omega; \mathbb{R}^d)} \leq B$ , where we used the fact  $\Omega$  is a bounded Lipschitz domain. ■

*Proof of Theorem 3.11 for the class of kernels satisfying Assumption 3.9(II).* Again we will use the Riesz–Kolmogorov–Fréchet criterion Lemma 3.1. The key ingredient is a good lower bound on the norm of Fourier symbol  $\lambda_{w_\delta}^\nu$ .

We claim that if  $w_\delta(z) = 2d\delta|z|^{\delta-d-1}$ , then there exist constants  $C_1(d) > 0$ ,  $C_2(d) > 0$  and  $\delta_0 \in (0, 1/2)$  such that for any  $\delta \in (0, \delta_0)$ ,

$$C_1(d)|\xi|^{1-\delta} \leq |\lambda_{w_\delta}^v(\xi)| \leq C_2(d)|\xi|^{1-\delta}, \quad \forall \xi \in \mathbb{R}^d. \quad (36)$$

We will prove (36) shortly but a consequence of it is that there exists  $C = C(d)$  such that for  $\tau \in (0, 1)$  and  $\delta \in (0, \delta_0)$ ,

$$|\eta_\tau(\xi)| \leq C\tau^{\frac{1}{2}}|\lambda_{w_\delta}^v(\xi)|, \quad \forall \xi \in \mathbb{R}^d, \quad (37)$$

from which the condition in Lemma 3.1 is satisfied.

To prove (37), we estimate it in two regimes. Let  $\delta_0$  be the constant used in (36). Then for any  $\delta \in (0, \delta_0)$  and  $|\xi| \leq 1/\tau$ , using (18) and (36) we know that there exists a constant  $C(d) > 0$  such that

$$\frac{|\eta_\tau(\xi)|}{|\lambda_{w_\delta}^v(\xi)|} \leq C(d) \frac{\tau|\xi|}{|\xi|^{1-\delta}} \leq C(d)\tau|\xi|^\delta \leq C(d)\tau^{1-\delta} \leq C(d)\tau^{\frac{1}{2}}.$$

Similarly, for  $|\xi| > 1/\tau$ , again using (18) and (36), we have

$$\frac{|\eta_\tau(\xi)|}{|\lambda_{w_\delta}^v(\xi)|} \leq C(d) \frac{1}{|\xi|^{1-\delta}} \leq C(d)\tau^{1-\delta} \leq C(d)\tau^{\frac{1}{2}}.$$

Let us now prove (36). By (10) and the form of  $w_\delta$ , for  $\delta \in (0, 1/2)$  and  $\xi \in \mathbb{R}^d$  we have

$$\begin{aligned} |\lambda_{w_\delta}^v(\xi)| &\leq 2d \int_{\mathbb{R}^d} \frac{\delta}{|z|^{d+1-\delta}} \left| e^{2\pi iz \cdot \xi} - 1 \right| dz \\ &= 2d|\xi|^{1-\delta} \int_{\mathbb{R}^d} \frac{\delta}{|z|^{d+1-\delta}} \left| e^{2\pi iz \cdot \xi/|\xi|} - 1 \right| dz \\ &\leq 2d\delta|\xi|^{1-\delta} \left( \int_{B_1(0)} \frac{2\pi|z|}{|z|^{d+1-\delta}} dz + \int_{B_1(0)^c} \frac{2}{|z|^{d+1-\delta}} dz \right) \\ &= 2d\delta|\xi|^{1-\delta} \omega_{d-1} \left( \frac{2\pi}{\delta} + \frac{2}{1-\delta} \right) \\ &\leq 8d\pi\omega_{d-1}|\xi|^{1-\delta}. \end{aligned}$$

For the lower bound, we recall from [20, Lemma 5.1] that when  $d \geq 2$  the norm of imaginary part of  $\lambda_{w_\delta}^v(\xi)$  is given by

$$\Im(\lambda_{w_\delta}^v)(\xi) = 2d\omega_{d-2} \int_0^{\frac{\pi}{2}} \cos(\theta) \sin^{d-2}(\theta) \int_0^\infty \frac{\delta}{r^{2-\delta}} \sin(2\pi|\xi|r \cos(\theta)) dr d\theta.$$

By a change of variable we have

$$|\lambda_{w_\delta}^v(\xi)| \geq 2d|\xi|^{1-\delta} \omega_{d-2} \int_0^{\frac{\pi}{2}} \cos^{2-\delta}(\theta) \sin^{d-2}(\theta) \int_0^\infty \frac{\delta}{r^{2-\delta}} \sin(2\pi r) dr d\theta.$$

By Proposition A.2 in Appendix A.3 we know as  $\delta \rightarrow 0$ ,

$$\int_0^{\frac{\pi}{2}} \cos^{2-\delta}(\theta) \sin^{d-2}(\theta) \int_0^\infty \frac{\delta}{r^{2-\delta}} \sin(2\pi r) dr d\theta \rightarrow 2\pi \int_0^{\frac{\pi}{2}} \cos^2(\theta) \sin^{d-2}(\theta) d\theta > 0.$$

Then there exist constants  $C_1(d) > 0$  and  $\delta_0 \in (0, 1/2)$  such that for any  $\delta \in (0, \delta_0)$ ,

$$|\lambda_{w_\delta}^v(\xi)| \geq C_1(d)|\xi|^{1-\delta}, \quad \forall \xi \in \mathbb{R}^d.$$

For  $d = 1$  the proof is similar as one can show  $|\lambda_{w_\delta}^v(\xi)| \geq 2|\xi|^{1-\delta} \int_0^\infty \frac{\delta}{r^{2-\delta}} \sin(2\pi r) dr$  for any  $\xi \in \mathbb{R}$ . Thus the inequality (36) is established. ■

To prove Theorem 3.11 when the set of kernels satisfies Assumption 3.9(III), we first establish lower bound estimates of the parameterized symbols  $\lambda_{w_\delta^c}^V(\xi)$  that is uniform in  $\delta$ . Here  $w_\delta^c(z)$  stands for the compactly supported kernel after truncation  $w_\delta^c(z) := w_\delta(z)\chi_{B_{R_2}(0)}(z)$ . We achieve this by separately estimating the real and imaginary parts of  $\lambda_{w_\delta^c}^V(\xi)$ . We recall that we can write

$$\lambda_{w_\delta^c}^V(\xi) = \Re(\lambda_{w_\delta^c}^V)(\xi) + i\Im(\lambda_{w_\delta^c}^V)(\xi)$$

where

$$\begin{aligned}\Re(\lambda_{w_\delta^c}^V)(\xi) &= \int_{B_{R_2}(0)} \chi_V(z) \frac{z}{|z|} w_\delta(z) (\cos(2\pi\xi \cdot z) - 1) dz \quad \text{and} \\ \Im(\lambda_{w_\delta^c}^V)(\xi) &= \int_{B_{R_2}(0)} \chi_V(z) \frac{z}{|z|} w_\delta(z) \sin(2\pi\xi \cdot z) dz.\end{aligned}$$

We estimate  $|\Im(\lambda_{w_\delta^c}^V)(\xi)|$  and  $|\Re(\lambda_{w_\delta^c}^V)(\xi)|$  for small  $|\xi|$  and large  $|\xi|$  respectively.

**Lemma 3.12.** *Let  $\{w_\delta : \delta \in (0, 1)\}$  be a family of kernels satisfying Assumption 3.9. Then there exist constants  $\bar{\delta}_1 = \bar{\delta}_1(h, R_1, R_2) \in (0, \bar{\delta}_0)$ ,  $C_0 = C_0(M_0, d)$ ,  $C_1 = C_1(d)$  and  $C_2 = C_2(R_1, M_0, d)$  such that for any  $\delta \in (0, \bar{\delta}_1)$ ,*

$$|\Im(\lambda_{w_\delta^c}^V)(\xi)| \geq C_1 |\xi|, \quad \forall |\xi| \leq \frac{C_0}{h(\delta)}, \quad (38)$$

and

$$|\Re(\lambda_{w_\delta^c}^V)(\xi)| \geq C_2 \int_{|z| \geq R_1 h(\delta)} w_\delta^c(z) dz > 0, \quad \forall |\xi| > \frac{C_0}{h(\delta)}. \quad (39)$$

*Proof.* Without loss of generality, we assume  $d \geq 2$ . The case  $d = 1$  can be proved similarly. In [20, Lemma 5.1] it is shown that

$$|\Im(\lambda_{w_\delta^c}^V)(\xi)| = \left| \omega_{d-2} \int_0^{\frac{\pi}{2}} \cos(\theta) \sin^{d-2}(\theta) \int_0^{R_2} r^{d-1} \overline{w_\delta}(r) \sin(2\pi|\xi|r \cos(\theta)) dr d\theta \right|.$$

Using the inequality  $\sin x \geq x - x^2/2$  for  $x \geq 0$ , one obtains that

$$\begin{aligned}|\Im(\lambda_{w_\delta^c}^V)(\xi)| &\geq \omega_{d-2} \int_0^{\frac{\pi}{2}} \cos(\theta) \sin^{d-2}(\theta) \cdot \\ &\quad \int_0^{R_2} r^{d-1} \overline{w_\delta}(r) (2\pi|\xi|r \cos(\theta) - 2\pi^2|\xi|^2 r^2 \cos^2(\theta)) dr d\theta \\ &= |\xi| \left( c_1(d) \int_{B_{R_2}(0)} |z| w_\delta(z) dz - c_2(d) |\xi| \int_{B_{R_2}(0)} |z|^2 w_\delta(z) dz \right)\end{aligned}$$

where  $c_1(d)$  and  $c_2(d)$  are positive constants. Since

$$\lim_{\delta \rightarrow 0} h(\delta) = 0 \quad \text{and} \quad \lim_{\delta \rightarrow 0} \int_{B_{R_2}(0)} |z| w_\delta(z) dz = 2d,$$

there exists  $\bar{\delta}_1 \in (0, \bar{\delta}_0)$  depending on  $R_1$  and  $R_2$  such that

$$R_1 h(\delta) < R_2 \quad \text{and} \quad \int_{B_{R_2}(0)} |z| w_\delta(z) dz > d$$

for any  $\delta \in (0, \bar{\delta}_1)$ . Choose  $C_0 = C_0(M_0, d)$  such that  $C_0 M_0 c_2(d) = c_1(d)d/2$ , then for any  $\delta \in (0, \bar{\delta}_1)$  and  $|\xi| \leq C_0/h(\delta)$ , by the calculation above and (31),

$$\Im(\lambda_{w_\delta^c}^V)(\xi) \geq |\xi| \left( c_1(d)d - c_2(d) \frac{C_0}{h(\delta)} \int_{B_{R_2}(0)} |z|^2 w_\delta(z) dz \right) \geq \frac{1}{2} c_1(d)d |\xi|.$$

Then (38) is shown where  $C_1 = c_1(d)d/2$ . By Lemma 2.2, taking  $N = C_0/h(\delta)$  and  $\epsilon = R_1h(\delta)$  there exist a constant  $C_2 = C_2(R_1, C_0, d) = C_2(R_1, M_0, d) > 0$  such that for  $|\xi| > C_0/h(\delta)$ ,

$$|\Re(\lambda_{w_\delta^c}^\nu)(\xi)| \geq C_2 \int_{|z| \geq \frac{C_0 R_1}{|\xi|}} w_\delta^c(z) dz \geq \int_{|z| \geq R_1 h(\delta)} w_\delta^c(z) dz.$$

Note that (30) implies that  $B_{R_1 h(\delta)}(\mathbf{0}) \subset \text{supp } w_\delta^c$  for  $\delta$  small enough. Thus without loss of generality one may assume that  $B_{R_1 h(\delta)}(\mathbf{0}) \subset \text{supp } w_\delta$  for  $\delta \in (0, \bar{\delta}_0)$ . The proof is then complete. ■

*Remark 3.13.* Following the exact proof of Lemma 2.5, one may show that for  $\{w_\delta : \delta \in (0, 1)\}$  be a family of kernels satisfying (1), the function spaces  $S_{w_\delta}^\nu(\mathbb{R}^d)$  and  $S_{w_\delta^c}^\nu(\mathbb{R}^d)$  are equivalent with equivalent norms. For a fixed constant  $R > 0$ , if we denote  $w_\delta^c(z) := w_\delta(z)\chi_{B_R(\mathbf{0})}(z)$  as the compactly supported kernel obtained from a truncation of  $w_\delta$ , then there exist constants  $\bar{\delta}_2 \in (0, 1)$  depending on  $R$  such that for any  $\delta \in (0, \bar{\delta}_2)$

$$\frac{1}{\sqrt{2}} \|u\|_{S_{w_\delta}^\nu(\mathbb{R}^d)} \leq \|u\|_{S_{w_\delta^c}^\nu(\mathbb{R}^d)} \leq \sqrt{2} \|u\|_{S_{w_\delta}^\nu(\mathbb{R}^d)}, \quad \forall u \in S_{w_\delta}^\nu(\mathbb{R}^d). \quad (40)$$

We are now ready to prove Theorem 3.11 for the last class of kernels.

*Proof of Theorem 3.11 for the class of kernels satisfying Assumption 3.9(III).* Denote  $w_{\delta_n}^c(z) := w_{\delta_n}(z)\chi_{B_{R_2}(\mathbf{0})}(z)$  where  $R_2$  is given by (31). Similar to the proof of Theorem 3.6 and in view of Remark 3.13, one can show that there exists a constant  $\tilde{B}$  depending only on  $B$  and  $\sup_n \|u_n\|_{L^2(\mathbb{R}^2)}$  such that

$$\sup_n \int_{\mathbb{R}^d} |\lambda_{w_{\delta_n}^c}^\nu(\xi) \hat{u}_n(\xi)|^2 d\xi \leq \tilde{B}^2.$$

As in the proof of Theorem 3.6, it suffices to show

$$\lim_{\tau \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\xi \neq 0} \frac{|\eta_\tau(\xi)|}{|\lambda_{w_{\delta_n}^c}^\nu(\xi)|} = 0. \quad (41)$$

Choose  $\bar{\delta}_1$  as in Lemma 3.12. Since  $\delta_n \rightarrow 0$  as  $n \rightarrow \infty$ , there exists  $N > 0$  such that for any  $n > N$ ,  $\delta_n \in (0, \bar{\delta}_1)$ . For  $n > N$ , we consider two cases  $|\xi| \leq C_0/h(\delta_n)$  and  $|\xi| > C_0/h(\delta_n)$ .

For  $|\xi| \leq C_0/h(\delta_n)$ , using (38) in Lemma 3.12 and (18) we know that there exists a constant  $C(d) > 0$  such that

$$\frac{|\eta_\tau(\xi)|}{|\lambda_{w_{\delta_n}^c}^\nu(\xi)|} \leq C(d) \frac{\tau |\xi|}{|\xi|} \leq C(d)\tau.$$

Similarly, for  $|\xi| > C_0/h(\delta_n)$ , (39) in Lemma 3.12 and (18) yield

$$\frac{|\eta_\tau(\xi)|}{|\lambda_{w_{\delta_n}^c}^\nu(\xi)|} \leq \frac{C(d)}{C_2(R_1, M_0, d) \int_{|z| > R_1 h(\delta_n)} w_{\delta_n}^c(z) dz} \rightarrow 0, \quad n \rightarrow \infty,$$

where we used (30) and (25) to obtain

$$\int_{|z| > R_1 h(\delta_n)} w_{\delta_n}^c(z) dz = \int_{|z| > R_1 h(\delta_n)} w_{\delta_n}(z) dz - \int_{|z| > R_2} w_{\delta_n}(z) dz \rightarrow \infty, \quad n \rightarrow \infty.$$

Therefore, (41) holds and the rest of the proof follows is similar to the proof given for the other class of kernels. ■

We conclude this section by presenting an application of the compactness result to prove a uniform Poincaré inequality applicable the parameterized function spaces  $S_{w_\delta}^v(\Omega)$ . The proof of this inequality follows the exact procedures as the proof of Theorem 3.7 and thus omitted.

**Theorem 3.14.** *Let  $\{w_\delta : \delta \in (0, 1)\}$  be a family of kernels given by Assumption 3.9.*

*Assume  $\Omega$  is a bounded Lipschitz domain. Then there exist  $\delta_0 > 0$  and  $C(\delta_0) > 0$  such that for any  $\delta \in (0, \delta_0)$ ,*

$$\|u\|_{L^2(\Omega)} \leq C(\delta_0) \|\mathfrak{G}_{w_\delta}^v u\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)}, \quad \forall u \in S_{w_\delta}^v(\Omega). \quad (42)$$

## 4 | APPLICATION I: CONVERGENCE OF PARAMETERIZED PROBLEMS AND ASYMPTOTICALLY COMPATIBLE SCHEMES

### 4.1 | Parameterized problems

In this section, we apply the two compactness results proved in the previous section to study the convergence properties of parameterized problems and the asymptotic compatibility of some numerical schemes for solving these problems. Let  $\Omega$  be a bounded Lipschitz domain in  $\mathbb{R}^d$  and the function  $A \in L^\infty(\mathbb{R}^d; \mathbb{R}^{d \times d})$  be a symmetric matrix such that there exists a constant  $\mu > 0$  such that

$$\xi^T A(x) \xi \geq \mu |\xi|^2, \quad \forall x \in \mathbb{R}^d, \quad \xi \in \mathbb{R}^d.$$

We would like to investigate the parameterized nonlocal equations

$$\begin{cases} -\mathfrak{D}_{w_n}^{-v}(A \mathfrak{G}_{w_n}^v u) = f & \text{in } \Omega, \\ u = 0 & \text{in } \mathbb{R}^d \setminus \Omega, \end{cases} \quad (43)$$

where the sequence of kernels  $\{w_n\}$  will have some specific properties. From the ellipticity of the coefficient  $A$ , we see that the energy space for the problem is  $S_{w_n}^v(\Omega)$ . The weak form of (43) can be expressed as follows. Given  $f \in L^2(\Omega)$ , find  $u \in S_{w_n}^v(\Omega)$  such that

$$B_{w_n}(u, v) = (f, v)_{L^2(\Omega)}, \quad \forall v \in S_{w_n}^v(\Omega), \quad (44)$$

where the bilinear form  $B_{w_n} : S_{w_n}^v(\Omega) \times S_{w_n}^v(\Omega) \rightarrow \mathbb{R}$  is given by

$$B_{w_n}(u, v) := \int_{\mathbb{R}^d} A(x) \mathfrak{G}_{w_n}^v u(x) \cdot \mathfrak{G}_{w_n}^v v(x) dx. \quad (45)$$

We analyze these parameterized nonlocal equation for two classes of sequence of kernels.

**Class I.** For a kernel  $w(x)$  satisfying Assumption 3.2, we consider a sequence of radial kernels  $\{w_n\}$  where

$$0 \leq w_n(z) \nearrow w(z) \quad \text{as } n \rightarrow \infty \quad \text{for a.e. } z \in \mathbb{R}^d.$$

These are exactly the kernels considered in Theorem 3.6. We also notice that using Lax–Milgram theorem, corresponding to  $f \in L^2(\Omega)$  and for each  $n$ , there is a unique solution  $u_n \in S_{w_n}^v(\Omega)$  to (44), see [20], with the stability estimate

$$|u_n|_{S_{w_n}^v(\Omega)} \leq C \|f\|_{L^2(\Omega)}$$

for some  $C > 0$  (independent of  $n$ ) where we have used the uniform Poincaré inequality Theorem 3.7. The result we are going to state shortly asserts that  $u_n \rightarrow u$  strongly in  $L^2(\Omega)$  where  $u \in S_w^v(\Omega)$  solves

$$B_w(u, v) = (f, v)_{L^2(\Omega)}, \quad \forall v \in S_w^v(\Omega). \quad (46)$$

Again a unique solution  $u$  to (46) exists and, generally, it is expected to be more regular than  $u_n$  since  $w$  is nonintegrable and  $S_w^v(\Omega)$  is compactly contained in  $L^2(\Omega)$ . We will also establish connection with discrete approximations of (44) with the solution  $u$  of (46). To that end, consider the Galerkin approximation of (44) where we look for  $u_{n,h} \in V_{n,h}$  such that

$$B_{w_n}(u_{n,h}, v) = (f, v)_{L^2(\Omega)}, \quad \forall v \in V_{n,h}. \quad (47)$$

where the finite element spaces  $\{V_{n,h}\}_h \subset S_{w_n}^v(\Omega)$  are the conforming spaces of piecewise polynomials, that is,

$$V_{n,h} := \{v \in S_{w_n}^v(\Omega) : v|_T \in P(T), \forall T \in \mathcal{T}_h\},$$

where  $P(T) = \mathcal{P}_k(T)$  is the space of polynomials on  $T \in \mathcal{T}_h$  with degree less or equal than a given nonnegative integer  $k$ , and  $\mathcal{T}_h$  is a quasi-uniform triangulation on  $\Omega$ .

The theorem we state below says that the discrete solutions  $u_{n,h} \rightarrow u$  in  $L^2(\Omega)$  and as such the discretized problem (47) can be viewed as an approximation to the nonlocal problem (46). Moreover, since  $S_w^v(\Omega)$  is compactly contained in  $L^2(\Omega)$ , the finite element space  $V_{n,h}$  which is conforming in the rough space  $S_w^v(\Omega)$ , may not be conforming in the regular space  $S_w^v(\Omega)$  for a sufficiently singular kernel  $w$ . We then conclude that the Galerkin approximation scheme (47) can be considered as a (possibly discontinuous) Galerkin (DG) scheme for (46).

**Theorem 4.1.** *Assume the degree of polynomials  $k \geq 1$ . Let  $\{u_n\}_n$ ,  $u$  and  $\{u_{n,h}\}_{n,h}$  be the solutions to problems (44), (46) and (47) respectively. Then there exists a constant  $C > 0$  such that for each  $n \in \mathbb{N}$ ,*

$$\|u_n - u_{n,h}\|_{S_{w_n}^v(\Omega)} \leq C \inf_{v_{n,h} \in V_{n,h}} \|u_n - v_{n,h}\|_{S_{w_n}^v(\Omega)} \rightarrow 0, \quad h \rightarrow 0,$$

and, moreover,

$$\|u_n - u\|_{L^2(\Omega)} \rightarrow 0, \quad n \rightarrow \infty,$$

and

$$\|u_{n,h} - u\|_{L^2(\Omega)} \rightarrow 0, \quad n \rightarrow \infty, \quad h \rightarrow 0.$$

**Class II.** We next consider a family of kernels  $\{w_{\delta_n}\}_n$  given in Assumption 3.9 where  $\delta_n \rightarrow 0$  as  $n \rightarrow \infty$ .

As before, for each  $n$ , we may apply Lax–Milgram theorem to show that a unique solution  $u_{\delta_n} \in S_{w_{\delta_n}}^v(\Omega)$  exists corresponding to  $f \in L^2(\Omega)$  to (44) for kernels of this class. There will also be a positive constant  $C > 0$ , independent of  $n$ , where for each  $n$ ,

$$|u_{\delta_n}|_{S_{w_{\delta_n}}^v(\Omega)} \leq C \|f\|_{L^2(\Omega)}.$$

For the above estimate we have used the uniform Poincaré inequality Theorem 3.14. We may now use the compactness result Theorem 3.11 to deduce the existence of  $u_0 \in H_0^1(\Omega)$  such that (up to a subsequence)  $u_{\delta_n} \rightarrow u_0$  in  $L^2(\Omega)$  as  $n \rightarrow \infty$ . As we will state in the next theorem, it turns out that  $u_0 \in H_0^1(\Omega)$  solves

$$B_0(u_0, v) = (f, v)_{L^2(\Omega)}, \quad \forall v \in H_0^1(\Omega), \quad (48)$$

where  $B_0(u, v) := \int_{\Omega} A(\mathbf{x}) \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x}) d\mathbf{x}$  for any  $u, v \in H_0^1(\Omega)$ . Moreover, considering a conforming finite element spaces  $\{V_{\delta_n,h}\}_h \subset S_{w_{\delta_n}}^v(\Omega)$  which are the spaces of piecewise polynomials of degree less or equal than  $k$ , that is,

$$V_{\delta_n,h} := \{v \in S_{w_{\delta_n}}^v(\Omega) : v|_K \in \mathcal{P}_k(K), \forall K \in \mathcal{T}_h\},$$

and the Galerkin approximation  $u_{\delta_n,h} \in V_{\delta_n,h}$  such that

$$B_{\delta_n}(u_{\delta_n,h}, v) = (f, v)_{L^2(\Omega)}, \quad \forall v \in V_{\delta_n,h}, \quad (49)$$

the theorem asserts that  $u_{n,h} \rightarrow u_0$  as  $n \rightarrow \infty$  and  $h \rightarrow 0$ . The following theorem states these results precisely.

**Theorem 4.2.** *Assume that  $\{w_{\delta_n}\}_n$  is given in Assumption 3.9 where  $\delta_n \rightarrow 0$  as  $n \rightarrow \infty$ . Assume also that the degree of polynomials  $k \geq 1$ . Let  $\{u_{\delta_n}\}_n$ ,  $u_0$  and  $\{u_{\delta_n,h}\}_{n,h}$  be the solutions to problems (44), (48) and (49) respectively. Then there exists a constant  $C > 0$  such that for each  $n \in \mathbb{N}_+$ ,*

$$\|u_{\delta_n} - u_{\delta_n,h}\|_{S_{w_{\delta_n}}(\Omega)} \leq C \inf_{v_{n,h} \in V_{\delta_n,h}} \|u_{\delta_n} - v_{n,h}\|_{S_{w_{\delta_n}}(\Omega)} \rightarrow 0, \quad h \rightarrow 0,$$

and, moreover,

$$\|u_{\delta_n} - u_0\|_{L^2(\Omega)} \rightarrow 0, \quad n \rightarrow \infty,$$

and

$$\|u_{\delta_n,h} - u_0\|_{L^2(\Omega)} \rightarrow 0, \quad n \rightarrow \infty, \quad h \rightarrow 0.$$

We defer the proofs of Theorems 4.1 and 4.2 to the next subsection but note that these two theorems are particular instances of a more abstract result obtained in [35, 36] on the convergence of asymptotically compatible (AC) schemes for parameterized linear problems. A slightly modified version of the abstract formulation is presented next.

## 4.2 | Asymptotically compatible schemes for linear equations with bounded and measurable coefficients

Let  $\{S_\sigma : \sigma \in \Sigma\}$  be a family of Hilbert spaces over  $\mathbb{R}$ , where the index set  $\Sigma$  is a subset of  $\mathbb{R}$  such that  $0 \in \Sigma$  and  $\sup \Sigma = \infty$ . Here we identify the dual space of  $S_0$  with itself,  $S_0^* = S_0$ . (In applications we will take  $S_0$  as the space of  $L^2$  functions.) Let  $S_\infty$  and  $\mathcal{H}$  be two Hilbert spaces over  $\mathbb{R}$  which are continuously embedded into  $S_0$ . Associated with the spaces  $\{S_\sigma : \sigma \in \Sigma \cup \{\infty\}\}$ , there are bilinear forms  $\{B_\sigma : S_\sigma \times S_\sigma \rightarrow \mathbb{R}\}$  respectively.

Given  $f \in S_0$ , we consider the parameterized variational problems: for each  $\sigma \in \Sigma$ , find  $u_\sigma \in S_\sigma$  such that

$$B_\sigma(u_\sigma, v) = (f, v)_{S_0}, \quad \forall v \in S_\sigma, \quad (50)$$

and the limiting problem: find  $u_\infty \in S_\infty$  such that

$$B_\infty(u_\infty, v) = (f, v)_{S_0}, \quad \forall v \in S_\infty. \quad (51)$$

Let  $\{V_{\sigma,h} \subset S_\sigma : \sigma \in \Sigma, h \in (0, h_0)\}$  be a family of finite-dimensional Hilbert spaces and consider the Galerkin approximation of (50): for each  $\sigma \in \Sigma$  and  $h \in (0, h_0)$ , find  $u_{\sigma,h} \in V_{\sigma,h}$  such that

$$B_\sigma(u_{\sigma,h}, v) = (f, v)_{S_0}, \quad \forall v \in V_{\sigma,h}. \quad (52)$$

The main question we would like to address is whether  $u_{\sigma,h} \rightarrow u_\infty$  in  $S_0$  as  $(\sigma, h) \rightarrow (\infty, 0)$ . As has been discussed extensively in [35, 36], in general the answer is negative. However, under addition assumptions, convergence is shown to hold.

**Assumption 4.3.** *Assume that there exist four positive constants  $M_1$ ,  $M_2$ ,  $C_1$ , and  $C_2$  independent of  $\sigma$  such that the following conditions hold:*

(A) *Uniform embeddings and asymptotically compact embeddings:*

- (Ai) For any  $\sigma \in \Sigma$ ,  $\mathcal{H} \subset \mathcal{S}_\sigma$ ,  $M_1 \|u\|_{\mathcal{S}_0} \leq \|u\|_{\mathcal{S}_\sigma}$  for any  $u \in \mathcal{S}_\sigma$ , and  $\|u\|_{\mathcal{S}_\sigma} \leq M_2 \|u\|_{\mathcal{H}}$  for any  $u \in \mathcal{H}$ .
- (Aii) For any set  $\{u_\sigma : \sigma \in \Sigma\}$ , if  $\sup\{\|u_\sigma\|_{\mathcal{S}_\sigma} : \sigma \in \Sigma\}$  is finite, then  $\{u_\sigma : \sigma \in \Sigma\}$  is relatively compact in  $\mathcal{S}_0$  with each limit point in  $\mathcal{S}_\infty$ .

(B) *Boundedness and coercivity of the bilinear forms  $\{B_\sigma : \sigma \in \Sigma \cup \{\infty\}\}$ :*

- (Bi)  $|B_\sigma(u, v)| \leq C_2 \|u\|_{\mathcal{S}_\sigma} \|v\|_{\mathcal{S}_\sigma}$ ,  $\forall u, v \in \mathcal{S}_\sigma$ ,  $\forall \sigma \in \Sigma \cup \{\infty\}$ .
- (Bii)  $B_\sigma(u, u) \geq C_1 \|u\|_{\mathcal{S}_\sigma}^2$ ,  $\forall u \in \mathcal{S}_\sigma$ ,  $\forall \sigma \in \Sigma \cup \{\infty\}$ .

(C) *Existence of a subset  $\mathcal{S}_*$  of  $\mathcal{H}$  such that*

- (Ci)  $\mathcal{S}_* \subset \cap_{\sigma \in \Sigma \cup \{\infty\}} \mathcal{S}_\sigma$  and  $\mathcal{S}_*$  is dense in  $\mathcal{S}_\sigma$  for each  $\sigma \in \Sigma \cup \{\infty\}$  with respect to their norms,

- (Cii) for any sequence  $\{u_{\sigma_k} \in \mathcal{S}_{\sigma_k}\}_{k \in \mathbb{N}_+}^{\Sigma \ni \sigma_k \rightarrow \infty}$  and  $u \in \mathcal{S}_\infty$  satisfying

$$\sup_{k \in \mathbb{N}_+} \|u_{\sigma_k}\|_{\mathcal{S}_{\sigma_k}} < \infty$$

and  $u_{\sigma_k} \rightarrow u$  in  $\mathcal{S}_0$  as  $k \rightarrow \infty$  one has

$$B_{\sigma_k}(u_{\sigma_k}, \phi) \rightarrow B_\infty(u, \phi), \quad k \rightarrow \infty$$

for any  $\phi \in \mathcal{S}_*$ .

(D) *Approximation of nonlocal space and asymptotic density in limiting space:*

- (Di) For given  $\sigma \in \Sigma$  and  $v \in \mathcal{S}_\sigma$ ,  $\inf_{v_{\sigma,h} \in V_{\sigma,h}} \|v - v_{\sigma,h}\|_{V_{\sigma,h}} \rightarrow 0$  as  $h \rightarrow 0$ .
- (Dii) For any  $v \in \mathcal{S}_*$ , there exists  $\{v_k \in V_{\sigma_k, h_k}\}_{h_k \rightarrow 0}^{\sigma_k \rightarrow \infty}$  such that  $\|v - v_k\|_{\mathcal{H}} \rightarrow 0$  as  $k \rightarrow \infty$ .

It should be noted that assumptions (A) and (C) in the above relax the conditions specified in the original paper [35]. Specifically, assumption (A) eliminates the necessity for the continuous embedding of  $\mathcal{S}_\infty$  into  $\mathcal{S}_\sigma$  and assumption (C) substitutes the requirement for strong convergence of operators on a dense subset with the requirement for the convergence of bilinear forms in a weak sense. We refer the readers to a recent work [15] that discusses extensions of the original AC framework to study nonlinear problems.

**Theorem 4.4** (main convergence theorem [35]). *Suppose Assumption 4.3 holds and let  $\{u_\sigma\}$ ,  $u_\infty$  and  $\{u_{\sigma,h}\}$  be the solutions to problems (50), (51), and (52) respectively. Then given properties (A)-(D), we have that for each  $\sigma \in \Sigma$ ,*

$$\|u_\sigma - u_{\sigma,h}\|_{\mathcal{S}_\sigma} \leq \frac{C_2}{C_1} \inf_{v_{\sigma,h} \in V_{\sigma,h}} \|u_\sigma - v_{\sigma,h}\|_{\mathcal{S}_\sigma} \rightarrow 0, \quad h \rightarrow 0, \quad (53)$$

and, moreover,

$$\|u_\sigma - u_\infty\|_{\mathcal{S}_0} \rightarrow 0, \quad \sigma \rightarrow \infty, \quad (54)$$

and

$$\|u_{\sigma,h} - u_\infty\|_{\mathcal{S}_0} \rightarrow 0, \quad \sigma \rightarrow \infty, \quad h \rightarrow 0. \quad (55)$$

*Proof.* We sketch the proof briefly. The first inequality is the best approximation and the convergence follows from the standard conforming Galerkin approximation theory. To

show (54), first note that

$$\sup\{\|u_\sigma\|_{S_\sigma} : \sigma \in \Sigma\} \leq \frac{\|f\|_{S_0}}{C_1}$$

by coercivity (Bii) and embedding of  $S_\sigma$  into  $S_0$ . Then by asymptotically compact embedding property (Aii) we know there exists  $\tilde{u} \in S_0$  such that  $u_\sigma \rightarrow \tilde{u}$  as  $\sigma \rightarrow \infty$  up to a subsequence. Then it suffices to show that  $\tilde{u}$  is the unique solution to (51) to conclude (54). By density of  $S_*$  in  $S$ , it suffices to show (51) holds (with  $\tilde{u}$  in place of  $u_\infty$ ) for any  $v \in S_*$ . Fix some  $v \in S_*$ , since  $S_* \subset S_\sigma$  for each  $\sigma \in \Sigma$ , one knows that

$$B_\sigma(u_\sigma, v) = (f, v)_{S_0}.$$

Letting  $\sigma \rightarrow \infty$  and using (Cii) yield  $B_\infty(\tilde{u}, v) = (f, v)_{S_0}$  as desired.

Finally, to show (55), first use the bound

$$\sup\{\|u_{\sigma,h}\|_{S_\sigma} : \sigma \in \Sigma, h \in (0, h_0)\} \leq \frac{\|f\|_{S_0}}{C_1}$$

and asymptotically compact embedding property to show there exists  $\tilde{u} \in S_0$  such that  $u_{\sigma,h} \rightarrow \tilde{u}$  as  $\sigma \rightarrow \infty, h \rightarrow 0$  up to a subsequence. Then again by density it suffices to show that  $B_\infty(\tilde{u}, v) = (f, v)_{S_0}$  for any  $v \in S_*$ . Fix  $v \in S_*$ , by asymptotic density, there exists  $\{v_k \in V_{\sigma_k, h_k}\}_{h_k \rightarrow 0}^{\sigma_k \rightarrow \infty}$  such that  $\|v - v_k\|_{\mathcal{H}} \rightarrow 0$  as  $k \rightarrow \infty$ . Note that

$$B_{\sigma_k}(u_{\sigma_k, h_k}, v_k) = (f, v_k)_{S_0}.$$

Then the final conclusion follows from the computation below:

$$\begin{aligned} |B_\infty(\tilde{u}, v) - (f, v)_{S_0}| &\leq |B_\infty(\tilde{u}, v) - B_{\sigma_k}(u_{\sigma_k, h_k}, v)| + |B_{\sigma_k}(u_{\sigma_k, h_k}, v - v_k)| + |B_{\sigma_k}(u_{\sigma_k, h_k}, v_k) - (f, v)_{S_0}| \\ &\leq |B_\infty(\tilde{u}, v) - B_{\sigma_k}(u_{\sigma_k, h_k}, v)| + C_2 \|u_{\sigma_k, h_k}\|_{S_{\sigma_k}} \|v - v_k\|_{S_{\sigma_k}} + |(f, v - v_k)_{S_0}| \\ &\leq |B_\infty(\tilde{u}, v) - B_{\sigma_k}(u_{\sigma_k, h_k}, v)| + C_2 \frac{\|f\|_{S_0}}{C_1} M_2 \|v - v_k\|_{\mathcal{H}} + C_{\mathcal{H}} \|f\|_{S_0} \|v - v_k\|_{\mathcal{H}} \\ &\rightarrow 0, \quad k \rightarrow \infty, \end{aligned}$$

where we used (Cii), (Ai) and the embedding inequality  $\|v\|_{S_0} \leq C_{\mathcal{H}} \|v\|_{\mathcal{H}}$  for any  $v \in \mathcal{H}$ . ■

With this abstract result at hand, we are now ready to prove Theorems 4.1 and 4.2.

*Proof of Theorem 4.1.* It suffices to verify assumptions (A)-(D) in Theorem 4.4. Using the notations in Theorem 4.4, we let  $\Sigma = \mathbb{N}_+$ ,  $S_\sigma = S_{w_n}^\nu(\Omega)$ ,  $S_\infty = S_w^\nu(\Omega)$ ,  $S_* = C_c^\infty(\Omega)$  and  $\mathcal{H} = H_0^1(\Omega)$ , where functions in  $H_0^1(\Omega)$  are understood to be in  $H^1(\mathbb{R}^d)$  and vanish outside of  $\Omega$  as the boundary of  $\Omega$  is regular enough [40, Theorem 3.7]. Then  $M_1 = 1$ ,  $C_2 = \|A\|_{L^\infty(\mathbb{R}^d; \mathbb{R}^{d \times d})}$  and Theorem 3.6 verifies (Aii).

The second inequality in (Ai) amounts to showing that there exists  $M_2 > 0$  such that

$$\|u\|_{S_{w_n}^\nu(\Omega)} \leq M_2 \|u\|_{H^1(\mathbb{R}^d)}, \quad \forall u \in H_0^1(\Omega).$$

But this follows from Proposition 2.4 where the norm of  $S_{w_n}^\nu(\Omega)$  can be expressed as

$$\|u\|_{S_{w_n}^\nu(\Omega)}^2 = \|u\|_{S_{w_n}^\nu(\mathbb{R}^d)}^2 = \int_{\mathbb{R}^d} (1 + |\lambda_{w_n}^\nu(\xi)|^2) |\hat{u}(\xi)|^2 d\xi, \quad \forall u \in S_{w_n}^\nu(\Omega).$$

Now using (11), one obtains that

$$|\lambda_{w_n}^\nu(\xi)| \leq 2\sqrt{2\pi} M_{w_n}^1 |\xi| + \sqrt{2} M_{w_n}^2 \leq 2\sqrt{2\pi} M_{w_n}^1 |\xi| + \sqrt{2} M_{w_n}^2, \quad \xi \in \mathbb{R}^d. \quad (56)$$

Thus the second inequality in (Ai) holds true for some  $M_2$  independent of  $n$  but dependent of  $w$ , using the fact that

$$\|u\|_{H^1(\mathbb{R}^d)}^2 = \int_{\mathbb{R}^d} (1 + |\xi|^2) |\hat{u}(\xi)|^2 d\xi.$$

For (Bii), one may choose  $C_1$  independent of  $N$  but dependent of some  $N_0$  determined by Theorem 3.7. Note that (Ci) follows immediately from [20, Theorem 3.2].

To show (Cii), suppose  $v_n \in S_{w_n}^V(\Omega)$ ,  $\sup_{n \in \mathbb{N}_+} \|v_n\|_{S_{w_n}^V(\Omega)} < \infty$ , and  $v_n \rightarrow v$  in  $L^2(\Omega)$  for  $v \in S_w^V(\Omega)$ . We need to prove that for any  $\phi \in C_c^\infty(\Omega)$ ,

$$(A\mathfrak{G}_{w_n}^V v_n, \mathcal{G}_{w_n}^V \phi)_{L^2(\mathbb{R}^d; \mathbb{R}^d)} \rightarrow (A\mathfrak{G}_w^V v, \mathcal{G}_w^V \phi)_{L^2(\mathbb{R}^d; \mathbb{R}^d)}, \quad n \rightarrow \infty. \quad (57)$$

To that end, first, using (2.8) and (2.9) from [20, Lemma 2.1], one may show that  $\mathcal{G}_{w_n}^V \phi \rightarrow \mathcal{G}_w^V \phi$  in  $L^2(\mathbb{R}^d; \mathbb{R}^d)$  as  $n \rightarrow \infty$ , and if  $\psi \in C_c^\infty(\Omega; \mathbb{R}^d)$  then  $D_{w_n}^{-V} \psi \rightarrow D_w^{-V} \psi$  in  $L^2(\mathbb{R}^d)$  as  $n \rightarrow \infty$ . Second, we show that  $\mathfrak{G}_{w_n}^V v_n \rightharpoonup \mathfrak{G}_w^V v$  in  $L^2(\mathbb{R}^d; \mathbb{R}^d)$ . If this holds, then since  $A \in L^\infty(\mathbb{R}^d; \mathbb{R}^{d \times d})$ , it follows that  $A\mathfrak{G}_{w_n}^V v_n \rightharpoonup A\mathfrak{G}_w^V v$  in  $L^2(\mathbb{R}^d; \mathbb{R}^d)$ , hence, combining the first step, one can conclude that (57) holds, and consequently, (Cii) holds. To show  $\mathfrak{G}_{w_n}^V v_n \rightharpoonup \mathfrak{G}_w^V v$  in  $L^2(\mathbb{R}^d; \mathbb{R}^d)$ , fix some  $w \in L^2(\mathbb{R}^d; \mathbb{R}^d)$  and  $\epsilon > 0$ , by density, there exists  $\psi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$  such that  $\|w - \psi\|_{L^2(\mathbb{R}^d; \mathbb{R}^d)} < \epsilon$ . Then

$$\begin{aligned} |(\mathfrak{G}_{w_n}^V v_n, w)_{L^2(\mathbb{R}^d; \mathbb{R}^d)} - (\mathfrak{G}_w^V v, w)_{L^2(\mathbb{R}^d; \mathbb{R}^d)}| &\leq |(\mathfrak{G}_{w_n}^V v_n, w - \psi)_{L^2(\mathbb{R}^d; \mathbb{R}^d)}| \\ &\quad + |(\mathfrak{G}_{w_n}^V v_n, \psi)_{L^2(\mathbb{R}^d; \mathbb{R}^d)} - (\mathfrak{G}_w^V v, \psi)_{L^2(\mathbb{R}^d; \mathbb{R}^d)}| \\ &\quad + |(\mathfrak{G}_w^V v, \psi - w)_{L^2(\mathbb{R}^d; \mathbb{R}^d)}| \\ &\leq \epsilon \sup_{n \in \mathbb{N}_+} \|v_n\|_{S_{w_n}^V(\Omega)} + |(v_n, D_{w_n}^{-V} \psi)_{L^2(\mathbb{R}^d)} - (v, D_w^{-V} \psi)_{L^2(\mathbb{R}^d)}| \\ &\quad + \epsilon \|v\|_{S_w^V(\Omega)}. \end{aligned}$$

Since  $D_{w_n}^{-V} \psi \rightarrow D_w^{-V} \psi$  in  $L^2(\mathbb{R}^d)$ , together with  $v_n \rightarrow v$  in  $L^2(\mathbb{R}^d)$ , one can conclude that  $(v_n, D_{w_n}^{-V} \psi)_{L^2(\mathbb{R}^d)} \rightarrow (v, D_w^{-V} \psi)_{L^2(\mathbb{R}^d)}$  as  $n \rightarrow \infty$  and thus  $\mathfrak{G}_{w_n}^V v_n \rightharpoonup \mathfrak{G}_w^V v$  in  $L^2(\mathbb{R}^d; \mathbb{R}^d)$  and (Cii) holds.

Finally, since  $\mathcal{H} = H_0^1(\Omega)$  and  $\{V_{n,h}\}_{n,h}$  are spaces of piecewise polynomials of degree  $k \geq 1$ , one may show that (D) is satisfied as in [35, Theorem 3.8]. ■

*Proof of Theorem 4.2.* The proof is quite similar to that of Theorem 4.1. We only point out a few key differences. Using the notations in Theorem 4.4, we let  $\Sigma = (\delta_0^{-1}, \infty)$ ,  $S_\sigma = S_{w_{1/\sigma}}^V(\Omega)$ ,  $S_\infty = H_0^1(\Omega)$ ,  $S_* = C_c^\infty(\Omega)$  and  $\mathcal{H} = H_0^1(\Omega)$ . Then Theorem 3.11 verifies (Aii).

For the second inequality in (Ai), one may argue as before for  $\{w_\delta\}$  in Assumption 3.9 by proving

$$|\lambda_{w_\delta}^V(\xi)| \leq D_1 |\xi| + D_2, \quad \forall \xi \in \mathbb{R}^d,$$

where  $D_1, D_2 \geq 0$  are constants independent of  $\delta$ . In fact the inequality is even true for kernels satisfying just Equations (1), (25), and (26) and follows from (11), where

$$|\lambda_{w_\delta}^V(\xi)| \leq 2\sqrt{2\pi} M_{w_\delta}^1 |\xi| + \sqrt{2} M_{w_\delta}^2, \quad \forall \xi \in \mathbb{R}^d$$

and that

$$M_{w_\delta}^1 \rightarrow 2d \quad \text{and} \quad M_{w_\delta}^2 \rightarrow 0 \quad \text{as } \delta \rightarrow 0.$$

Assumption (Bii) can be verified by Theorem 3.14.

To show (Cii), suppose  $v_n \in S_{w_{\delta_n}}^{\nu}(\Omega)$ ,  $\sup_{n \in \mathbb{N}_+} \|v_n\|_{S_{w_{\delta_n}}^{\nu}(\Omega)} < \infty$ , and  $v_n \rightarrow v$  in  $L^2(\Omega)$  for  $v \in H_0^1(\Omega)$ . We need to prove that for any  $\phi \in C_c^{\infty}(\Omega)$  as  $n \rightarrow \infty$  ( $\delta_n \rightarrow 0$ )

$$(A\mathfrak{G}_{w_{\delta_n}}^{\nu} v_n, \mathcal{G}_{w_{\delta_n}}^{\nu} \phi)_{L^2(\mathbb{R}^d; \mathbb{R}^d)} \rightarrow (A\nabla v, \nabla \phi)_{L^2(\mathbb{R}^d; \mathbb{R}^d)}, \quad n \rightarrow \infty. \quad (58)$$

This holds true by the same argument for (57) if one notices that, according to Proposition 3.8,  $\mathcal{G}_{w_{\delta_n}}^{\nu} \phi \rightarrow \nabla \phi$  in  $L^2(\mathbb{R}^d; \mathbb{R}^d)$ , and if  $\psi \in C_c^{\infty}(\Omega; \mathbb{R}^d)$  then  $D_{w_{\delta_n}}^{-\nu} \psi \rightarrow \operatorname{div} \psi$  in  $L^2(\mathbb{R}^d)$  as  $n \rightarrow \infty$ .

The rest of the proof is similar to that of Theorem 4.1 and thus omitted.  $\blacksquare$

## 5 | APPLICATION II. OPTIMAL CONTROL PROBLEM OF A NONLOCAL EQUATION

### 5.1 | Parameterized continuous and discrete optimal control problems

In this section, we apply the compactness results proved in Section 3 to analyze the well-posedness, the numerical approximation, and the limiting behavior of solutions of a parameterized optimal control problem of linear nonlocal equations. Optimal control problems with constraints that involve nonlocal equations have been a focus of recent research interest. While this work is inspired by work [25] which deals with an optimal control problem of linearized peridynamics, we should mention that such types of problems where the state equation is either a fractional or nonlocal equation have been investigated in literature. To cite a few, the paper [14] rigorously analyzed optimal control problems when the state equation is of the form

$$-\mathcal{L}_{\delta}u = g, \quad \text{where} \quad \mathcal{L}_{\delta}u(x) = 2 \int_{\Omega} (u(y) - u(x)) \gamma_{\delta}(x, y) dy \quad (59)$$

subject to some volumetric boundary condition outside of  $\Omega$ . In the above,  $\gamma_{\delta}(x, y)$  serves as a kernel. The paper also presents a finite element approximation and numerical simulations which illustrate the theoretical results. The papers [17, 27] have also studied optimal control problems with the state equation being the parameterized nonlocal equations of the form (59) and demonstrated rigorously the convergence of optimal pairs to an optimal pair of an optimal control problem with a local (PDE) based state equation. Mathematical analysis and numerical approximation of optimal control problems with the fractional equation as the state equation have also been investigated in [1–3, 9, 12, 28]. To put our work in a clear perspective, we should mention that the parameterized state equation constraints we will be dealing with is different from the classical fractional equations and the nonlocal equations of the form (59).

To properly describe the problem we study, we assume that  $\{w_{\delta}\}_{\delta>0}$  is a family of kernels given in Assumption 3.9. We fix  $A \in L^{\infty}(\mathbb{R}^d; \mathbb{R}^{d \times d})$  which is uniformly elliptic and symmetric. That is, for any  $x \in \Omega$ ,  $A(x) = A(x)^T$  and for some  $\mu > 0$  we have

$$\xi^T A(x) \xi \geq \mu |\xi|^2, \quad \forall x \in \mathbb{R}^d, \quad \xi \in \mathbb{R}^d.$$

Given  $\Omega$ , a bounded polygonal domain in  $\mathbb{R}^d$ , and  $\nu \in \mathbb{S}^{d-1}$ , consider the parameterized nonlocal equations

$$\begin{cases} -\mathfrak{D}_{w_{\delta}}^{-\nu}(A\mathfrak{G}_{w_{\delta}}^{\nu} u) = f & \text{in } \Omega, \\ u = 0 & \text{in } \mathbb{R}^d \setminus \Omega, \end{cases} \quad (60)$$

where the right-hand side data  $f$  comes from an admissible class that satisfies a box condition. Given the functions  $\alpha, \beta \in C(\bar{\Omega})$  such that  $\alpha(x) < \beta(x)$  for all  $x \in \Omega$ , we introduce the admissible class of

right-hand side

$$Z_{\text{ad}} = \{g \in L^2(\Omega) : \alpha(\mathbf{x}) \leq g(\mathbf{x}) \leq \beta(\mathbf{x}), \text{ a.e. } \mathbf{x} \in \Omega\}.$$

We remark that there exists a constant  $B > 0$  such that

$$\sup_{g \in Z_{\text{ad}}} \|g\|_{L^2(\Omega)} \leq B, \quad (61)$$

since for any  $g \in Z_{\text{ad}}$ ,  $|g(\mathbf{x})| \leq \max\{|\alpha(\mathbf{x})|, |\beta(\mathbf{x})|\}$  a.e.  $\mathbf{x} \in \Omega$  and  $\Omega$  is bounded. We note that corresponding to each  $f \in L^2(\Omega)$ , a unique solution  $u_\delta(f) \in S_{w_\delta}^v(\Omega)$  exists that solve the weak form of (60)

$$B_\delta(u_\delta, v) = \langle f, v \rangle, \quad \forall v \in S_{w_\delta}^v(\Omega),$$

where  $B_\delta : S_{w_\delta}^v(\Omega) \times S_{w_\delta}^v(\Omega) \rightarrow \mathbb{R}$  is the bilinear form given in (45) with  $w_\delta$  in place of  $w_n$  and  $\langle \cdot, \cdot \rangle$  is the inner product on  $L^2(\Omega)$ . The solution map  $\mathfrak{S}_\delta : L^2(\Omega) \rightarrow S_{w_\delta}^v(\Omega)$  given by  $\mathfrak{S}_\delta(f) = u_\delta(f)$  is a bounded linear map with the estimate

$$|\mathfrak{S}_\delta(f)|_{S_{w_\delta}^v(\Omega)} = |u_\delta(f)|_{S_{w_\delta}^v(\Omega)} \leq \frac{C(\delta_0)}{\mu} \|f\|_{L^2(\Omega)}, \quad \forall f \in L^2(\Omega), \delta \in (0, \delta_0), \quad (62)$$

where the constants  $\delta_0 > 0$  and  $C(\delta_0) > 0$  are chosen as in Theorem 3.14. Indeed, By taking the test function  $v = u_\delta$  in the weak form of (60) and using the ellipticity of  $A$ , we know

$$\mu |u_\delta|_{S_{w_\delta}^v(\Omega)}^2 \leq B_\delta(u_\delta, u_\delta) = \langle f, u_\delta \rangle \leq \|f\|_{L^2(\Omega)} \|u_\delta\|_{L^2(\Omega)}.$$

Combing this with the inequality  $\|u_\delta\|_{L^2(\Omega)} \leq C(\delta_0) |u_\delta|_{S_{w_\delta}^v(\Omega)}$  from Theorem 3.14, one obtains (62) and, in addition, the following  $L^2$  estimate

$$\|\mathfrak{S}_\delta(f)\|_{L^2(\Omega)} \leq \frac{C(\delta_0)^2}{\mu} \|f\|_{L^2(\Omega)}, \quad \forall f \in L^2(\Omega), \delta \in (0, \delta_0). \quad (63)$$

Hence  $\mathfrak{S}_\delta : L^2(\Omega) \rightarrow L^2(\Omega)$  is a bounded linear operator and  $\|\mathfrak{S}_\delta\|_{L^2(\Omega) \rightarrow L^2(\Omega)} \leq C(\delta_0)^2 / \mu$  for all  $\delta \in (0, \delta_0)$ .

As we vary  $f$  in  $Z_{\text{ad}}$ , the goal of the optimal control problem is to find a pair  $(u_\delta(f), f) \in S_{w_\delta}^v(\Omega) \times Z_{\text{ad}}$ , called the *optimal solution pair*, that minimizes a certain objective functional  $I_\delta(u, g) : S_{w_\delta}^v(\Omega) \times Z_{\text{ad}} \rightarrow \mathbb{R}$ . In this case, the right-hand side  $f$  will serve as a *control variable* and  $u$  will be the *state variable* satisfying the state equation (60). Following the work in [25], we consider the objective functional  $I_\delta = I|_{S_{w_\delta}^v(\Omega) \times Z_{\text{ad}}}$  where  $I : L^2(\Omega) \times Z_{\text{ad}} \rightarrow \mathbb{R}$  is given by

$$I(u, g) := \int_{\Omega} F(\mathbf{x}, u(\mathbf{x})) d\mathbf{x} + \frac{\lambda}{2} \int_{\Omega} \Gamma(\mathbf{x}) |g(\mathbf{x})|^2 d\mathbf{x},$$

where  $\lambda > 0$ ,  $\Gamma \in L^1(\Omega)$  is positive and  $F : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$  satisfies the following properties:

- (1) For all  $\xi \in \mathbb{R}$  the mapping  $\mathbf{x} \mapsto F(\mathbf{x}, \xi)$  is measurable.
- (2) For a.e.  $\mathbf{x} \in \Omega$  the mapping  $\xi \mapsto F(\mathbf{x}, \xi)$  is convex (and continuous).
- (3) There exist a constant  $c_1 > 0$  and a function  $\ell \in L^1(\Omega)$  for which

$$|F(\mathbf{x}, \xi)| \leq c_1 |\xi|^2 + \ell(\mathbf{x})$$

for a.e.  $\mathbf{x} \in \Omega$  and all  $\xi \in \mathbb{R}$ .

- (4) There exist two functions  $c \in L^1(\Omega)$  and  $d \in L^2(\Omega)$  such that

$$F(\mathbf{x}, \xi) \geq c(\mathbf{x}) + d(\mathbf{x})\xi$$

for a.e.  $\mathbf{x} \in \Omega$  and all  $\xi \in \mathbb{R}$ .

(5)  $F$  is continuously differentiable in the second argument. That is, for a.e.  $x \in \Omega$ ,  $F_\xi(x, \cdot) \in C^0(\mathbb{R})$  where  $F_\xi(x, \xi) := \frac{\partial F}{\partial \xi}(x, \xi)$  for  $x \in \Omega$  and  $\xi \in \mathbb{R}$ . Moreover, there exist a constant  $c_2 > 0$  and a function  $e \in L^2(\Omega)$  such that

$$|F_\xi(x, \xi)| \leq c_2|\xi| + e(x), \quad \text{a.e. } x \in \Omega, \forall \xi \in \mathbb{R}. \quad (64)$$

A typical objective functional is  $I(u, g) = \|u - u_{\text{des}}\|_{L^2(\Omega)}^2 + \frac{\lambda}{2}\|g\|_{L^2}^2$  where we are looking for an optimal pair  $(u, g)$  where the control has the smallest  $L^2(\Omega)$ -norm and the state  $u$  is the closest to a given desired state  $u_{\text{des}} \in L^2(\Omega)$ .

There are now a number of outstanding questions we would like to address in this section. First, we would like to address whether the problem is well-posed, that is, whether there is a unique optimal pair  $(\bar{u}_\delta, \bar{g}_\delta)$  to the optimal control problem. Second, we would like to study the asymptotic behavior, with respect to appropriate topology, of the sequence of optimal pairs  $(\bar{u}_\delta, \bar{g}_\delta)$  as  $\delta \rightarrow 0$ , and if there is a limit point  $(u, f)$ , we would like to determine whether it solves an optimal control problem. And finally, we would like to study the discretization of the parameterized optimal control problems and establish the existence of asymptotically compatible schemes.

The following result states the well-posedness of the parameterized optimal control problem as well as the convergence of solutions. It is one of the main results of this section and its proof is deferred for later.

**Theorem 5.1** (continuous problems and convergence of solutions). *Assume all the conditions we stated at the beginning of this section on  $\Omega$ , the sequence of kernels  $\{w_\delta\}_{\delta>0}$ , the control admissible set  $Z_{\text{ad}}$ , and the objective functional  $I$ . Then for each  $\delta > 0$ , there exists a unique optimal pair  $(\bar{u}_\delta, \bar{g}_\delta) \in S_{w_\delta}^\nu(\Omega) \times Z_{\text{ad}}$  such that*

$$I(\bar{u}_\delta, \bar{g}_\delta) = \min I(u_\delta, g), \quad (65)$$

where the minimization is over pairs  $(u_\delta, g) \in S_{w_\delta}^\nu(\Omega) \times Z_{\text{ad}}$  that satisfy

$$B_\delta(u_\delta, v) = \langle g, v \rangle, \quad \forall v \in S_{w_\delta}^\nu(\Omega). \quad (66)$$

Moreover, there exists a unique pair  $(\bar{u}, \bar{g}) \in H_0^1(\Omega) \times Z_{\text{ad}}$  such that  $\bar{u}_\delta \rightarrow \bar{u}$  in  $L^2(\Omega)$  and  $\bar{g}_\delta \rightarrow \bar{g}$  in  $L^2(\Omega)$  and  $(\bar{u}, \bar{g})$  solves

$$I(\bar{u}, \bar{g}) = \min I(u, g), \quad (67)$$

where the minimization is over pairs  $(u, g) \in H_0^1(\Omega) \times Z_{\text{ad}}$  that satisfy

$$B_0(u, v) = \langle g, v \rangle, \quad \forall v \in H_0^1(\Omega). \quad (68)$$

Here  $B_0(u, v) := \int_{\Omega} A(x) \nabla u(x) \cdot \nabla v(x) dx$ .

To state the corresponding result for discretized problems, let us introduce some notations. Let  $\{\mathcal{T}_h\}_{h>0}$  be a quasi-uniform mesh of size  $h$  on  $\Omega$ . Let  $X_h$  be the space of continuous piecewise linear functions subject to the mesh with zero nonlocal boundary data:

$$X_h := \{w_h \in C^0(\bar{\Omega}) \mid w_h|_T \in \mathcal{P}_1(T), \forall T \in \mathcal{T}_h, w_h = 0 \text{ on } \mathbb{R}^d \setminus \Omega\}.$$

We equip  $X_h$  with  $H^1(\Omega)$ -norm. For the nonlocal discrete problem, we use the space of piecewise linear functions that are in  $S_{w_\delta}^\nu(\Omega)$ :

$$X_{\delta, h} = \{w_h \in S_{w_\delta}^\nu(\Omega) \mid w_h|_T \in \mathcal{P}_1(T), \forall T \in \mathcal{T}_h\}$$

equipped with  $S_{w_\delta}^v(\Omega)$ -norm. Similarly, let  $Z_h$  denote the piecewise constant functions with respect to the mesh  $\{\mathcal{T}_h\}_{h>0}$ , that is,

$$Z_h := \{z_h \in L^\infty(\Omega) \mid z_h|_T \in \mathcal{P}_0(T), \forall T \in \mathcal{T}_h\}.$$

where  $\mathcal{P}_m(T)$ , as before, is the space of polynomials of degree  $m$ . Notice that since  $\alpha(\mathbf{x}) < \beta(\mathbf{x})$  for all  $\mathbf{x} \in \Omega$ , there exists a constant  $h_0 > 0$  such that  $Z_h \cap Z_{\text{ad}} \neq \emptyset$  for any  $h \in (0, h_0)$ . Hereafter, we will always assume  $h \in (0, h_0)$  implicitly.

**Theorem 5.2** (Discrete problems and convergence of solutions). *Assume all the conditions used in Theorem 5.1. Then for any  $\delta > 0$ , there exists a unique optimal pair  $(\overline{u_{\delta,h}}, \overline{g_{\delta,h}}) \in X_{\delta,h} \times (Z_h \cap Z_{\text{ad}})$  such that*

$$I(\overline{u_{\delta,h}}, \overline{g_{\delta,h}}) = \min I(u_{\delta,h}, g_{\delta,h}) \quad (69)$$

where the minimization is over pairs  $(u_{\delta,h}, g_h) \in X_{\delta,h} \times (Z_h \cap Z_{\text{ad}})$  that satisfy

$$B_\delta(u_{\delta,h}, v_{\delta,h}) = \langle g_h, v_{\delta,h} \rangle, \quad \forall v_{\delta,h} \in X_{\delta,h}. \quad (70)$$

Moreover, there is a unique pair  $(\overline{u_h}, \overline{g_h}) \in X_h \times (Z_h \cap Z_{\text{ad}})$  such that  $\overline{u_h} \rightarrow \overline{u_h}$  in  $L^2(\Omega)$ ,  $\overline{g_h} \rightarrow \overline{g_h}$  in  $L^2(\Omega)$ , as  $\delta \rightarrow 0$ , and in addition,  $(\overline{u_h}, \overline{g_h})$  solves the local discrete optimal control

$$I(\overline{u_h}, \overline{g_h}) = \min I(u_h, g_h) \quad (71)$$

where the minimization is over pairs  $(u_h, g_h) \in X_h \times (Z_h \cap Z_{\text{ad}})$  that satisfy

$$B_0(u_h, v_h) = \langle g_h, v_h \rangle, \quad \forall v_h \in X_h. \quad (72)$$

Before we present the proofs of Theorems 5.1 and 5.2, we note that applying the nonlocal and local Poincaré inequality and Lax–Milgram theorem, we know the state Equations (66), (68), (70) and (72) are uniquely solvable in their corresponding energy spaces. Like the continuous case, for Equation (70) we also introduce the discrete solution operator  $\mathfrak{S}_{\delta,h} : L^2(\Omega) \rightarrow X_{\delta,h} \subset S_{w_\delta}^v(\Omega) \subset L^2(\Omega)$ . One may check the same uniform estimates (62) and (63) hold for  $\mathfrak{S}_{\delta,h}$  for all  $\delta \in (0, \delta_0)$  and  $h \in (0, h_0)$ . Since the techniques to prove the well-posedness of the four minimization problems as well as the convergence results are essentially the same for the two theorems, we present the proof of Theorem 5.1 and omit the proof of Theorem 5.2.

*Proof of Theorem 5.1.* We will prove the theorem in several steps.

**Well-posedness:** We show existence of a minimizer to the objective functional subject to the nonlocal constraint using the direct method of calculus of variations. Using the solution operator  $\mathfrak{S}_\delta$ , we will work on the reduced cost functional  $j(g) = I(\mathfrak{S}_\delta(g), g)$  given by

$$j(g) := \mathfrak{F}(\mathfrak{S}_\delta(g)) + \frac{\lambda}{2} \int_{\Omega} \Gamma(\mathbf{x}) |g(\mathbf{x})|^2 d\mathbf{x}, \quad (73)$$

where  $\mathfrak{F} : L^2(\Omega) \rightarrow \mathbb{R}$  is defined by

$$\mathfrak{F}(v) := \int_{\Omega} F(\mathbf{x}, v(\mathbf{x})) d\mathbf{x}.$$

We apply the direct method of calculus of variations to the problem of finding a minimizer to  $\inf_{g \in Z_{\text{ad}}} j(g)$ .

We first notice that  $Z_{\text{ad}}$  is a closed, convex and bounded subset of Hilbert space  $L^2(\Omega)$ , by [37, Theorem 2.11]  $Z_{\text{ad}}$  is weakly sequentially compact. In addition,  $j$  is bounded from

below on  $Z_{\text{ad}}$ . Indeed, it suffices to show that  $\mathfrak{F} \circ \mathfrak{S}_\delta : Z_{\text{ad}} \rightarrow \mathbb{R}$  is bounded from below, since the second term in  $j$  is nonnegative. To that end, using item (3) of the assumption on  $F$ , (61) and (63), we have that for any  $g \in Z_{\text{ad}}$ ,

$$\begin{aligned}\mathfrak{F}(\mathfrak{S}_\delta(g)) &\geq -c_1 \|\mathfrak{S}_\delta(g)\|_{L^2(\Omega)}^2 - \|\ell\|_{L^1(\Omega)} \\ &\geq -c_1 \frac{C(\delta_0)^4}{\mu^2} \|g\|_{L^2(\Omega)}^2 - \|\ell\|_{L^1(\Omega)} \geq -c_1 B^2 \frac{C(\delta_0)^4}{\mu^2} - \|\ell\|_{L^1(\Omega)}.\end{aligned}$$

We henceforth denote  $j_0 = \inf_{g \in Z_{\text{ad}}} j(g)$ . It remains to show that there exists  $\tilde{g} \in Z_{\text{ad}}$  such that  $j_0 = j(\tilde{g})$ . To this end, first we find a sequence  $\{g_n\}_{n=1}^\infty \subset Z_{\text{ad}}$  such that  $\lim_{n \rightarrow \infty} j(g_n) = j_0$ . Again by [37, Theorem 2.11] there exists a subsequence  $\{g_{n_k}\}_{k=1}^\infty \subset Z_{\text{ad}}$  converging to some  $\tilde{g} \in Z_{\text{ad}}$  weakly in  $L^2(\Omega)$ . Since  $|\sqrt{\Gamma(x)}g_{n_k}(x)| \leq |\sqrt{\Gamma(x)}| \min\{|a(x)|, |b(x)|\}$ ,  $\{\sqrt{\Gamma(x)}g_{n_k}(x)\}_{k=1}^\infty$  is uniformly bounded in  $L^2(\Omega)$ . Then one can further obtain a subsequence, still denoted by  $\{g_{n_k}\}_{k=1}^\infty$ , converging weakly in  $L^2(\Omega)$ . By a density argument, it is not hard to show this weak limit is  $\sqrt{\Gamma(x)}\tilde{g}(x)$ . Also, since  $F(x, \cdot)$  is convex in  $\mathbb{R}$  a.e.  $x \in \Omega$  and  $F$  is bounded from below by an affine map according to items (2) and (4) of the assumption on  $F$ , [18, Theorem 6.54] implies that  $\mathfrak{F}$  is sequentially weakly lower semicontinuous. Moreover, since  $\mathfrak{S}_\delta$  is continuous on  $L^2(\Omega)$ , it is weakly continuous on  $L^2(\Omega)$  and therefore,  $\mathfrak{F} \circ \mathfrak{S}_\delta : L^2(\Omega) \rightarrow \mathbb{R}$  is also sequentially weakly lower semicontinuous on  $L^2(\Omega)$ . Combining the above, we have that

$$\begin{aligned}j(\tilde{g}) &= \mathfrak{F}(\mathfrak{S}_\delta \tilde{g}) + \frac{\lambda}{2} \int_{\Omega} \Gamma(x) |\tilde{g}(x)|^2 dx \\ &\leq \liminf_{k \rightarrow \infty} \mathfrak{F}(\mathfrak{S}_\delta g_{n_k}) + \liminf_{k \rightarrow \infty} \frac{\lambda}{2} \int_{\Omega} \Gamma(x) |g_{n_k}(x)|^2 dx \\ &\leq \liminf_{k \rightarrow \infty} \left( \mathfrak{F}(\mathfrak{S}_\delta g_{n_k}) + \frac{\lambda}{2} \int_{\Omega} \Gamma(x) |g_{n_k}(x)|^2 dx \right) = \lim_{k \rightarrow \infty} j(g_{n_k}) = j_0.\end{aligned}$$

Then the existence of minimizer is proved. The uniqueness of the optimal pair follows from the fact that  $j$  is strictly convex because  $\mathfrak{F}$  is convex and  $\lambda > 0$ .

**Compactness:** Having shown that, for every horizon  $\delta > 0$ , the optimal control problem has a unique solution  $(\bar{u}_\delta, \bar{g}_\delta)$ , we now study the behavior of the pair as  $\delta \rightarrow 0$ . Since  $\bar{u}_\delta = \mathfrak{S}_\delta \bar{g}_\delta$ , by (62) and (63) one obtains

$$|\bar{u}_\delta|_{S_{w_\delta}^v(\Omega)} \leq \frac{C(\delta_0)}{\mu} \|\bar{g}_\delta\|_{L^2(\Omega)} \quad \text{and} \quad \|\bar{u}_\delta\|_{L^2(\Omega)} \leq \frac{C(\delta_0)^2}{\mu} \|\bar{g}_\delta\|_{L^2(\Omega)}.$$

Therefore, using (61) we have

$$\sup_{\delta \in (0, \delta_0)} \|\bar{u}_\delta\|_{S_{w_\delta}^v(\Omega)} \leq \sqrt{1 + C(\delta_0)^2} \frac{C(\delta_0)}{\mu} B.$$

By Theorem 3.11 one concludes that, up to a subsequence,  $\bar{u}_\delta$  converges strongly in  $L^2(\Omega)$  to some  $u \in H_0^1(\Omega)$ . Also, since  $Z_{\text{ad}}$  is weakly sequentially compact, then there exists  $g \in Z_{\text{ad}}$  such that  $\bar{g}_\delta \rightharpoonup g$  in  $L^2(\Omega)$  up to a subsequence. Without loss of generality, we assume  $\bar{u}_\delta \rightarrow u$  in  $L^2(\Omega)$  and  $\bar{g}_\delta \rightarrow g$  in  $L^2(\Omega)$ .

**Convergence:** Next we show that  $(u, g)$  solves

$$B_0(u, v) = \langle g, v \rangle, \quad \forall v \in H_0^1(\Omega).$$

By density, it suffices to show this for all  $v \in C_c^\infty(\Omega)$ . Fix  $v \in C_c^\infty(\Omega)$ . On the one hand, due to the weak convergence  $\bar{g}_\delta \rightharpoonup g$  in  $L^2(\Omega)$  it holds that  $\langle \bar{g}_\delta, v \rangle \rightarrow \langle g, v \rangle$  as  $\delta \rightarrow 0$ . On

the other hand, arguing as in the proof of (Cii) for Theorem 4.2 one knows

$$(A\mathfrak{G}_{w_\delta}^v \bar{u}_\delta, \mathfrak{G}_{w_\delta}^v v)_{L^2(\mathbb{R}^d; \mathbb{R}^d)} \rightarrow (A\nabla u, \nabla v)_{L^2(\mathbb{R}^d; \mathbb{R}^d)}, \quad \delta \rightarrow 0. \quad (74)$$

Therefore, letting  $\delta \rightarrow 0$  in

$$B_\delta(\bar{u}_\delta, v) = \langle \bar{g}_\delta, v \rangle$$

finishes the proof of convergence.

**Limit point is the optimal solution:** Finally, we show that the limit point pair  $(u, g)$  is an optimal pair in the sense that

$$I(u, g) \leq I(v, f),$$

where  $(v, f) \in H_0^1(\Omega) \times Z_{ad}$  solves  $B_0(v, \phi) = \langle f, \phi \rangle$ ,  $\forall \phi \in H_0^1(\Omega)$ . Given such a pair  $(v, f) \in H_0^1(\Omega) \times Z_{ad}$ , we denote  $v_\delta := \mathfrak{G}_\delta f \in S_{w_\delta}^v(\Omega)$  for  $\delta \in (0, \delta_0)$ . Since  $\|v_\delta\|_{S_{w_\delta}^v(\Omega)} \leq C\|f\|_{L^2(\Omega)}$ , by compactness, Theorem 3.11, as  $\delta \rightarrow 0$ , up to a subsequence,  $v_\delta$  strongly converges in  $L^2(\Omega)$  to some function belonging to  $H_0^1(\Omega)$ , and applying a similar argument as in the previous step, we can in fact show that  $v_\delta \rightarrow v$  in  $L^2(\Omega)$ . By [8, Theorem 4.9] we may take a further subsequence, such that for some  $h(\mathbf{x}) \in L^2$ ,  $|v_\delta(\mathbf{x})| \leq h(\mathbf{x})$  for all  $\delta$ , and  $v_\delta \rightarrow v$  almost everywhere. We then have by item (3) of the assumption on  $F$  that

$$|F(\mathbf{x}, v_\delta(\mathbf{x}))| \leq C|h(\mathbf{x})|^2 + \ell(\mathbf{x}) \text{ and } F(\mathbf{x}, v_\delta(\mathbf{x})) \rightarrow F(\mathbf{x}, v(\mathbf{x})) \quad \text{a.e. } \mathbf{x} \in \Omega.$$

By Lebegue dominated convergence theorem, we have

$$\lim_{\delta \rightarrow 0} \int_{\Omega} F(\mathbf{x}, v_\delta(\mathbf{x})) d\mathbf{x} = \int_{\Omega} F(\mathbf{x}, v(\mathbf{x})) d\mathbf{x}.$$

Thus we have

$$\lim_{\delta \rightarrow 0} I(v_\delta, f) = I(v, f).$$

Recall that  $\bar{g}_\delta \rightharpoonup g$  in  $L^2(\Omega)$ . Using a density argument as before, one can show that  $\sqrt{\Gamma \bar{g}_\delta} \rightharpoonup \sqrt{\Gamma g}$  in  $L^2(\Omega)$  as  $\delta \rightarrow 0$ . Then it follows that

$$\int_{\Omega} \Gamma(\mathbf{x}) |g(\mathbf{x})|^2 d\mathbf{x} \leq \liminf_{\delta \rightarrow 0} \int_{\Omega} \Gamma(\mathbf{x}) |\bar{g}_\delta(\mathbf{x})|^2 d\mathbf{x}$$

and

$$\begin{aligned} I(u, g) &= \int_{\Omega} F(\mathbf{x}, u(\mathbf{x})) d\mathbf{x} + \frac{\lambda}{2} \int_{\Omega} \Gamma(\mathbf{x}) |g(\mathbf{x})|^2 d\mathbf{x} \\ &\leq \liminf_{\delta \rightarrow 0} \left( \int_{\Omega} F(\mathbf{x}, \bar{u}_\delta(\mathbf{x})) d\mathbf{x} + \frac{\lambda}{2} \int_{\Omega} \Gamma(\mathbf{x}) |\bar{g}_\delta(\mathbf{x})|^2 d\mathbf{x} \right) \\ &= \liminf_{\delta \rightarrow 0} I(\bar{u}_\delta, \bar{g}_\delta), \end{aligned}$$

as  $\bar{u}_\delta \rightarrow u$  in  $L^2(\Omega)$  and  $\bar{g}_\delta \rightharpoonup g$  in  $L^2(\Omega)$ . Since  $(\bar{u}_\delta, \bar{g}_\delta)$  is a minimizer of  $I$  subject to the nonlocal state equation, we have that  $I(\bar{u}_\delta, \bar{g}_\delta) \leq I(v_\delta, f)$ . Combining all the above inequalities yields

$$I(u, g) \leq \liminf_{\delta \rightarrow 0} I(\bar{u}_\delta, \bar{g}_\delta) \leq \lim_{\delta \rightarrow 0} I(v_\delta, f) = I(v, f).$$

Having shown now that  $(u, g)$  is an optimal solution we will use the bar notation and write  $(\bar{u}, \bar{g})$ . This finishes the proof. ■

## 5.2 | First order optimality and asymptotic compatibility

As we have seen in the proof of Theorem 5.1, one can rewrite optimal control problem (65)–(66) as a variational problem involving the reduced objective functional  $j(g)$  defined in (73): find  $\overline{g}_\delta \in Z_{\text{ad}}$  such that

$$j(\overline{g}_\delta) = \min_{g \in Z_{\text{ad}}} j(g). \quad (75)$$

It is clear that if  $\overline{g}_\delta$  is the unique solution to (75), then  $(\overline{u}_\delta, \overline{g}_\delta)$  is the unique solution to (65)–(66) where  $\overline{u}_\delta := \mathfrak{S}_\delta \overline{g}_\delta$ , and conversely, if  $(\overline{u}_\delta, \overline{g}_\delta)$  is the unique solution to (65)–(66) then  $\overline{g}_\delta$  is the unique solution to (75).

By [37, Lemma 2.21], the first order optimality condition of the minimization problem (75) is

$$\langle j'(\overline{g}_\delta), q - \overline{g}_\delta \rangle \geq 0, \quad \forall q \in Z_{\text{ad}}.$$

By calculating  $j'(\overline{g}_\delta)$  the optimality condition can be written as

$$\langle \mathfrak{S}_\delta^* F_\xi(\cdot, \mathfrak{S}_\delta \overline{g}_\delta(\cdot)) + \lambda \Gamma \overline{g}_\delta, q - \overline{g}_\delta \rangle \geq 0, \quad \forall q \in Z_{\text{ad}}. \quad (76)$$

where  $\mathfrak{S}_\delta^* : L^2(\Omega) \rightarrow L^2(\Omega)$  is the adjoint of  $\mathfrak{S}_\delta$  in the  $L^2$ -sense. It is immediate that  $\mathfrak{S}_\delta$  is self-adjoint. Also, the map  $F_\xi(\cdot, \mathfrak{S}_\delta \overline{g}_\delta(\cdot)) := (x \mapsto F_\xi(x, \mathfrak{S}_\delta \overline{g}_\delta(x))) \in L^2(\Omega)$  so that  $\mathfrak{S}_\delta^*$  can act on it. Indeed, by assumption  $|F_\xi(x, \xi)| \leq C|\xi| + e(x)$ , one has

$$|F_\xi(x, \mathfrak{S}_\delta \overline{g}_\delta(x))| \leq C|\mathfrak{S}_\delta \overline{g}_\delta(x)| + e(x) \in L^2(\Omega).$$

By introducing a new notation  $\overline{p}_\delta$ , the above equation can be rewritten as the system

$$\begin{cases} \langle \overline{p}_\delta + \lambda \Gamma \overline{g}_\delta, q - \overline{g}_\delta \rangle \geq 0, & \forall q \in Z_{\text{ad}}, \\ \overline{p}_\delta = \mathfrak{S}_\delta^* F_\xi(\cdot, \overline{u}_\delta(\cdot)), \\ \overline{u}_\delta = \mathfrak{S}_\delta \overline{g}_\delta. \end{cases} \quad (77)$$

Note that if  $\Gamma = 1$ , then the first inequality amounts to the following identity

$$\overline{g}_\delta = -\frac{1}{\lambda} \Pi_{Z_{\text{ad}}} \overline{p}_\delta, \quad (78)$$

where  $\Pi_E : L^2(\Omega) \rightarrow E$  denotes the  $L^2$ -projection onto the bounded, convex and closed set  $E \subset L^2(\Omega)$ , that is, for any  $f \in L^2(\Omega)$ ,  $\Pi_E f$  is the unique solution to the minimization problem

$$\min_{g \in E} \|g - f\|_{L^2(\Omega)}.$$

Since  $\mathfrak{S}_\delta$  is self-adjoint, one obtains that  $\overline{p}_\delta = \mathfrak{S}_\delta F_\xi(\cdot, \overline{u}_\delta(\cdot))$ .

Note that the objective functional is strictly convex, the first order necessary condition is also sufficient. Therefore, we have the following proposition.

**Proposition 5.3** (optimality conditions). *For every  $\delta > 0$ , the pair  $(\overline{u}_\delta, \overline{g}_\delta) \in S_{W_\delta}^V(\Omega) \times Z_{\text{ad}}$  is a solution to (65)–(66) if and only if (77) holds.*

The optimality conditions for the nonlocal discrete problem read:

$$\begin{cases} \langle \overline{p}_{\delta,h} + \lambda \Gamma \overline{g}_{\delta,h}, q_h - \overline{g}_{\delta,h} \rangle \geq 0, & \forall q_h \in Z_{\text{ad}} \cap Z_h, \\ \overline{p}_{\delta,h} = \mathfrak{S}_{\delta,h} F_\xi(\cdot, \overline{u}_{\delta,h}(\cdot)), \\ \overline{u}_{\delta,h} = \mathfrak{S}_{\delta,h} \overline{g}_{\delta,h}, \end{cases} \quad (79)$$

where  $\mathfrak{S}_{\delta,h} : L^2(\Omega) \rightarrow L^2(\Omega)$  is the discrete solution operator, and  $\mathfrak{S}_{\delta,h}^* : L^2(\Omega) \rightarrow L^2(\Omega)$  is its  $L^2$ -adjoint operator. Note that in the second equation we used the fact that  $\mathfrak{S}_{\delta,h}$  is self-adjoint in the  $L^2$  sense.

Similarly, the continuous local problem is well-posed if and only if the corresponding optimality conditions hold:

$$\begin{cases} \langle \bar{p} + \lambda \Gamma \bar{g}, q - \bar{g} \rangle \geq 0, & \forall q \in Z_{\text{ad}}, \\ \bar{p} = \mathfrak{S}_0 F_\xi(\cdot, \bar{u}(\cdot)), \\ \bar{u} = \mathfrak{S}_0 \bar{g}, \end{cases} \quad (80)$$

where  $\mathfrak{S}_0 : L^2(\Omega) \rightarrow L^2(\Omega)$  is the solution operator defined as  $\mathfrak{S}_0 g := u$  where  $u \in H_0^1(\Omega) \subset L^2(\Omega)$  is the unique solution to (68). Here again we used  $\mathfrak{S}_0$  is self-adjoint in the second equation.

Finally we state and prove the results for asymptotic compatibility. We first recall [25, Definition 7.1] which is the definition of asymptotic compatibility of a scheme to the optimal control problem.

**Definition 5.4** (asymptotic compatibility). We say that the family of solutions  $\{(\bar{u}_{\delta,h}, \bar{g}_{\delta,h})\}_{h>0, \delta>0}$  to (69) is **asymptotically compatible** in  $\delta, h > 0$  if for any sequences  $\{\delta_k\}_{k=1}^\infty, \{h_k\}_{k=1}^\infty$  with  $\delta_k, h_k \rightarrow 0$ , we have that  $\bar{u}_{\delta_k, h_k} \rightarrow \bar{u}$  strongly in  $L^2(\Omega)$  and  $\bar{g}_{\delta_k, h_k} \rightarrow \bar{g}$  weakly in  $L^2(\Omega)$ . Here  $(\bar{u}, \bar{g}) \in H_0^1(\Omega) \times Z_{\text{ad}}$  denotes the optimal solution for (71).

Recall that in the proof of Theorem 4.2 we have checked the conditions (A)–(D) hold to prove asymptotic compatibility for parameterized problems. It turns out that for the optimal control (69) we can establish similar asymptotic compatibility as well in the sense of Definition 5.4. Such type of result is proved in [25, Theorem 7.3]. We demonstrate below that the parameterized discrete optimal control problem, under the constraint of the nonlocal state equation we consider here, has similar compatibility behavior. The following is the main result of the subsection. Its proof is similar to that of [25, Theorem 7.3] with appropriate modification to fit our setting.

**Theorem 5.5** (asymptotic compatibility). *The solution to (69) is asymptotically compatible in  $\delta, h > 0$ , in the sense of Definition 5.4.*

*Proof.* We denote  $\{(\bar{u}_k, \bar{g}_k, \bar{p}_k)\}_{k=1}^\infty := \{(\bar{u}_{\delta_k, h_k}, \bar{g}_{\delta_k, h_k}, \bar{p}_{\delta_k, h_k})\}_{k=1}^\infty$ , which is the sequence of triples solving (79). We consider an arbitrary, non-relabeled subsequence of the triples  $\{(\bar{u}_k, \bar{g}_k, \bar{p}_k)\}_{k=1}^\infty$ , and show that it has a further subsequence which always converges to the same limit point. Moreover, this limit solves (80) and, since this uniquely characterizes the solution to (71), asymptotic compatibility will follow.

Since  $\{\bar{g}_k\}_{k=1}^\infty \subset Z_{\text{ad}}$ , again by [37, Theorem 2.11] there exists a subsequence (non-relabeled) and a function  $g_* \in Z_{\text{ad}}$  such that  $\bar{g}_k \rightharpoonup g_*$  in  $L^2(\Omega)$ . Meanwhile, using (61), (62), and (63) with  $\mathfrak{S}_{\delta_k, h_k}$  in place of  $\mathfrak{S}_\delta$  we know there exists a constant  $C_1 := \sqrt{1 + C(\delta_0)^2} C(\delta_0) B / \mu$  such that  $\|\bar{u}_k\|_{S_{\mathfrak{w}_{\delta_k}}(\Omega)} \leq C_1$  for all  $k$  large enough. By Theorem 3.11, upon taking a further non-relabeled subsequence, there exists a limit point  $u_* \in H_0^1(\Omega)$  such that  $\bar{u}_k \rightarrow u_*$  in  $L^2(\Omega)$ . Since  $\{(\bar{u}_k, \bar{g}_k)\}_{k=1}^\infty$  are the pair of functions solving (69), we have that

$$B_{\delta_k}(\bar{u}_k, v_k) = \langle \bar{g}_k, v_k \rangle, \quad \forall v_k \in X_{\delta_k, h_k}.$$

Arguing as in the last paragraph of the proof of Theorem 4.4, with a slight modification that  $(f, v_k)_{L^2}$  and  $(f, v)_{L^2}$  are replaced by  $\langle \bar{g}_k, v_k \rangle$  and  $\langle g_*, v \rangle$  respectively, one can show that

$$B_0(u_*, v) = \langle g_*, v \rangle, \quad \forall v \in H_0^1(\Omega). \quad (81)$$

Recall that  $\bar{p}_k$  satisfies the second equation in (79), that is,

$$B_{\delta_k}(\bar{p}_k, v_k) = \langle F_\xi(\cdot, \bar{u}_k(\cdot)), v_k \rangle, \quad \forall v_k \in X_{\delta_k, h_k}.$$

Since

$$\|F_\xi(\cdot, \bar{u}_k(\cdot))\|_{L^2(\Omega)}^2 \leq 2c_2^2 \|\bar{u}_k\|_{L^2(\Omega)}^2 + 2\|e\|_{L^2(\Omega)}^2 \leq 2c_2^2 C_1^2 + 2\|e\|_{L^2(\Omega)}^2 =: C_2,$$

repeating the argument above for the uniform bound on  $\|\bar{u}_k\|_{S_{w,\delta_k}^\nu(\Omega)}$ , we have  $\|\bar{p}_k\|_{S_{w,\delta_k}^\nu(\Omega)} \leq C_1 \|F_\xi(\cdot, \bar{u}_k(\cdot))\|_{L^2(\Omega)}/B \leq C_1 \sqrt{C_2}/B$  for all  $k$  large enough. Moreover, up to a subsequence, using [8, Theorem 4.9], item (5) of the assumption on  $F$ , and the Lebesgue dominated convergence theorem, we have that as  $k \rightarrow \infty$ ,

$$F_\xi(\cdot, \bar{u}_k(\cdot)) \rightarrow F_\xi(\cdot, u_*(\cdot)) \text{ in } L^2(\Omega).$$

Repeating the analysis for  $\{\bar{u}_k\}_{k=1}^\infty$ , we identify  $p_* \in H_0^1(\Omega)$  such that up to a subsequence  $\bar{p}_k \rightarrow p_*$  in  $L^2(\Omega)$  and

$$B_0(p_*, v) = \langle F_\xi(\cdot, u_*(\cdot)), v \rangle, \quad \forall v \in H_0^1(\Omega). \quad (82)$$

Finally we show that

$$\langle p_* + \lambda \Gamma g_*, q - g_* \rangle \geq 0, \quad \forall q \in Z_{\text{ad}}. \quad (83)$$

Since  $\alpha, \beta \in C(\bar{\Omega})$ , it is not hard to show the following approximation result: for any  $q \in Z_{\text{ad}}$ , there exists a sequence  $\{q_k \in Z_{h_k} \cap Z_{\text{ad}}\}_{h_k>0}$  such that  $q_k \rightarrow q$  in  $L^2(\Omega)$  as  $k \rightarrow \infty$ ; see Lemma A.3 for its proof. From the first equation of (79) we know

$$\langle \bar{p}_k + \lambda \Gamma \bar{g}_k, q_k - \bar{g}_k \rangle \geq 0.$$

That is,

$$\langle \bar{p}_k, q_k - \bar{g}_k \rangle + \langle \lambda \Gamma \bar{g}_k, q_k \rangle \geq \lambda \int_{\Omega} \Gamma(x) |\bar{g}_k(x)|^2 dx.$$

By the weak lower semicontinuity, one has

$$\liminf_{k \rightarrow \infty} \int_{\Omega} \Gamma(x) |\bar{g}_k(x)|^2 dx \geq \int_{\Omega} \Gamma(x) |g_*(x)|^2 dx.$$

On the other hand, letting  $k \rightarrow \infty$  one has

$$\langle \bar{p}_k, q_k - \bar{g}_k \rangle + \langle \lambda \Gamma \bar{g}_k, q_k \rangle \rightarrow \langle p_*, q - g_* \rangle + \langle \lambda \Gamma g_*, q \rangle.$$

Therefore,

$$\langle p_*, q - g_* \rangle + \langle \lambda \Gamma g_*, q \rangle \geq \lambda \int_{\Omega} \Gamma(x) |g_*(x)|^2 dx,$$

and (83) holds. From (83), (82), and (81) we know that  $(u_*, g_*, p_*)$  solves the equations (80). Since this system has a unique triple of solution  $(\bar{u}, \bar{g}, \bar{p})$ , it follows that  $(u_*, g_*, p_*) = (\bar{u}, \bar{g}, \bar{p})$ . Therefore, asymptotic compatibility holds. ■

## 6 | CONCLUSION

In this work we establish two Bourgain–Brezis–Mironescu [7] type compactness theorems regarding the nonlocal Sobolev spaces associated with half-space gradient operators [20]. An equivalent Fourier characterization of the nonlocal function spaces is shown, which highlights the key role of the Fourier symbol  $\lambda_w^\nu(\xi)$  in the studies of nonlocal function spaces  $S_w^\nu(\Omega)$ . Using an improved lower bound estimate on  $|\lambda_w^\nu(\xi)|$ , we are able to show the locally compact embedding of  $S_w^\nu(\mathbb{R}^d)$  into  $L^2(\mathbb{R}^d)$  for a nonintegrable kernel  $w$ . For the sequence of kernels  $\{w_n\}$  such that  $0 \leq w_n \nearrow w$  where  $w$  is nonintegrable, we prove the sequential compactness result in Theorem 3.6. Another compactness

result holds for a different sequence of kernels satisfying Assumption 3.9, while the result is shown in Theorem 3.11. Based on these compactness theorems we show the uniform Poincaré inequalities generalizing the results in [20], and study the convergence and asymptotic compatibility of nonlocal diffusion problems and their approximations, respectively. Finally we apply the theoretical results to optimal control problems with nonlocal equations as constraints.

A key tool to show  $L^p$ -compactness is the Riesz–Kolmogorov–Fréchet compactness criterion. Here, however, we can only show compactness for  $p = 2$  as in such case one can use Plancherel's theorem to convert the  $L^2$  norm of a function into that of its Fourier transform. It is intriguing to explore whether the compactness result holds for general  $1 < p < \infty$ . For  $p = 2$ , one crucial ingredient we frequently use is the lower bound estimate of  $|\lambda_w^v(\xi)|$ . For large  $|\xi|$  the lower bound has been significantly improved compared to [20, Lemma 5.1]. For small  $|\xi|$ , we apply the technique of truncation at infinity to make  $w$  compactly supported so that Lemma 2.1(2)(a) is applicable to derive  $|\lambda_w^v(\xi)| \geq C|\xi|$  for  $|\xi|$  small.

Note that Assumption 3.9 provides sufficient conditions to derive the crucial lower bound estimates for the Fourier symbols associated with a sequence of kernels in Lemma 3.12. However, it is worth exploring whether these assumptions can be relaxed. An open question remains: can the compactness result in Theorem 3.11 be established under the weakest assumptions, specifically Equations (1), (25) and (26)? To illustrate, consider a kernel defined by

$$w_\delta(z) = \frac{2d}{\omega_{d-1}} \frac{1}{|\log(\delta)|} \frac{1}{|z|^{d+1}} \chi_{\{\delta < |z| < 1\}}(z).$$

This kernel appeared in [29, Corollary 2.3] has the feature that it is truncated near origin. Our numerical results show that both the real part and imaginary part of the Fourier symbol are oscillating but eventually the real part stabilizes at a positive value while the imaginary part converges to zero. However, we could not prove the compactness results because we were unable to determine the critical value  $|\xi|$  such that we use the lower bound of imaginary and real part separately across this critical value. We nevertheless conjecture that Theorem 3.11 will remain valid under the assumptions Equations (1), (25) and (26) and hope to explore this in future work.

## FUNDING INFORMATION

ZH and XT's research is supported in part by NSF grants DMS-2111608 and DMS-2240180. TM's research is supported by NSF grant DMS-2206252. The early stages of this project were conducted during a SQuaRE at the American Institute for Mathematics (AIM) titled *Variational methods for multiscale and nonlinear nonlocal models with applications to peridynamics in San Jose, CA, in May 2023*. The authors thank AIM for providing a supportive and mathematically rich environment.

## DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

## ORCID

Zhaolong Han  <https://orcid.org/0009-0004-1970-4638>

Xiaochuan Tian  <https://orcid.org/0000-0002-4539-6702>

## REFERENCES

- [1] H. Antil, T. Brown, R. Khatri, A. Onwunta, D. Verma, and M. Warma, “Optimal control, numerics, and applications of fractional PDEs,” *Handbook of numerical analysis*, Numerical Control: Part A, Vol 23, E. Trélat and E. Zuazua (eds.), Elsevier, Amsterdam, 2022, pp. 87–114.

[2] H. Antil, J. Pfefferer, and M. Warma, *A note on semilinear fractional elliptic equation: Analysis and discretization*, *ESAIM Math. Modell. Numer. Anal.* 51 (2017), no. 6, 2049–2067.

[3] H. Antil, D. Verma, and M. Warma, *Optimal control of fractional elliptic PDEs with state constraints and characterization of the dual of fractional-order Sobolev spaces*, *J. Optim. Theory Appl.* 186 (2020), no. 1, 1–23.

[4] A. Arroyo-Rabasa, Functional and variational aspects of nonlocal operators associated with linear PDEs. arXiv preprint, arXiv:2210.13161 2022.

[5] 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 (2023), no. 1, 1–48.

[6] J. C. Bellido, C. Mora-Corral, and H. Schönberger, *Nonlocal gradients: Fundamental theorem of calculus*, *Poincaré Inequalities and Embeddings*. arXiv preprint, arXiv:2402.16487 (2024).

[7] J. Bourgain, H. Brezis, and P. Mironescu, *Another look at Sobolev spaces*, IOS, Amsterdam, 2001.

[8] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Springer, New York, NY, 2011.

[9] O. Burkowska, C. Glusa, and M. D’Elia, *An optimization-based approach to parameter learning for fractional type nonlocal models*, *Comput. Math. Appl.* 116 (2022), 229–244.

[10] J. Cueto, C. Kreisbeck, and H. Schönberger, *A variational theory for integral functionals involving finite-horizon fractional gradients*, *Fract. Calculus Appl. Anal.* 26 (2023), no. 5, 2001–2056.

[11] J. Cueto, C. Kreisbeck, and H. Schönberger,  $\Gamma$ -convergence involving nonlocal gradients with varying horizon: Recovery of local and fractional models. arXiv preprint, arXiv:2404.18509 2024.

[12] M. D’Elia, C. Glusa, and E. Otárola, *A priori error estimates for the optimal control of the integral fractional Laplacian*, *SIAM J. Control. Optim.* 57 (2019), no. 4, 2775–2798.

[13] 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. Calculus Appl. Anal.* 25 (2022), no. 6, 2488–2531.

[14] M. D’Elia and M. Gunzburger, *Optimal distributed control of nonlocal steady diffusion problems*, *SIAM J. Control. Optim.* 52 (2014), no. 1, 243–273.

[15] Q. Du, J. M. Scott, and X. Tian, Asymptotically compatible schemes for nonlinear variational models via gamma-convergence and applications to nonlocal problems. arXiv preprint, arXiv:2402.07749 2024.

[16] Q. Du and X. Tian, *Stability of nonlocal Dirichlet integrals and implications for peridynamic correspondence material modeling*, *SIAM J. Appl. Math.* 78 (2018), no. 3, 1536–1552.

[17] A. Evgrafov and J. C. Bellido, *Nonlocal control in the conduction coefficients: Well-posedness and convergence to the local limit*, *SIAM J. Control. Optim.* 58 (2020), no. 4, 1769–1794.

[18] I. Fonseca and G. Leoni, *Modern methods in the calculus of variations:  $L^p$  spaces*, Springer Monographs in Mathematics, New York, NY, 2006.

[19] J. T. Foster, S. A. Silling, and W. W. Chen, *Viscoplasticity using peridynamics*, *Int. J. Numer. Methods Eng.* 81 (2010), no. 10, 1242–1258.

[20] Z. Han and X. Tian, *Nonlocal half-ball vector operators on bounded domains: Poincaré inequality and its applications*, *Math. Models Methods Appl. Sci.* 33 (2023), no. 12, 2507–2556.

[21] H. Lee and Q. Du, *Nonlocal gradient operators with a nonspherical interaction neighborhood and their applications*, *ESAIM Math. Modell. Numer. Anal.* 54 (2020), no. 1, 105–128.

[22] T. Mengesha, *Nonlocal Korn-type characterization of Sobolev vector fields*, *Commun. Contemp. Math.* 14 (2012), no. 04, 1250028.

[23] T. Mengesha and Q. Du, *Nonlocal constrained value problems for a linear Peridynamic Navier equation*, *J. Elast.* 116 (2014), no. 1, 27–51.

[24] T. Mengesha and Q. Du, *Characterization of function spaces of vector fields and an application in nonlinear peridynamics*, *Nonlinear Anal.* 140 (2016), 82–111.

[25] T. Mengesha, A. J. Salgado, and J. M. Siktar, *On the optimal control of a linear Peridynamics model*, *Appl. Math. Optim.* 88 (2023), no. 3, 70.

[26] T. Mengesha and D. Spector, *Localization of nonlocal gradients in various topologies*, *Calc. Var.* 52 (2015), no. 1–2, 253–279.

[27] J. Muñoz, *Local and nonlocal optimal control in the source*, *Mediterr. J. Math.* 19 (2022), no. 1, 1–24.

[28] E. Otárola, R. Rankin, and A. J. Salgado, *Maximum-norm a posteriori error estimates for an optimal control problem*, *Comput. Optim. Appl.* 73 (2019), no. 3, 997–1017.

[29] A. Ponce, *An estimate in the spirit of Poincaré’s inequality*, *J. Eur. Math. Soc.* 6 (2004), 1–15.

[30] T.-T. Shieh and D. E. Spector, *On a new class of fractional partial differential equations*, *Adv. Calculus Variat.* 8 (2015), no. 4, 1–16.

[31] T.-T. Shieh and D. E. Spector, *On a new class of fractional partial differential equations II*, *Adv. Calculus Variat.* 11 (2018), no. 3, 289–307.

[32] S. A. Silling, *Reformulation of elasticity theory for discontinuities and long-range forces*, *J. Mech. Phys. Solids* 48 (2000), no. 1, 175–209.

[33] S. A. Silling, *Stability of peridynamic correspondence material models and their particle discretizations*, *Comput. Methods Appl. Mech. Eng.* 322 (2017), 42–57.

[34] S. A. Silling, M. Epton, O. Weckner, J. Xu, and E. Askari, *Peridynamic states and constitutive modeling*, *J. Elast.* 88 (2007), no. 2, 151–184.

[35] X. Tian and Q. Du, *Asymptotically compatible schemes and applications to robust discretization of nonlocal models*, *SIAM J. Numer. Anal.* 52 (2014), no. 4, 1641–1665.

[36] X. Tian and Q. Du, *Nonconforming discontinuous Galerkin methods for nonlocal Variational problems*, *SIAM J. Numer. Anal.* 53 (2015), no. 2, 762–781.

[37] F. Tröltzsch, “Optimal control of partial differential equations: Theory, methods, and applications,” *Graduate studies in mathematics 1065–7339 (print)*, online-aug ed., American Mathematical Society, Providence, R.I., 2010.

[38] M. R. Tupek and R. Radovitzky, *An extended constitutive correspondence formulation of peridynamics based on nonlinear bond-strain measures*, *J. Mech. Phys. Solids* 65 (2014), 82–92.

[39] T. L. Warren, S. A. Silling, A. Askari, O. Weckner, M. A. Epton, and J. Xu, *A non-ordinary state-based peridynamic method to model solid material deformation and fracture*, *Int. J. Solids Struct.* 46 (2009), no. 5, 1186–1195.

[40] J. Wloka, *Partial differential equations*, Cambridge University Press, Cambridge, 1987.

**How to cite this article:** Z. Han, T. Mengesha, and X. Tian, *Compactness results for a Dirichlet energy of nonlocal gradient with applications*, *Numer. Methods Partial Differ. Eq.* 40 (2024), e23149. <https://doi.org/10.1002/num.23149>

## APPENDIX

### A.1 | PROOFS OF LEMMA 2.2 AND PROPOSITION 3.8

*Proof of Lemma 2.2.* Recall from the proof of [20, Lemma 5.1] that

$$|\Re(\lambda_w^\nu)(\xi)| \geq \int_{z_1 > 0} \frac{z_1}{|z|} w(z) (1 - \cos(2\pi R_\nu^T \xi \cdot z)) dz,$$

where  $R_\nu$  is an orthogonal matrix such that  $\nu = R_\nu e_1$ . Denote  $\mu := R_\nu^T \xi / |R_\nu^T \xi|$ . Then

$$\begin{aligned} |\Re(\lambda_w^\nu)(\xi)| &\geq \int_{z_1 > 0} \frac{z_1}{|z|} w(z) (1 - \cos(2\pi |\xi| \mu \cdot z)) dz \\ &=: I(\xi). \end{aligned}$$

Our goal is to find constant  $c_\lambda = c_\lambda(N\epsilon, d) > 0$  such that

$$I(\xi) \geq c_\lambda \int_{|z| > \frac{N\epsilon}{|\xi|}} w(z) dz, \quad \forall |\xi| > N, \xi \in \mathbb{R}^d. \quad (\text{A1})$$

We distinguish two cases  $d = 1$  and  $d \geq 2$ .

**Case I:**  $d = 1$ . Then  $\mu = e_1$  or  $\mu = -e_1$ . In either case one has

$$I(\xi) = \int_0^\infty w(z) (1 - \cos(2\pi |\xi| z)) dz.$$

Since the right-hand side is even in  $\xi$ , it suffices to show that for  $\xi > N$ ,

$$I(\xi) = \int_0^\infty w(z) (1 - \cos(2\pi \xi z)) dz \geq c_\lambda \int_{\frac{N\epsilon}{|\xi|}}^\infty w(z) dz. \quad (\text{A2})$$

Fix  $R > 1/(4N) + \epsilon$ . Without loss of generality, assume  $N < 1/(4\epsilon)$ . On the one hand, using the nonincreasing assumption on  $w$  and [16, Lemma 2] we have

$$\begin{aligned} \int_{\frac{1}{4\xi}}^R w(z) \cos(2\pi\xi z) dz &= \frac{1}{2\pi\xi} \int_{\frac{\pi}{2}}^{2\pi\xi R} w\left(\frac{z}{2\pi\xi}\right) \cos(z) dz \\ &= -\frac{1}{2\pi\xi} \int_0^{2\pi\xi R - \frac{\pi}{2}} w\left(\frac{1}{2\pi\xi}\left(z + \frac{\pi}{2}\right)\right) \sin(z) dz \leq 0, \end{aligned}$$

thus

$$\begin{aligned} \int_{\frac{1}{4\xi}}^R w(z) dz &= \int_{\frac{1}{4\xi}}^R w(z)(1 - \cos(2\pi\xi z)) dz + \int_{\frac{1}{4\xi}}^R w(z) \cos(2\pi\xi z) dz \\ &\leq \int_{\frac{1}{4\xi}}^R w(z)(1 - \cos(2\pi\xi z)) dz. \end{aligned} \tag{A3}$$

On the other hand, again by nonincreasing property of  $w$  one has

$$\begin{aligned} \int_{\frac{N\epsilon}{\xi}}^{\frac{1}{4\xi}} w(z) dz &= \xi^{-1} \int_{N\epsilon}^{\frac{1}{4}} w(\xi^{-1}z) dz \leq \xi^{-1} \frac{1 - 4N\epsilon}{2N\epsilon} \int_{\frac{N\epsilon}{2}}^{N\epsilon} w(\xi^{-1}z) dz \\ &\leq \xi^{-1} \frac{1 - 4N\epsilon}{2N\epsilon} \frac{1}{1 - \cos(N\epsilon)} \int_{\frac{N\epsilon}{2}}^{N\epsilon} w(\xi^{-1}z)(1 - \cos(2\pi z)) dz \\ &= \frac{1 - 4N\epsilon}{2N\epsilon} \frac{1}{1 - \cos(N\epsilon)} \int_{\frac{N\epsilon}{2\xi}}^{\frac{N\epsilon}{\xi}} w(z)(1 - \cos(2\pi\xi z)) dz \\ &\leq \frac{1 - 4N\epsilon}{2N\epsilon} \frac{1}{1 - \cos(N\epsilon)} \int_0^{\frac{1}{4\xi}} w(z)(1 - \cos(2\pi\xi z)) dz \end{aligned} \tag{A4}$$

Combining (A3) and (A4) yields

$$I(\xi) \geq \int_0^R w(z)(1 - \cos(2\pi\xi z)) dz \geq c_\lambda \int_{\frac{N\epsilon}{\xi}}^R w(z) dz, \quad \xi > N,$$

where

$$c_\lambda = \left(1 + \frac{1 - 4N\epsilon}{2N\epsilon(1 - \cos(N\epsilon))}\right)^{-1}.$$

Letting  $R \rightarrow +\infty$  yields (A2).

**Case II:**  $d \geq 2$ . Let  $\mu = (\mu_1, \dots, \mu_d)$ . Notice that  $I(\xi)$  maintains if  $\mu$  is replaced by  $-\mu$ . Without loss of generality, we assume  $\mu_1 \geq 0$ . Denote  $\mu = (\mu_1, \mu')$  where  $\mu' \in \mathbb{R}^{d-1}$ .

**Case II(i):**  $\mu_1 = 0$ . This corresponds to the case where  $\xi \perp \nu$ . Fix  $R > \epsilon/2$ . By doing a change of variable and using the Fubini–Tonelli theorem, one obtains

$$\begin{aligned}
 I(\xi) &\geq \int_{|z'| < R} (1 - \cos(2\pi|\xi|\mu' \cdot z')) \int_0^{\sqrt{R^2 - |z'|^2}} \frac{z_1}{|z|} \bar{w}(|z|) dz_1 dz' \\
 &= \int_{|z'| < R} (1 - \cos(2\pi|\xi|\mu' \cdot z')) \int_{|z'|}^R \bar{w}(r) dr dz' \\
 &= \int_0^R \bar{w}(r) \int_{|z'| < r} (1 - \cos(2\pi|\xi|\mu' \cdot z')) dz' dr \\
 &\geq \int_{\frac{N\epsilon}{|\xi|}}^R \bar{w}(r) \int_{|z'| < r} (1 - \cos(2\pi|\xi|\mu' \cdot z')) dz' dr \\
 &= \int_{\frac{N\epsilon}{|\xi|}}^R \bar{w}(r) \int_{B_r^{d-1}(\mathbf{0})} (1 - \cos(2\pi|\xi|y_1)) dy dr \\
 &= \int_{\frac{N\epsilon}{|\xi|}}^R r^{d-1} \bar{w}(r) \int_{B_1^{d-1}(\mathbf{0})} (1 - \cos(2\pi r|\xi|y_1)) dy dr,
 \end{aligned} \tag{A5}$$

where in the last two steps we used change-of-variable formula.

Denote

$$J(r) := \int_{B_1^{d-1}(\mathbf{0})} (1 - \cos(2\pi r y_1)) dy.$$

It suffices to show  $J(r)$  has a positive lower bound  $c_J = c_J(N\epsilon, d)$  over  $[N\epsilon, \infty)$  to prove (A1). Indeed, using (A5) and this positive lower bound one obtains

$$I(\xi) \geq \int_{\frac{N\epsilon}{|\xi|}}^R r^{d-1} \bar{w}(r) J(r|\xi|) dr \geq c_J \int_{\frac{N\epsilon}{|\xi|}}^R r^{d-1} \bar{w}(r) dr = \frac{c_J}{\omega_{d-1}} \int_{\frac{N\epsilon}{|\xi|} < |z| < R} w(z) dz$$

for any  $|\xi| > N$  with  $\xi \in \mathbb{R}^d$ . Letting  $R \rightarrow +\infty$  gives (A1). Now we focus on the proof of  $J(r) \geq c_J$  for  $r \geq N\epsilon$ .

Note that for any  $r > 0$ ,  $J(r) > 0$  as the integrand is nonnegative. By a change of variable, one obtains

$$\begin{aligned}
 J(r) &= \int_0^1 (1 - \cos(2\pi r y_1)) \int_{|y'| < \sqrt{1-y_1^2}} dy' dy_1 \\
 &= V_{d-2} \int_0^1 (1 - \cos(2\pi r y_1)) (1 - y_1^2)^{\frac{d-2}{2}} dy_1 \\
 &\geq V_{d-2} \left( \frac{\sqrt{3}}{2} \right)^{d-2} \int_0^{\frac{1}{2}} (1 - \cos(2\pi r y)) dy \\
 &\rightarrow V_{d-2} \left( \frac{\sqrt{3}}{2} \right)^{d-2} \frac{1}{2} > 0, \quad r \rightarrow +\infty,
 \end{aligned}$$

where in the last step we used Riemann–Lebesgue lemma. Since  $J(r)$  is continuous, there exists  $c_J = c_J(N\epsilon, d) > 0$  such that  $J(r) \geq c_J$  for  $r \geq N\epsilon$ . This concludes the proof of Case II(i).

**Case II(ii):**  $\mu_1 > 0$ . By Proposition A.1 in Appendix A.2, there exist  $\mathbf{v}_1, \dots, \mathbf{v}_{d-1} \in \mathbb{S}^{d-1}$  such that

$$\mathbf{v}_k^{(1)} := \mathbf{v}_k \cdot \mathbf{e}_1 \geq 0, \quad 1 \leq k \leq d-1,$$

and  $\mathbf{v}_1, \dots, \mathbf{v}_{d-1}, \boldsymbol{\mu}$  form an orthonormal basis for  $\mathbb{R}^d$ . Moreover,

$$\mathbf{v}_k^{(1)} = \frac{\mu_1^{k-1} |\mu_{k+1}|}{\sqrt{\mu_1^2 + \dots + \mu_k^2} \sqrt{\mu_1^2 + \dots + \mu_{k+1}^2}}, \quad 1 \leq k \leq d-1.$$

Denote  $A_{\boldsymbol{\mu}} := (\mathbf{v}_1, \dots, \mathbf{v}_{d-1}, \boldsymbol{\mu})$ . Then  $A_{\boldsymbol{\mu}}$  is a  $d \times d$  orthogonal matrix. Consider the following change of variable  $\mathbf{z} = A_{\boldsymbol{\mu}} \mathbf{y}$  where  $\mathbf{y} \in \mathbb{R}^d$ . Then

$$z_1 = \mu_1 y_d + \sum_{k=1}^{d-1} \mathbf{v}_k^{(1)} y_k > 0$$

if  $y_1, \dots, y_d > 0$ . Thus, for  $R > 1$ ,

$$\begin{aligned} I(\xi) &\geq \int_{z_1 > 0, |\mathbf{z}| < R} \frac{z_1}{|\mathbf{z}|} w(\mathbf{z}) (1 - \cos(2\pi |\xi| \boldsymbol{\mu} \cdot \mathbf{z})) d\mathbf{z} \\ &\geq \int_{y_1 > 0, \dots, y_d > 0, |\mathbf{y}| < R} \left( \mu_1 y_d + \sum_{k=1}^{d-1} \mathbf{v}_k^{(1)} y_k \right) \frac{1 - \cos(2\pi |\xi| y_d)}{|\mathbf{y}|} w(\mathbf{y}) d\mathbf{y}. \end{aligned} \quad (\text{A7})$$

**Case II(ii)(a):** If  $\mu_1$  is the largest among  $\{\mathbf{v}_1^{(1)}, \dots, \mathbf{v}_{d-1}^{(1)}, \mu_1\}$ , then  $\mu_1 \geq 1/\sqrt{d}$  and (A7) yields

$$\begin{aligned} I(\xi) &\geq \frac{1}{\sqrt{d}} \int_{y_1 > 0, \dots, y_d > 0, |\mathbf{y}| < R} y_d \frac{1 - \cos(2\pi |\xi| y_d)}{|\mathbf{y}|} w(\mathbf{y}) d\mathbf{y} \\ &= \frac{1}{\sqrt{d}} \int_0^R y_d (1 - \cos(2\pi |\xi| y_d)) \int_{y_1 > 0, \dots, y_{d-1} > 0, y_1^2 + \dots + y_{d-1}^2 < R^2 - y_d^2} \frac{w(\mathbf{y})}{|\mathbf{y}|} d\tilde{\mathbf{y}} dy_d, \end{aligned}$$

where  $\tilde{\mathbf{y}} = (y_1, \dots, y_{d-1})$ . Using polar coordinates one may compute

$$\begin{aligned} \int_{y_1 > 0, \dots, y_{d-1} > 0, y_1^2 + \dots + y_{d-1}^2 < R^2 - y_d^2} \frac{w(\mathbf{y})}{|\mathbf{y}|} d\tilde{\mathbf{y}} &= \frac{1}{2^{d-1}} \int_{|\tilde{\mathbf{y}}| < \sqrt{R^2 - y_d^2}} \frac{1}{(y_d^2 + |\tilde{\mathbf{y}}|^2)^{\frac{1}{2}}} \bar{w}\left((y_d^2 + |\tilde{\mathbf{y}}|^2)^{\frac{1}{2}}\right) d\tilde{\mathbf{y}} \\ &= \frac{\omega_{d-2}}{2^{d-1}} \int_0^{\sqrt{R^2 - y_d^2}} \frac{r^{d-2}}{(y_d^2 + r^2)^{\frac{1}{2}}} \bar{w}\left((y_d^2 + r^2)^{\frac{1}{2}}\right) dr \\ &= \frac{\omega_{d-2}}{2^{d-1}} \int_{y_d}^R \bar{w}(r) (r^2 - y_d^2)^{\frac{d-3}{2}} dr. \end{aligned}$$

Therefore,

$$\begin{aligned} I(\xi) &\geq \frac{1}{\sqrt{d}} \frac{\omega_{d-2}}{2^{d-1}} \int_0^R y_d (1 - \cos(2\pi |\xi| y_d)) \int_{y_d}^R \bar{w}(r) (r^2 - y_d^2)^{\frac{d-3}{2}} dr dy_d \\ &= \frac{1}{\sqrt{d}} \frac{\omega_{d-2}}{2^{d-1}} \int_0^R \bar{w}(r) \int_0^r y_d (1 - \cos(2\pi |\xi| y_d)) (r^2 - y_d^2)^{\frac{d-3}{2}} dy_d dr \\ &= \frac{1}{\sqrt{d}} \frac{\omega_{d-2}}{2^{d-1}} \int_0^R r^{d-1} \bar{w}(r) \int_0^1 t (1 - \cos(2\pi r |\xi| t)) (1 - t^2)^{\frac{d-3}{2}} dt dr \\ &\geq \frac{1}{\sqrt{d}} \frac{\omega_{d-2}}{2^{d-1}} \int_{\frac{N\epsilon}{|\xi|}}^R r^{d-1} \bar{w}(r) H(r|\xi|) dr, \end{aligned} \quad (\text{A8})$$

where  $H : (0, \infty) \rightarrow (0, \infty)$  is defined by

$$H(r) := \int_0^1 t(1 - \cos(2\pi rt))(1 - t^2)^{\frac{d-3}{2}} dt. \quad (\text{A9})$$

It suffices to show there exists  $c_H = c_H(N\epsilon, d) > 0$  such that  $H(r) \geq c_H$  for  $r \geq N\epsilon$  to show (A1). Indeed, from (A8) and the lower bound  $c_H$  one obtains

$$I(\xi) \geq c_H \frac{1}{\sqrt{d}} \frac{\omega_{d-2}}{2^{d-1}} \int_{\frac{N\epsilon}{|\xi|}}^R r^{d-1} \bar{w}(r) dr = \frac{c_H}{\omega_{d-1}} \frac{1}{\sqrt{d}} \frac{\omega_{d-2}}{2^{d-1}} \int_{\frac{N\epsilon}{|\xi|} < |z| < R} w(z) dz$$

for any  $|\xi| > N$  with  $\xi \in \mathbb{R}^d$ . Letting  $R \rightarrow +\infty$  yields (A1). Noticing that  $H(r)$  is continuous and positive for any  $r > 0$ , it suffices to show

$$\lim_{r \rightarrow +\infty} H(r) > 0$$

to conclude the existence of the positive lower bound  $c_H$ . The proof uses again Riemann-Lebesgue lemma as follows.

When  $d = 2$ , since the function  $t \mapsto t(1 - t^2)^{-\frac{1}{2}}$  is increasing over  $(0, 1)$ , one has

$$H(r) \geq \frac{1}{\sqrt{3}} \int_{\frac{1}{2}}^1 (1 - \cos(2\pi rt)) dt \rightarrow \frac{1}{2\sqrt{3}}, \quad r \rightarrow +\infty.$$

When  $d \geq 3$ , since the function  $t \mapsto (1 - t^2)^{\frac{d-3}{2}}$  is nonincreasing over  $(0, 1)$ , one has

$$H(r) \geq \left( \frac{\sqrt{3}}{2} \right)^{d-3} \int_0^{\frac{1}{2}} t(1 - \cos(2\pi rt)) dt \rightarrow \left( \frac{\sqrt{3}}{2} \right)^{d-3} \frac{1}{8}, \quad r \rightarrow +\infty.$$

Then the proof for Case II(ii)(a) is done.

**Case II(ii)(b):** If  $v_j^{(1)}$  is the largest among  $\{v_1^{(1)}, \dots, v_{d-1}^{(1)}, \mu_1\}$  where  $1 \leq j \leq d-1$ , then  $v_j^{(1)} \geq 1/\sqrt{d}$ . Denoting  $\tilde{\mathbf{y}} := (y_1, \dots, \hat{y}_j, \dots, y_d) = (\tilde{y}_1, \dots, \tilde{y}_{d-1}) \in \mathbb{R}^{d-1}$  (with  $\tilde{y}_{d-1} = y_d$ ) we continue from (A7) to compute

$$\begin{aligned} I(\xi) &\geq \frac{1}{\sqrt{d}} \int_{y_1 > 0, \dots, y_d > 0, |\mathbf{y}| < R} y_j \frac{1 - \cos(2\pi|\xi|y_d)}{|\mathbf{y}|} w(\mathbf{y}) d\mathbf{y} \\ &= \frac{1}{\sqrt{d}} \int_{\tilde{y}_1 > 0, \dots, \tilde{y}_{d-1} > 0, |\tilde{\mathbf{y}}| < R} (1 - \cos(2\pi|\xi|\tilde{y}_{d-1})) \int_0^{\sqrt{R^2 - |\mathbf{y}|^2}} \frac{y_j}{|\mathbf{y}|} \bar{w}(|\mathbf{y}|) dy_j d\tilde{\mathbf{y}} \\ &= \frac{1}{\sqrt{d}} \int_{\tilde{y}_1 > 0, \dots, \tilde{y}_{d-1} > 0, |\tilde{\mathbf{y}}| < R} (1 - \cos(2\pi|\xi|\tilde{y}_{d-1})) \int_{|\tilde{\mathbf{y}}|}^R \bar{w}(r) dr d\tilde{\mathbf{y}} \\ &= \frac{1}{\sqrt{d}} \frac{1}{2^{d-1}} \int_{B_1^{d-1}(\mathbf{0})} (1 - \cos(2\pi|\xi|z_1)) \int_{|z|}^R \bar{w}(r) dr dz \\ &= \frac{1}{\sqrt{d}} \frac{1}{2^{d-1}} \int_0^R \bar{w}(r) \int_{B_r^{d-1}(\mathbf{0})} (1 - \cos(2\pi|\xi|z_1)) dz dr \\ &\geq \frac{1}{\sqrt{d}} \frac{1}{2^{d-1}} \int_{\frac{N\epsilon}{|\xi|}}^R r^{d-1} \bar{w}(r) \int_{B_1^{d-1}(\mathbf{0})} (1 - \cos(2\pi r|\xi|z_1)) dz dr. \end{aligned}$$

Note that the above calculation is similar to (A5) and the final expression is just a constant multiple of that of (A5). Therefore, following the proof in Case (II)(i) after (A5) we complete the whole proof. ■

*Proof of Proposition 3.8.* We first prove the pointwise convergence result and then use generalized dominated convergence theorem to prove the  $L^p$  convergence. For simplicity we focus on  $\mathcal{G}_\delta^\nu$ .

Without loss of generality, we may assume that  $u \in C_c^2(\mathbb{R}^d)$  in the proof. The result for  $u \in C_c^1(\mathbb{R}^d)$  can then be obtained using a standard density argument so we omit the details. Since  $u \in C_c^2(\mathbb{R}^d)$ , there exists  $R > 0$  such that  $\text{supp } u \subset B_R(\mathbf{0}) =: B_R$ . Recall that

$$\mathcal{G}_\delta^\nu u(\mathbf{x}) = \int_{\mathbb{R}^d} \chi_\nu(z) \frac{z}{|z|} w_\delta(z) (u(\mathbf{x} + z) - u(\mathbf{x})) dz.$$

We discuss two cases  $|\mathbf{x}| \leq 2R$  and  $|\mathbf{x}| > 2R$ .

**Case (I):**  $|\mathbf{x}| \leq 2R$ . Then for  $|z| > 3R$  one has  $|\mathbf{x} + z| \geq |z| - |\mathbf{x}| > R$  and thus  $u(\mathbf{x} + z) = 0$ . Note that we have Taylor expansion

$$u(\mathbf{x} + z) - u(\mathbf{x}) = Du(\mathbf{x})z + \int_0^1 z^T D^2 u(\mathbf{x} + tz) z dt.$$

Hence,

$$\begin{aligned} \mathcal{G}_\delta^\nu u(\mathbf{x}) &= \int_{B_{3R}} \chi_\nu(z) \frac{z}{|z|} w_\delta(z) (u(\mathbf{x} + z) - u(\mathbf{x})) dz + \int_{B_{3R}^c} \chi_\nu(z) \frac{z}{|z|} w_\delta(z) (-u(\mathbf{x})) dz \\ &= \left( \int_{B_{3R}} \chi_\nu(z) |z| w_\delta(z) \frac{z}{|z|} \otimes \frac{z}{|z|} dz \right) \nabla u(\mathbf{x}) + \int_{B_{3R}^c} \chi_\nu(z) \frac{z}{|z|} w_\delta(z) (-u(\mathbf{x})) dz \\ &\quad + \int_{B_{3R}^c} \chi_\nu(z) \frac{z}{|z|} w_\delta(z) \left( \int_0^1 z^T D^2 u(\mathbf{x} + tz) z dt \right) dz \\ &=: I_1 + I_2 + I_3. \end{aligned}$$

Since  $w_\delta$  is radial, one can compute as in [20, Remark 2.2] that

$$\int_{B_{3R}} \chi_\nu(z) |z| w_\delta(z) \frac{z}{|z|} \otimes \frac{z}{|z|} dz = \frac{1}{2d} \left( \int_{B_{3R}} |z| w_\delta(z) dz \right) \mathbb{I}_d \rightarrow \mathbb{I}_d, \quad \delta \rightarrow 0^+,$$

where  $\mathbb{I}_d$  is the  $d \times d$  identity matrix and we used (26). Hence  $I_1 \rightarrow \nabla u(\mathbf{x})$  as  $\delta \rightarrow 0$ .

For  $I_2$  and  $I_3$ , we show that they converge to 0 as  $\delta \rightarrow 0$ . Indeed, as  $\delta \rightarrow 0$ , by Equations (25) and (27),

$$|I_2| \leq \|u\|_{C^0(\mathbb{R}^d)} \int_{B_{3R}^c} w_\delta(z) dz \rightarrow 0$$

and

$$|I_3| \leq \|u\|_{C^2(\mathbb{R}^d)} \int_{B_{3R}^c} |z|^2 w_\delta(z) dz \rightarrow 0.$$

Combining all the estimates for  $I_1$ ,  $I_2$  and  $I_3$  yields

$$\mathcal{G}_\delta^\nu u(\mathbf{x}) \rightarrow \nabla u(\mathbf{x}), \quad \delta \rightarrow 0, \quad \forall |\mathbf{x}| < 2R.$$

Moreover, there exist constants  $\delta_0 \in (0, 1)$  and  $C_0 = C_0(\|u\|_{C^2(\mathbb{R}^d)})$  such that  $|\mathcal{G}_\delta^\nu u(\mathbf{x})| \leq C_0$  for  $|\mathbf{x}| \leq 2R$  and  $\delta \in (0, \delta_0)$ .

**Case (II):**  $|\mathbf{x}| > 2R$ . Then  $u(\mathbf{x}) = 0$  and

$$\mathcal{G}_\delta^\nu u(\mathbf{x}) = \int_{\mathbb{R}^d} \chi_\nu(z) \frac{z}{|z|} w_\delta(z) u(\mathbf{x} + z) dz = \int_{B_R(-\mathbf{x})} \chi_\nu(z) \frac{z}{|z|} w_\delta(z) u(\mathbf{x} + z) dz.$$

Then for  $|\mathbf{x}| > 2R$ ,

$$|\mathcal{G}_\delta^\nu u(\mathbf{x})| \leq \|u\|_{C^0(\mathbb{R}^d)} \int_{B_R(-\mathbf{x})} w_\delta(z) dz.$$

Define  $g_\delta : \mathbb{R}^d \rightarrow \mathbb{R}$  by

$$g_\delta(\mathbf{x}) = \begin{cases} C_0, & |\mathbf{x}| \leq 2R, \\ \|u\|_{C^0(\mathbb{R}^d)} \int_{B_R(-\mathbf{x})} w_\delta(z) dz, & |\mathbf{x}| > 2R. \end{cases}$$

Then  $|\mathcal{G}_\delta^\nu u(\mathbf{x})| \leq g_\delta(\mathbf{x})$  for any  $\mathbf{x} \in \mathbb{R}^d$ . Note that for  $|\mathbf{x}| > 2R$ ,  $B_R(-\mathbf{x}) \cap B_R(\mathbf{0}) = \emptyset$ , hence

$$|\mathcal{G}_\delta^\nu u(\mathbf{x})| \leq g_\delta(\mathbf{x}) \leq \|u\|_{C^0(\mathbb{R}^d)} \int_{B_R^c} w_\delta(z) dz \rightarrow 0 = \nabla u(\mathbf{x}), \quad \delta \rightarrow 0.$$

Combining the pointwise limit in Case (I) we obtain

$$\mathcal{G}_\delta^\nu u(\mathbf{x}) \rightarrow \nabla u(\mathbf{x}), \quad \delta \rightarrow 0, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

It remains to show for  $p \in [1, \infty]$ ,  $\mathcal{G}_\delta^\nu u \rightarrow \nabla u$  in  $L^p(\mathbb{R}^d)$ . For  $p = \infty$  this follows from the estimates above. Now suppose  $1 \leq p < \infty$ . Notice that  $g_\delta(\mathbf{x}) \rightarrow C_0 \chi_{B_{2R}}(\mathbf{x}) =: g(\mathbf{x})$  as  $\delta \rightarrow 0$  for any  $\mathbf{x} \in \mathbb{R}^d$ , to apply the generalized dominated convergence theorem, we need to show that  $\|g_\delta\|_{L^p(\mathbb{R}^d)} \rightarrow \|g\|_{L^p(\mathbb{R}^d)}$  as  $\delta \rightarrow 0$ . In fact, we will show  $\|g_\delta - g\|_{L^p(\mathbb{R}^d)} \rightarrow 0$  as  $\delta \rightarrow 0$ . Note that

$$\|g_\delta - g\|_{L^p(\mathbb{R}^d)}^p = \|u\|_{C^0(\mathbb{R}^d)}^p \int_{B_{2R}^c} \left( \int_{B_R(-\mathbf{x})} w_\delta(z) dz \right)^p d\mathbf{x}.$$

Since  $B_R(-\mathbf{x}) \subset B_R^c$  for  $|\mathbf{x}| > 2R$ , one has

$$\begin{aligned} \int_{B_{2R}^c} \left( \int_{B_R(-\mathbf{x})} w_\delta(z) dz \right)^p d\mathbf{x} &= \int_{B_{2R}^c} \left( \int_{\mathbb{R}^d} \chi_{B_R(-\mathbf{x})}(z) w_\delta(z) dz \right)^p d\mathbf{x} \\ &= \int_{B_{2R}^c} \left( \int_{B_R(-\mathbf{x})} w_\delta(z) dz \right)^{p-1} \int_{B_R(-\mathbf{x})} w_\delta(z) dz d\mathbf{x} \\ &\leq \left( \int_{B_R^c} w_\delta(z) dz \right)^{p-1} \int_{B_{2R}^c} \int_{B_R^c} \chi_{B_R(-\mathbf{x})}(z) w_\delta(z) dz d\mathbf{x} \\ &= \left( \int_{B_R^c} w_\delta(z) dz \right)^{p-1} \int_{B_R^c} \int_{B_{2R}^c \cap B_R(-\mathbf{x})} d\mathbf{x} w_\delta(z) dz \\ &\leq |B_R| \left( \int_{B_R^c} w_\delta(z) dz \right)^p \rightarrow 0, \quad \delta \rightarrow 0, \end{aligned}$$

where we used Tonelli–Fubini theorem and (25). This implies  $\|g_\delta - g\|_{L^p(\mathbb{R}^d)} \rightarrow 0$  as  $\delta \rightarrow 0$  and by generalized dominated convergence theorem  $\mathcal{G}_\delta^\nu u \rightarrow \nabla u$  in  $L^p(\mathbb{R}^d)$  for  $p \in [1, \infty]$ .  $\blacksquare$

## A.2 | AN ORTHONORMAL BASIS IN $\mathbb{R}^d$

**Proposition A.1.** *Let  $d \geq 2$  and  $\mu = (\mu_1, \dots, \mu_d) \in \mathbb{S}^{d-1}$  with  $\mu_1 > 0$ . Then there exist  $\mathbf{v}_1, \dots, \mathbf{v}_{d-1} \in \mathbb{S}^{d-1}$  such that*

$$\mathbf{v}_k^{(1)} := \mathbf{v}_k \cdot \mathbf{e}_1 \geq 0, \quad 1 \leq k \leq d-1,$$

and  $v_1, \dots, v_{d-1}, \mu$  form an orthonormal basis for  $\mathbb{R}^d$ . Moreover, one may choose

$$v_k^{(1)} = \frac{\mu_1 |\mu_{k+1}|}{\sqrt{\mu_1^2 + \dots + \mu_k^2} \sqrt{\mu_1^2 + \dots + \mu_{k+1}^2}}, \quad 1 \leq k \leq d-1. \quad (\text{A10})$$

*Proof.* We prove it by induction on  $d$ .

If  $d = 2$ , then one may choose

$$v_1 = \begin{cases} (\mu_2, -\mu_1), & \text{if } \mu_2 \geq 0, \\ (-\mu_2, \mu_1), & \text{if } \mu_2 < 0. \end{cases}$$

Then  $v_1, \mu$  form an orthonormal basis for  $\mathbb{R}^2$  and  $v_1^{(1)} = |\mu_2|$  as desired.

Suppose the conclusion holds for some  $d \geq 2$ . Given  $\mu = (\mu_1, \dots, \mu_{d+1}) \in \mathbb{S}^d$  with  $\mu_1 > 0$ , one has  $|\mu_{d+1}| < 1$  and

$$\tilde{\mu} := \frac{1}{\sqrt{1 - \mu_{d+1}^2}} (\mu_1, \dots, \mu_d) \in \mathbb{S}^{d-1}.$$

By the inductive hypothesis, there exist  $w_1, \dots, w_{d-1} \in \mathbb{S}^{d-1}$  such that  $w_k^{(1)} \geq 0$  for  $1 \leq k \leq d-1$  and  $w_1, \dots, w_{d-1}, \tilde{\mu}$  form an orthonormal basis for  $\mathbb{R}^d$ . Define  $v_k := (w_k, 0) \in \mathbb{S}^d$  for  $1 \leq k \leq d-1$  and

$$v_d := \left( |\mu_{d+1}| \tilde{\mu}, -\text{sgn}(\mu_{d+1}) \sqrt{1 - \mu_{d+1}^2} \right) \in \mathbb{S}^d.$$

Then it is clear that  $v_1, \dots, v_d, \mu$  form an orthonormal basis for  $\mathbb{R}^{d+1}$ . Moreover, if  $w_1^{(1)}, \dots, w_{d-1}^{(1)}$  are assumed to have form in (A10), then it can be checked that  $v_1^{(1)}, \dots, v_d^{(1)}$  satisfy (A10) as well. By induction, the proposition holds for any  $d \geq 2$ . ■

### A.3 | A USEFUL LIMIT CONCERNING THE FOURIER SYMBOL WITH RESPECT TO RIESZ FRACTIONAL KERNEL

**Proposition A.2.** *The following equality holds:*

$$\lim_{\delta \rightarrow 0} \int_0^\infty \frac{\delta}{z^{2-\delta}} (e^{2\pi iz} - 1) dz = 2\pi i, \quad (\text{A11})$$

that is,

$$\lim_{\delta \rightarrow 0} \int_0^\infty \frac{\delta}{z^{2-\delta}} (\cos(2\pi z) - 1) dz = 0, \quad \lim_{\delta \rightarrow 0} \int_0^\infty \frac{\delta}{z^{2-\delta}} \sin(2\pi z) dz = 2\pi. \quad (\text{A12})$$

In addition, these three equalities hold with the upper limit  $\infty$  replaced by 1.

*Proof.* Note that the last assertion follows from the estimate

$$\left| \int_1^\infty \frac{\delta}{z^{2-\delta}} (e^{2\pi iz} - 1) dz \right| \leq 2 \int_1^\infty \frac{\delta}{z^{2-\delta}} dz \leq 2\delta \int_1^\infty \frac{1}{z^{3/2}} dz$$

for  $\delta \in (0, 1/2)$ , provided the following two equalities

$$\lim_{\delta \rightarrow 0} \int_0^1 \frac{\delta}{z^{2-\delta}} (\cos(2\pi z) - 1) dz = 0 \quad (\text{A13})$$

and

$$\lim_{\delta \rightarrow 0} \int_0^1 \frac{\delta}{z^{2-\delta}} \sin(2\pi z) dz = 2\pi \quad (\text{A14})$$

hold. Using the inequality  $|1 - \cos(x)| \leq x^2/2$  for any  $x \in \mathbb{R}$ , it is clear that (A13) holds as

$$\left| \int_0^1 \frac{\delta}{z^{2-\delta}} (\cos(2\pi z) - 1) dz \right| \leq 2\pi^2 \delta \int_0^1 z^\delta dz \leq 2\pi^2 \delta \rightarrow 0, \quad \delta \rightarrow 0.$$

To show (A14), first notice that  $\int_0^1 \frac{\delta}{z^{1-\delta}} dz = 1$ . Since  $\lim_{x \rightarrow 0^+} \sin(x)/x = 1$ , for any  $\epsilon > 0$ , there exists  $N_0 = N_0(\epsilon) \in (0, 1)$  such that for any  $z \in (0, N_0)$ ,

$$\left| \frac{\sin(2\pi z)}{2\pi z} - 1 \right| < \epsilon.$$

Then

$$\left| \int_0^{N_0} \frac{\delta}{z^{1-\delta}} \left( \frac{\sin(2\pi z)}{2\pi z} - 1 \right) dz \right| < \epsilon \int_0^1 \frac{\delta}{z^{1-\delta}} dz = \epsilon.$$

On the other hand, for  $\delta \in (0, 1)$  one has

$$\left| \int_{N_0}^1 \frac{\delta}{z^{1-\delta}} \left( \frac{\sin(2\pi z)}{2\pi z} - 1 \right) dz \right| \leq 2 \int_{N_0}^1 \frac{\delta}{z} dz = -2\delta \ln(N_0) \rightarrow 0, \quad \delta \rightarrow 0.$$

Thus, there exists  $\delta_0 = \delta_0(\epsilon)$  such that for  $\delta \in (0, \delta_0)$

$$\left| \int_0^1 \frac{\delta}{z^{2-\delta}} \sin(2\pi z) dz - 2\pi \right| = 2\pi \left| \int_0^1 \frac{\delta}{z^{1-\delta}} \left( \frac{\sin(2\pi z)}{2\pi z} - 1 \right) dz \right| < 2\epsilon.$$

This proves (A14) and the whole proposition. ■

#### A.4 | AN APPROXIMATION RESULT

**Lemma A.3.** *Let  $\Omega$  be a bounded polygonal domain in  $\mathbb{R}^d$  and  $\{\mathcal{T}_h\}_{h>0}$  be a quasi-uniform mesh of size  $h$  on  $\Omega$ . Given  $\alpha, \beta \in C(\bar{\Omega})$  such that  $\alpha(x) < \beta(x)$  for all  $x \in \bar{\Omega}$ , define  $Z_{\text{ad}} := \{f \in L^2(\Omega) : \alpha(x) \leq f(x) \leq \beta(x), \text{ a.e. } x \in \Omega\}$  and denote  $Z_h$  the piecewise constant functions with respect to the mesh  $\{\mathcal{T}_h\}_{h>0}$ , that is,*

$$Z_h := \{z_h \in L^\infty(\Omega) | z_h|_T \in \mathcal{P}_0(T), \forall T \in \mathcal{T}_h\}.$$

*Then for any  $q \in Z_{\text{ad}}$ , there exists  $\{q_h\}_{h>0} \subset Z_h \cap Z_{\text{ad}}$  such that  $q_h \rightarrow q$  in  $L^2(\Omega)$ .*

*Proof.* Without loss of generality, we assume that  $\alpha(x) < \beta(x)$  for all  $x \in \bar{\Omega}$ . Given  $q \in L^2(\Omega)$ , define the piecewise constant function  $p_h \in L^2(\Omega)$  by

$$p_h(x) := \frac{1}{|T|} \int_T q(y) dy, \quad \text{if } x \in \text{int}(T), \quad T \in \mathcal{T}_h.$$

where  $|T|$  denotes the Lebesgue measure of  $T$ . It is well-known that  $p_h \rightarrow q$  in  $L^2(\Omega)$ . However,  $p_h$  is not necessarily in  $Z_h$ . Consider another piecewise constant  $q_h \in L^2(\Omega)$  given by

$$q_h(x) := \min \left\{ \max \left\{ p_h(x), \sup_{y \in \bar{T}} \alpha(y) \right\}, \inf_{y \in \bar{T}} \beta(y) \right\}, \quad \text{if } x \in \text{int}(T), \quad T \in \mathcal{T}_h.$$

It is clear that  $q_h \in Z_h \cap Z_{\text{ad}}$ . To show  $q_h \rightarrow q$  in  $L^2(\Omega)$ , it suffices to show that  $\|q_h - p_h\|_{L^2(\Omega)} \rightarrow 0$  as  $h \rightarrow 0$ . To that end, consider the partition  $\mathcal{T}_h = \mathcal{T}_h^1 \cup \mathcal{T}_h^2 \cup \mathcal{T}_h^3$  of  $\mathcal{T}_h$  with respect to  $q$  defined by

$$\begin{aligned}\mathcal{T}_h^1 &:= \left\{ T \in \mathcal{T}_h : \frac{1}{|T|} \int_T q(\mathbf{y}) d\mathbf{y} \leq \sup_{\mathbf{y} \in \bar{T}} \alpha(\mathbf{y}) \right\}, \\ \mathcal{T}_h^2 &:= \left\{ T \in \mathcal{T}_h : \frac{1}{|T|} \int_T q(\mathbf{y}) d\mathbf{y} \geq \inf_{\mathbf{y} \in \bar{T}} \beta(\mathbf{y}) \right\}, \\ \mathcal{T}_h^3 &:= \left\{ T \in \mathcal{T}_h : \sup_{\mathbf{y} \in \bar{T}} \alpha(\mathbf{y}) < \frac{1}{|T|} \int_T q(\mathbf{y}) d\mathbf{y} < \inf_{\mathbf{y} \in \bar{T}} \beta(\mathbf{y}) \right\}.\end{aligned}$$

Then for  $T \in \mathcal{T}_h^3$ ,  $p_h(\mathbf{x}) = q_h(\mathbf{x})$  for all  $\mathbf{x} \in T$  and, therefore,

$$\begin{aligned}\|p_h - q_h\|_{L^2(\Omega)}^2 &= \sum_{T \in \mathcal{T}_h} \int_T |p_h(\mathbf{x}) - q_h(\mathbf{x})|^2 d\mathbf{x} \\ &= \sum_{T \in \mathcal{T}_h^1} |T| \left( \frac{1}{|T|} \int_T q(\mathbf{y}) d\mathbf{y} - \sup_{\mathbf{y} \in \bar{T}} \alpha(\mathbf{y}) \right)^2 + \sum_{T \in \mathcal{T}_h^2} |T| \left( \frac{1}{|T|} \int_T q(\mathbf{y}) d\mathbf{y} - \inf_{\mathbf{y} \in \bar{T}} \beta(\mathbf{y}) \right)^2 \\ &\leq \sum_{T \in \mathcal{T}_h^1} |T| (\sup_{\mathbf{y} \in \bar{T}} \alpha(\mathbf{y}) - \inf_{\mathbf{y} \in \bar{T}} \alpha(\mathbf{y}))^2 + \sum_{T \in \mathcal{T}_h^2} |T| (\sup_{\mathbf{y} \in \bar{T}} \beta(\mathbf{y}) - \inf_{\mathbf{y} \in \bar{T}} \beta(\mathbf{y}))^2,\end{aligned}$$

where in the last step we used  $\alpha(\mathbf{x}) \leq q(\mathbf{x}) \leq \beta(\mathbf{x})$  for a.e.  $\mathbf{x} \in \Omega$ . Since  $\alpha$  and  $\beta$  are continuous on  $\bar{\Omega}$ , they are uniformly continuous on  $\bar{\Omega}$ . Then for any  $\epsilon > 0$ , there exists  $\tau > 0$  such that for any  $0 < h < \tau$  and  $T \in \mathcal{T}_h$ ,

$$\sup_{\mathbf{y} \in \bar{T}} \alpha(\mathbf{y}) - \inf_{\mathbf{y} \in \bar{T}} \alpha(\mathbf{y}) < \epsilon, \quad \sup_{\mathbf{y} \in \bar{T}} \beta(\mathbf{y}) - \inf_{\mathbf{y} \in \bar{T}} \beta(\mathbf{y}) < \epsilon.$$

Therefore,

$$\|p_h - q_h\|_{L^2(\Omega)}^2 \leq \epsilon^2 \sum_{T \in \mathcal{T}_h^1} |T| + \sum_{T \in \mathcal{T}_h^2} |T| \leq \epsilon^2 |\Omega|.$$

This implies  $\|q_h - p_h\|_{L^2(\Omega)} \rightarrow 0$  as  $h \rightarrow 0$ . Thus  $q_h \rightarrow q$  in  $L^2(\Omega)$  as  $h \rightarrow 0$ . ■