



# Min–Max formulas for nonlocal elliptic operators on Euclidean Space<sup>☆</sup>

Nestor Guillen<sup>a</sup>, Russell W. Schwab<sup>b,\*</sup>

<sup>a</sup> Department of Mathematics University of Massachusetts, Amherst, Amherst, MA 01003-9305, United States

<sup>b</sup> Department of Mathematics Michigan State University, 619 Red Cedar Road, East Lansing, MI 48824, United States

## ARTICLE INFO

### Article history:

Received 30 November 2018

Accepted 19 February 2019

Communicated by Enzo Mitidieri

### MSC:

35J99

35R09

45K05

46T99

47G20

49L25

49N70

60J75

93E20

### Keywords:

Global comparison principle

Integro-differential operators

Isaacs equation

Whitney extension

Dirichlet-to-Neumann

Fully nonlinear equations

## ABSTRACT

An operator satisfies the Global Comparison Property if anytime a function touches another from above at some point, then the operator preserves the ordering at the point of contact. This is characteristic of degenerate elliptic operators, including nonlocal and nonlinear ones. In previous work, the authors considered such operators in Riemannian manifolds and proved they can be represented by a min–max formula in terms of Lévy operators. In this note we revisit this theory in the context of Euclidean space. With the intricacies of the general Riemannian setting gone, the ideas behind the original proof of the min–max representation become clearer. Moreover, we prove new results regarding operators that commute with translations or which otherwise enjoy some spatial regularity.

©2019 Elsevier Ltd. All rights reserved.

## 1. Introduction

A map  $I : C_b^2(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$  is said to satisfy the **Global Comparison Property** (GCP) if

$$u \leq v \text{ in } \mathbb{R}^d \text{ and } u(x) = v(x) \Rightarrow I(u, x) \leq I(v, x). \quad (1.1)$$

<sup>☆</sup> The authors gratefully acknowledge partial support from the National Science Foundation, United States while this work was in progress: N. Guillen DMS-1700307 and R. Schwab DMS-1665285. The authors also thank the anonymous referee for some suggestions that we believe improved the presentation of our results.

\* Corresponding author.

E-mail addresses: [nguillen@math.umass.edu](mailto:nguillen@math.umass.edu) (N. Guillen), [rschwab@math.msu.edu](mailto:rschwab@math.msu.edu) (R.W. Schwab).

The Laplacian operator, as well as its fractional powers  $-(-\Delta)^{\alpha/2}$  ( $\alpha \in (0, 2)$ ) all satisfy this property. More generally, given a Lévy measure  $\nu(dy)$  (a measure on  $\mathbb{R}^d \setminus \{0\}$  such that  $\min\{1, |y|^2\}$  is integrable with respect to  $\nu$ ) the operator

$$I(u, x) = \int_{\mathbb{R}^d} u(x + y) - u(x) - \chi_{B_1}(y) \nabla u(x) \cdot y \nu(dy),$$

will have the GCP. The GCP is also satisfied by Dirichlet-to-Neumann maps for elliptic equations, generators of Markov processes, Bellman–Isaacs operators in control and differential games, among many examples. When the operator is known a priori to be local, then nonlinear examples of maps with the GCP are of the form,

$$I(u, x) = F(D^2u(x), \nabla u(x), u(x)),$$

where  $F : \mathbb{S}_d \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$  is monotone in its first argument, and Lipschitz continuous in all arguments.

The main contribution of this article is to address when certain operators acting on  $C_b^2(\mathbb{R}^d)$  must necessarily enjoy a structure similar to those examples above. The canonical object used to address this question will be a linear operator we choose to say is “of Lévy type”: those operators for which there exist functions,  $A(x) \in \mathbb{S}_d$ ,  $B(x) \in \mathbb{R}^d$ ,  $C(x) \in \mathbb{R}$ , and measures  $\mu(x, dy)$  so that

$$\begin{aligned} L(u, x) &= \text{tr}(A(x)D^2u(x)) + B(x) \cdot \nabla u(x) + C(x)u(x) \\ &+ \int_{\mathbb{R}^d} u(x + y) - u(x) - \mathbb{1}_{B_1(0)}(y) \nabla u(x) \cdot y \mu(x, dy), \\ &\text{with } A(x) \geq 0, \text{ and } \sup_x \int_{\mathbb{R}^d} \min(|y|^2, 1) \mu(x, dy) < \infty. \end{aligned} \tag{1.2}$$

We will review some recent results that show for  $I : C_b^2(\mathbb{R}^d) \rightarrow C_b(\mathbb{R}^d)$  that enjoys the GCP, is Lipschitz, and has a natural structural constraint, there exists a family of functions,  $f_{ab}$  and linear operators of Lévy type,  $L_{ab}$ , so that

$$I(u, x) = \min_a \max_b \{f_{ab}(x) + L_{ab}(u, x)\}. \tag{1.3}$$

For linear operators, in the 1960s Courrège [19] showed that all of those that satisfy the GCP must have the form given in (1.2). All of our results here should be considered an extension of Courrège’s result to the nonlinear setting.

In our previous work, [28], we showed such a min–max representation in (1.3). The result in [28] in fact dealt with a more general situation where  $I : C_b^2(M) \rightarrow C_b^0(M)$  where  $M$  is a complete Riemannian manifold. We will review the proof of this result in the context of Euclidean space, where many of the arguments simplify greatly. Moreover, we prove two refinements of the main result from [28] relevant to the Euclidean case, one involving translation invariant operators and one for operators that behave continuously with respect to translation operators. Stated informally, our results are the following:

**Theorem 1.** *An operator  $I(u, x)$  that is Lipschitz and satisfies the GCP admits a min–max formula in terms of Lévy type operators.*

**Theorem 2.** *In the previous theorem, assume further that  $I(u, x)$  commutes with translations. Then the Lévy operators appearing in the min–max formula all commute with translations.*

**Theorem 3.** *Instead of translation invariance assume that the finite differences of  $I(u, x)$  commute with translations up to a certain error depending on a modulus of continuity  $\omega(\cdot)$ . Then the Lévy operators appearing in the min–max formula have continuous coefficients with common modulus of continuity of the form  $C\omega(2(\cdot))$ .*

Theorem 1 is a special case of the main result in [28], and Theorems 2 and 3 are new.

### 1.1. Assumptions and main results

Here are our main assumptions.

**Assumption 1.1.** The map  $I : C_b^2(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$  is Lipschitz continuous and has the Global Comparison Property (1.1).

**Assumption 1.2.** The map  $I : C_b^2(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$  is translation invariant. Namely, for any  $x, z \in \mathbb{R}^d$  and  $u \in C_b^2(\mathbb{R}^d)$  we have

$$I(\tau_z u, x) = I(u, x + z), \text{ where } \tau_z u(x) := u(x + z). \quad (1.4)$$

**Assumption 1.3.** There is a non-increasing function  $\rho : (0, \infty) \rightarrow \mathbb{R}$  with  $\rho(R) \rightarrow 0$  as  $R \rightarrow \infty$  such that if  $u, v \in C_b^2(\mathbb{R}^d)$  are such that  $u \equiv v$  in  $B_{2R}(x_0)$ , then

$$\|I(u) - I(v)\|_{L^\infty(B_R(x_0))} \leq \rho(R) \|u - v\|_{L^\infty(\mathbb{R}^d)}.$$

**Assumption 1.4.** There exists a modulus,  $\omega$ , for all  $v, u \in C_b^2(\mathbb{R}^d)$ ,  $x, z \in \mathbb{R}^d$ ,  $r > 0$ , we have

$$\begin{aligned} & |I(v + \tau_{-z} u, x + z) - I(v, x + z) - (I(v + u, x) - I(v, x))| \\ & \leq \omega(|z|) C(r) \left( \|u\|_{C^2(B_{2r}(x))} + \|u\|_{L^\infty(CB_r(x))} \right). \end{aligned}$$

It is allowed that  $C(r) \rightarrow \infty$  as  $r \rightarrow 0$ ; in some examples  $C(r)$  may be bounded and in some it may be unbounded.

The meaning of [Assumptions 1.1](#) and [1.2](#) is self-evident. [Assumption 1.3](#) seems rather technical, but it will be necessary to obtain compactness for a family of measures arising in the proof (and this assumption is satisfied by a broad family of examples). Note however that this assumption is not needed for the translation invariant case as well as the setting of [Theorem 1.9](#) as these two theorems are obtained with different methods.

Last but not least, [Assumption 1.4](#) can be thought of as a “coefficient regularity” assumption. For instance, in the linear and local case, in which  $I$  is a Lévy operator without integral part, [Assumption 1.4](#) is equivalent to the coefficients of the operator having modulus of continuity  $C\omega(\cdot)$  for some constant  $C > 0$ . In fact, [Assumption 1.4](#) is stated so that it indeed linearizes to this usual assumption that one expects in the linear case.

**Remark 1.5.** As mentioned above, one can check that for linear operators, [Assumption 1.4](#) is equivalent to the coefficients of the local part being uniformly continuous and the Lévy measures being uniformly continuous in the TV norm along shifts in the base point, i.e.

$$\|\mu(x + x, \cdot) - \mu(x, \cdot)\|_{TV(CB_r)} \leq C\omega(|z|).$$

By its design, [Assumption 1.4](#) is a technical artifact of our proof, and as such, it is unlikely to be sharp or even the most natural assumption. There is most likely room for improvement here. In fact, one indication of the possibility to make a more natural assumption lies in the fact that even when the original operator,  $I$ , is translation invariant (so the most regular dependence on  $x$ ), it does not necessarily follow that  $I$  also satisfies [Assumption 1.4](#). This also reflects the fact that we have taken two completely different methods of proof for the results that concern translation invariant operators, and ones that have a modulus with respect to translations.

**Remark 1.6.** In Section 6, we give a short list of some operators that fall within the scope of Assumptions 1.1–1.4 and Theorems 1.9–1.14. At the end of Section 6, we give a list of which assumptions each example satisfies.

**Remark 1.7.** We note that one subtle improvement of the current work upon our previous one in [28] is that because of a more streamlined proof for the translation invariant case, we were able to establish the non-translation invariant case, Theorem 1.9 (below), without the technical Assumption 1.3. This is purely an artifact of using an approximation scheme in [28] to treat all operators by the same method, and this turns out to have been not essential when one does not want the extra information provided by Theorems 1.11 and 1.14.

The first theorem uses the notion of “pointwise”  $C^2$  or  $C^1$ , and so we will define that property here.

**Definition 1.8.** For a fixed  $x$  we say that  $u \in C^2(x)$  (“pointwise  $C^2$  at  $x$ ”) if there exists a vector,  $\nabla u(x)$ , and a symmetric matrix,  $D^2u(x)$ , such that

$$\text{as } y \rightarrow x, \quad \left| u(y) - u(x) - \nabla u(x) \cdot (y - x) - \frac{1}{2}(y - x) \cdot (D^2u(x)(y - x)) \right| \leq o(|y - x|^2).$$

Similarly if  $u$  only enjoys the existence of  $\nabla u(x)$  and

$$\text{as } y \rightarrow x, \quad |u(y) - u(x) - \nabla u(x) \cdot (y - x)| \leq o(|y - x|),$$

we say that  $u \in C^1(x)$  (“pointwise  $C^1$  at  $x$ ”).

Now we can restate Theorems 1–3, in more precise terms.

**Theorem 1.9.** If  $I : C_b^2(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$  satisfies Assumption 1.1, then, for each  $x$ , there exists a family of linear functionals on  $C^2(x)$  that depend on  $I$  and  $x$ , called  $\mathcal{K}(I)_x$ , so that for all  $u \in C^2(x)$

$$I(u, x) = \min_{v \in C_b^2(\mathbb{R}^d)} \max_{L \in \mathcal{K}(I)_x} \{I(v, x) + L(u - v)\}.$$

Here, each  $L \in \mathcal{K}(I)_x$ , has the form

$$L(u) = \text{tr}(A_x D^2u(x)) + B_x \cdot \nabla u(x) + C_x u(x) + \int_{\mathbb{R}^d} u(x + y) - u(x) - \mathbb{1}_{B_1(0)}(y) \nabla u(x) \cdot y \mu_x(dy),$$

and for some universal  $C$ , the terms also satisfy the bound for all  $x$ :

$$|A_x| + |B_x| + |C_x| + \int_{\mathbb{R}^d} \min\{1, |y|^2\} \mu_x(dy) \leq C \|I\|_{Lip, C_b^2 \rightarrow C_b^0}.$$

The proof of Theorem 1.9 appears in Section 3.1, which is at the end of Section 3.

We want to point out to the reader that the notation in Theorem 1.9 is intentional in its use of subscripts for e.g.  $A_x$ , etc. This is because our construction does not actually produce  $L$  as a linear mapping  $C_b^2 \rightarrow C_b^0$ , and so it is not correct to think of having a family of  $L$  whose coefficients are actually *functions* of  $x$ . Rather, it just says that at each  $x$  there is a family of functionals that have the desired structure, but it is not clear that they can be put together across all  $x$  to make a family of  $x$ -dependent operators.

This situation changes under other assumptions, and in the next two theorems, our method produces a family of linear operators mapping  $C_b^2(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$ , all of the form (1.2).

**Theorem 1.10.** *If  $I : C_b^2(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$  satisfies Assumptions 1.1 and 1.2 then there exists a family,  $\{f_{ab}, L_{ab}\}_{a,b \in \mathcal{K}(I)}$ , that depends only on  $I$ , where for all  $a, b$ ,  $f_{ab}$  are constants, and  $L_{ab}$  are linear translation invariant operators mapping  $C_b^2(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$  of the form (1.2) (i.e. constant coefficients), and for all  $u \in C_b^2(\mathbb{R}^d)$  and  $x \in \mathbb{R}^d$  we have*

$$I(u, x) = \min_a \max_b \{f_{ab} + L_{ab}(u, x)\}.$$

Furthermore, for a universal  $C$ , for all  $f_{ab}$  and  $L_{ab}$ ,

$$|f_{ab}| + |A_{ab}| + |B_{ab}| + |C_{ab}| + \int_{\mathbb{R}^d} \min\{1, |y|^2\} \mu_{ab}(dy) \leq C \|I\|_{Lip, C_b^2 \rightarrow C_b^0}.$$

The proof of Theorem 1.10 appears in Section 3.1, which is at the end of Section 3.

**Theorem 1.11.** *If  $I : C_b^2(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$  satisfies Assumptions 1.1, 1.3, and 1.4, then, there exists a family,  $\{f_{ab}, L_{ab}\}_{a,b \in \mathcal{K}(I)}$ , that depends only on  $I$ , where for all  $a, b$ ,  $f_{ab} \in C_b^0(\mathbb{R}^d)$  are functions, and  $L_{ab}$  are linear operators mapping  $C_b^2(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$  of the form (1.2), and for all  $u \in C_b^2(\mathbb{R}^d)$ , we have*

$$I(u, x) = \min_a \max_b \{f_{ab}(x) + L_{ab}(u, x)\},$$

and for a universal  $C$ , for all  $f_{ab}$  and  $L_{ab}$ ,

$$\|f_{ab}\|_{L^\infty} + \|A_{ab}\|_{L^\infty} + \|B_{ab}\|_{L^\infty} + \|C_{ab}\|_{L^\infty} + \sup_x \int_{\mathbb{R}^d} \min\{1, |y|^2\} \mu_{ab}(x, dy) \leq C \|I\|_{Lip, C_b^2 \rightarrow C_b^0}.$$

Furthermore, if  $\omega$  is as in Assumption 1.4, then the functions  $f_{ab}, A_{ab}, B_{ab}, C_{ab}$ , all have a modulus of continuity  $C\omega(2\cdot)$ , while for each  $r > 0$  we have the estimate,

$$\|\mu_{ab}(x_1) - \mu_{ab}(x_2)\|_{\text{TV}(CB_r)} \leq C(r)\omega(2|x_1 - x_2|), \quad (1.5)$$

where as above,  $C(r) > 0$ , is a constant that may possibly (but not necessarily) have the property that  $C(r) \rightarrow \infty$  as  $r \rightarrow 0$ .

The proof of Theorem 1.11 appears in Section 5.5, which is at the end of Section 5.

Finally, we give a theorem that reduces the possible terms in the min–max over (1.2). Namely, there are instances in which there may be no second order terms or first order terms. To state this, we abuse notation slightly, and we give a shorthand as  $C_b^\beta(\mathbb{R}^d)$  to mean the following:

$$\begin{aligned} &\text{if } \beta = 2 + \gamma, \text{ for } \gamma \in (0, 1), \text{ then, we mean } C_b^\beta(\mathbb{R}^d) = C_b^{2,\gamma}(\mathbb{R}^d); \\ &\text{if } \beta = 2^+, \text{ then, we mean } C_b^\beta(\mathbb{R}^d) = C_b^2(\mathbb{R}^d); \\ &\text{if } \beta = 2, \text{ then, we mean } C_b^\beta(\mathbb{R}^d) = C_b^{1,1}(\mathbb{R}^d); \\ &\text{if } \beta = 1 + \gamma, \text{ for } \gamma \in (0, 1), \text{ then, we mean } C_b^\beta(\mathbb{R}^d) = C_b^{1,\gamma}(\mathbb{R}^d); \\ &\text{if } \beta = 1^+, \text{ then, we mean } C_b^\beta(\mathbb{R}^d) = C_b^1(\mathbb{R}^d); \\ &\text{if } \beta = 1, \text{ then, we mean } C_b^\beta(\mathbb{R}^d) = C_b^{0,1}(\mathbb{R}^d); \\ &\text{if } \beta = \gamma, \text{ for } \gamma \in (0, 1), \text{ then, we mean } C_b^\beta(\mathbb{R}^d) = C_b^{0,\gamma}(\mathbb{R}^d). \end{aligned} \quad (1.6)$$

**Definition 1.12.** For a fixed  $x$ , we say that  $u \in C^\beta(x)$  (“pointwise  $C^\beta(x)$ ”) if the same requirements of Definition 1.8 hold, but the estimate on the right hand side takes into account the different decay as follows:

- if,  $\beta = 2 + \gamma$ , then  $u$  has a second order Taylor expansion and the right hand side is  $O(|y - x|^{2+\gamma})$ ;
- if,  $\beta = 2^+$ , then  $u$  has a second order Taylor expansion and the right hand side is  $o(|y - x|^2)$ ;

- if,  $\beta = 2$ , then we include this in the previous case whenever  $u$  has a second order Taylor expansion at  $x$ ;
- if,  $\beta = 1 + \gamma$ , then  $u$  has a first order Taylor expansion and the right hand side is  $O(|y - x|^{1+\gamma})$ ;
- if,  $\beta = 1^+$ , then  $u$  has a first order Taylor expansion and the right hand side is  $o(|y - x|)$ ;
- if,  $\beta = 1$ , then we include this in the previous case whenever  $u$  has a first order Taylor expansion at  $x$ ;
- if,  $\beta = \gamma \in (0, 1)$ , then  $|u(y) - u(x)| \leq C|y - x|^\gamma$ .

**Assumption 1.13.** All of [Assumptions 1.1–1.4](#) hold, but with all instances of  $C_b^2(\mathbb{R}^d)$  replaced by  $C_b^\beta(\mathbb{R}^d)$ .

**Theorem 1.14.** *For each of [Theorems 1.9, 1.10, 1.11](#), we have the following variation: in each case assume that  $I$  satisfies [Assumption 1.13](#), for some  $\beta \in [0, 2^+]$  (as enumerated above). Then, taking into account [Definition 1.12](#) for [Theorem 1.9](#), the min–max formula holds in each of the previous results with the following additions: if  $\beta < 2$  then  $A_{ab} = 0$  for all  $a, b$ , while if  $\beta < 1$  then  $B_{ab} = 0$  for all  $a, b$  and the operators  $L_{ab}$  take the form*

$$L_{ab}(u, x) = C_{ab}(x)u(x) + \int_{\mathbb{R}^d} u(x + y) - u(x) \mu_{ab}(x, dy).$$

Moreover, the smaller  $\beta$ , the more regular the Lévy measures  $\mu_{ab}$  are at  $y = 0$ , namely, we have

$$\sup_{a, b, x} \int_{\mathbb{R}^d} \min\{1, |y|^\beta\} \mu_{ab}(x, dy) < \infty.$$

The proof of [Theorem 1.14](#) appears in [Section 5.5](#), which is at the end of [Section 5](#).

**Remark 1.15.** In [Sections 4](#) and [5](#), one can see that at its heart, the fact that the modulus for  $I$  is passed onto the coefficient functions in [\(1.2\)](#) is a consequence of our choice to use a Whitney extension in an approximation to  $I$ , and the Whitney extension is well known to preserve a modulus of continuity. The actual details are a bit more involved, but that is the main reason. We note the presence of the factor of 2 in the new modulus is a consequence of the Whitney Extension method; the interested reader can see [\[54, Chapter VI\]](#).

A further comment regarding the assumptions is in order. Suppose that  $I$  satisfies [Assumption 1.4](#) with  $\omega \equiv 0$ . In this case, taking  $v \equiv 0$  the assumption says that

$$I(\tau_{-h}u, x + h) - I(0, x + h) = I(u, x) - I(0, x),$$

and if we further assume that  $I(0, x)$  is constant (i.e.  $I$  applied to the zero function returns a constant), then we have

$$I(\tau_{-h}u, x + h) = I(u, x),$$

that is,  $I$  is translation invariant. However, at first sight it is not clear what happens in the reverse direction. That is, we do not know how to show that a translation-invariant operator automatically satisfies [Assumption 1.4](#) with  $\omega \equiv 0$ , and in fact we expect that this assumption can be modified so that it seamlessly includes the translation invariant operators as well.

## 1.2. Notation

For the readers' convenience, a summary of symbols used in the paper is presented below.

Notation	Definition
$d$	Space dimension
$C_b^2$	Twice differentiable functions $f$ with bounded $f, \nabla f$ , and $D^2 f$
$C_b^\beta$	Bounded functions of class $C^\beta$ , see (1.6) for definition
$\mathbb{S}_d$	Symmetric matrices of size $d \times d$
$\ \cdot\ _{\text{TV}}$	Total variation norm for a measure
$L(X, Y)$	Space of bounded linear operators from $X$ to $Y$
$\text{c.h.}(E)$	The convex hull of a set $E$
$CE$	Complement of a subset of $\mathbb{R}^d$
$F^0(x, v)$	Upper gradient of a Lipschitz function (Definition 2.1)
$\partial F(x)$	Generalized gradient of $F$ at $x$ (Definition 2.3)
$G_n$	Grid with step size $2^{-n}$
$C(G_n)$	Space of real valued functions defined in $G_n$ (Definition 4.1)
$C_*(G_n)$	Subset of $C(G_n)$ of functions vanishing outside $[-2^n, 2^n] \cap G_n$ (Definition 4.1)
$(\nabla_n)^1 u(x)$	Discrete gradient for step size $2^{-n}$ (Definition 4.7)
$(\nabla_n)^2 u(x)$	Discrete Hessian for step size $2^{-n}$ (Definition 4.8)

### 1.3. Background

There were roughly two reasons that motivated the results we present in this paper. First of all, the link between elliptic equations and a min–max formula for operators has a long history, and it has been exploited extensively in the case of *local* operators. Until [28], the connection was not known for nonlocal, nonlinear operators. Even so, the link between the two was natural enough that there are at least a few results that assumed a structure like (1.3), including [5,35,40,47,48,51], among many others. Thus the theorems here and in [28] give a sort of a posteriori justification to min–max assumptions that appeared in earlier works. Secondly, a formula such as (1.3) can be very useful in connecting results about the integro-differential theory (of which, there has been a large volume recently) with some other pursuits that may not obviously relate to operators such as (1.2). Two recent projects that exploit or were motivated by the min–max formulas are on some Hele-Shaw type free boundary evolutions in [16] and some Neumann homogenization problems [30,31]. Both of these relate to linear and nonlinear Dirichlet-to-Neumann maps, studied in [26], and there is plenty more to learn about the integro-differential structure in the nonlinear setting. The choice to pursue continuity properties such as the dependence given in (1.5), although a posteriori seems straightforward, was not initially obvious, and it was motivated by recent results about comparison theorems for viscosity solutions of integro-differential equations in [27].

As mentioned earlier, for linear operators, the representation of (1.2) goes back to Courrèges [19]. This was naturally connected with generators of Markov processes and boundary excursion processes for reflected diffusions. Hsu [32] provides a similar representation for the Dirichlet to Neumann map for the Laplacian in a smooth domain  $\Omega$ , and this corresponds to studying the boundary process for a reflected Brownian motion. If  $I$  is not necessarily linear but happens to satisfy the stronger *local* comparison principle, there are min–max results by many authors, e.g. Evans [21], Souganidis [53], Evans–Souganidis [22] and Katsoulakis [38]. In this case, the operator takes the form,

$$I(u, x) = F(x, u(x), \nabla u(x), D^2 u(x)),$$

which can be expressed as in Theorem 1.9, but with  $\mu(x, dh) \equiv 0$ . This was extended to even include the possibility of weak solutions acting as a *local* semi-group on  $BUC(\mathbb{R}^d)$ , related to image processing, in Alvarez–Guichard–Lions–Morel [1], and to weak solutions of sets satisfying an order preserving set flow by Barles–Souganidis in [6]. In [1] it was shown under quite general assumptions that certain nonlinear semigroups must be represented as the unique viscosity solution to a degenerate parabolic equation.

Although it is still too early to tell, one hopes that theorems like those presented here can create a bridge between some nonlocal equations for which regularity questions arise and the known results about such equations when a min–max structure is known to hold. In the *local* setting, there are a number of results that leverage the min–max to shed new light on certain issues, and it would be interesting to see if similar things can be done for the nonlocal theory (see the discussion in [28, Section 1] for an incomplete list of such results). The types of regularity results that could find new applications via the min–max theorems here fall into roughly three categories: Krylov–Safonov type results; regularity for translation invariant equations; and Schauder type regularity results. For Krylov–Safonov, this means that solutions of fully nonlinear equations can be shown to enjoy Hölder estimates depending only on the  $L^\infty$  norm of the solution; some examples are: [10, 14, 15, 37], and [49], among many others. For translation invariant equations, these are the results that show solutions to translation invariant equations very often enjoy  $C^{1,\alpha}$  regularity under mild assumptions; some examples are: [10, 17, 41, 44, 50], among others. Finally, for Schauder regularity, we mean results that show that for  $x$ -dependent operators, under certain regularity for the coefficients (such as Dini), solutions will have as much regularity as those equations with “constant coefficients”; some examples are: [20, 36, 43], among others. On top of questions of the type of Krylov–Safonov regularity mentioned above, there is another family of regularity results that accompanies existence and uniqueness techniques for viscosity solutions of elliptic partial-differential/integro-differential equations, and it is typically referred to as the Ishii–Lions method, going back to [34]. Both this Ishii–Lions regularity and comparison results could connect well with the operators treated in this paper, as many of the existing works on nonlocal equations assume a min–max. The types of results that could be applicable are like those in [2–5], and [35], among others.

There is some more discussion of related works and background inside of the examples that we list in Section 6.

#### 1.4. Another description of operators satisfying the GCP

Let us describe an elementary but useful way to view operators satisfying the GCP, which is also related to the min–max representation. First, we introduce a family of functional spaces.

**Definition 1.16.** For  $\beta \in [0, 2^+]$  (using the abuse of notation in (1.6)) we define the space  $L_\beta^\infty$  as follows. First, if  $\beta \neq 1^+$ ,

$$L_\beta^\infty := \{h \in L^\infty(\mathbb{R}^d) \mid |h(y)| = O(|y|^\beta) \text{ as } |y| \rightarrow 0\},$$

while for  $\beta = 1^+$ ,

$$L_\beta^\infty := \{h \in L^\infty(\mathbb{R}^d) \mid |h(y)| = o(|y|^\beta) \text{ as } y \rightarrow 0\}.$$

(We note the first space requires “Big-O”, while the second space requires “little-o”.) The spaces  $L_\beta^\infty$  are Banach spaces, with norms given by

$$\sup_y |h(y)| \min\{1, |y|^\beta\}^{-1}.$$

Now, suppose we are given a continuous function

$$F : L_\beta^\infty(\mathbb{R}^d) \times \mathbb{S}_d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}.$$

Assume that this function is monotone (non-decreasing) with respect to the first two variables. Then, given  $u \in C_b^\beta(\mathbb{R}^d)$  define

$$I(u, x) := F(\delta_x u, D^2 u(x), \nabla u(x), u(x), x)$$

where we are using the notation  $\delta_x u(y) := u(x + y) - u(x) - \nabla u(x) \cdot y \chi_{B_1(0)}(y)$  for  $\beta \geq 1$ , and  $\delta_x u(y) := u(x + y) - u(x)$  for  $\beta < 1$ . It is clear the operator  $I$  thus defined has the GCP.

Do all operators with the GCP arise in this form? It is easy to see that the answer is positive, at least when  $\beta < 2$ . Given  $I : C^\beta(\mathbb{R}^d) \rightarrow C^0(\mathbb{R})$ , with  $\beta < 2$ , we define a function

$$F : L_\beta^\infty(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R},$$

by the formula  $F(h, p, u, x) := I(\tau_{-x}h + \tau_{-x}p \cdot (\cdot)\chi_{B_1} + u, x)$ . It is straightforward to see that for  $F$  so defined and  $u \in C_b^\beta(\mathbb{R}^d)$  we have

$$I(u, x) = F(\delta_x u, \nabla u(x), u(x), x).$$

## 2. Real valued Lipschitz functions on Banach spaces

In this section we review various well known facts about Lipschitz functions on Banach spaces, following Clarke's book [18, Chapter 2]. We will refer most of the proofs to the relevant section in [18]. The section ends with **Theorem 2.6** which yields a min–max formula for any real valued, Lipschitz  $F$ , such a result is neither new nor surprising, but we present it here in complete detail for the sake of completeness.

We fix a Banach Space, denoted by  $X$ , an open convex subset  $\mathcal{K} \subset X$ , and a function

$$F : \mathcal{K} \subset X \rightarrow \mathbb{R},$$

which is assumed Lipschitz with constant  $L > 0$ , that is

$$|F(x) - F(y)| \leq L\|x - y\| \quad \forall x, y \in \mathcal{K}. \quad (2.1)$$

**Definition 2.1.** The upper gradient of  $F$  at  $x \in \mathcal{K}$  in the direction of  $v \in X$ , is defined as

$$F^0(x, v) := \limsup_{t \searrow 0} \frac{F(x + tv) - F(x)}{t}.$$

This can be seen as a function  $F^0 : \mathcal{K} \times X \rightarrow \mathbb{R}$ .

**Proposition 2.2.** *The function  $F^0(x, v)$  has the following properties*

- (1) *For any  $x \in \mathcal{K}, v \in X$ , and  $\lambda > 0$  we have  $F^0(x, \lambda v) = \lambda F^0(x, v)$ .*
- (2) *For any  $x \in \mathcal{K}$ , and  $v, w \in X$  we have  $|F^0(x, v) - F^0(x, w)| \leq L\|v - w\|$ .*
- (3) *If  $(x_k, v_k) \rightarrow (x, v)$  then  $\limsup F^0(x_k, v_k) \leq F^0(x, v)$ .*
- (4)  *$F^0(x, -v) = (-F)^0(x, v)$ .*

**Proof.** We refer the reader to [18, Proposition 2.1.1].  $\square$

**Definition 2.3.** The generalized gradient of  $F$  at  $x \in \mathcal{K}$  is the subset of  $X^*$  given by

$$\partial F(x) := \{\ell \in X^* \mid F^0(x, v) \geq \langle \ell, v \rangle \quad \forall v \in X\}.$$

We will denote by  $\partial F$  the convex hull of the union of  $\partial F(x)$ ,

$$\partial F := \text{c.h.} \left( \bigcup_{x \in \mathcal{K}} \partial F(x) \right).$$

**Proposition 2.4.** *The set  $\partial F(x)$ ,  $x \in \mathcal{K}$ , has the following properties*

- (1)  $\partial F(x)$  is a non-empty, convex, weak\*-compact subset of  $X^*$ .
- (2)  $\|\ell\| \leq L$  for every  $\ell \in \partial F(x)$ .
- (3) For any  $v \in X$ , we have that

$$F^0(x, v) = \max_{\ell \in \partial F(x)} \langle \ell, v \rangle.$$

**Proof.** We refer the reader to [18, Proposition 2.1.2].  $\square$

The following theorem, due to Lebourg, is a generalization of the mean value theorem for differentiable functions.

**Theorem 2.5** (Lebourg's Theorem). *Let  $x, y$  be points in  $\mathcal{K}$ . Then there exist  $z$  of the form  $z = tx + (1 - t)y$  for some  $t \in [0, 1]$ , such that for some  $\ell \in \partial F(z)$*

$$F(x) - F(y) = \langle \ell, x - y \rangle.$$

**Proof.** We refer the reader to [18, Theorem 2.3.7].  $\square$

Using the generalized gradient and Lebourg's theorem we can easily prove a min–max formula for Lipschitz functionals. Observe this is a general result for Lipschitz functionals in general Banach spaces, and it does not involve anything like GCP (functionals with the GCP on  $C_b^\beta(\mathbb{R}^d)$  are considered in the next section).

**Theorem 2.6.** *Let  $F : \mathcal{K} \subset X \rightarrow \mathbb{R}$  be a Lipschitz function, with  $\mathcal{K}$  convex, then for all  $x \in \mathcal{K}$ ,*

$$F(x) = \min_{y \in \mathcal{K}} \max_{\ell \in \partial F} \{F(y) + \langle \ell, y - x \rangle\}.$$

**Proof.** According to Theorem 2.5, given  $x, y \in \mathcal{K}$  there is some  $\ell \in \partial F$  such that

$$F(x) - F(y) = \langle \ell, x - y \rangle.$$

In other words, for any  $x$  and  $y$  in  $\mathcal{K}$  we have the inequality

$$F(x) \leq \max_{\ell \in \partial F} \{F(y) + \langle \ell, x - y \rangle\}.$$

This also yields an equality for  $y = x$ , thus  $F(x) = \min_{y \in \mathcal{K}} \max_{\ell \in \partial F} \{F(y) + \langle \ell, x - y \rangle\}$ .  $\square$

### 3. Functionals with the GCP, revisited

Throughout this section  $\mathcal{K}$  denotes an open convex set of  $C_b^\beta(\mathbb{R}^d)$  (see (1.6)). Moreover, for  $\rho > 0$ , we shall write

$$\mathcal{K}_\rho = \{u \in C_b^\beta(\mathbb{R}^d) \mid \|v - u\|_{C^\beta} < \rho \Rightarrow v \in \mathcal{K}\}.$$

**Definition 3.1.** Let  $F$  be a map  $F : \mathcal{K} \subset C_b^\beta(\mathbb{R}^d) \rightarrow \mathbb{R}$  and  $x \in \mathbb{R}^d$ . Such a functional is said to have the Global Comparison Property with respect to  $x$  if  $F(u) \leq F(v)$  for any pair of functions  $u, v \in \mathcal{K}$  such that  $u(y) \leq v(y)$  for all  $y$  and  $u(x) = v(x)$  –we will say in such a case that  $v$  touches  $u$  from above at  $x$ .

The following two auxiliary functions will be useful throughout the section: Fix  $\phi_0 : \mathbb{R} \rightarrow \mathbb{R}$ , a nondecreasing  $C^\infty$  function such that  $0 \leq \phi_0 \leq 1$ ,  $\phi_0(x) = 0$  for  $x \leq 0$ ,  $\phi_0(x) = 1$  for  $x \geq 1$ . Then, given  $r, R > 0$  we define the functions

$$\phi_{r,R}(y) := \phi_0\left(\frac{|y| - R}{r}\right) \quad (3.1)$$

$$\psi_{r,R}(y) := 1 - \phi_{r,R}(y) \quad (3.2)$$

The following Proposition was first proved in [28, Lemma 4.15, Corollary 4.16], we review the proof here for the reader's convenience.

**Proposition 3.2.** *Suppose that  $F : \mathcal{K} \subset C_b^\beta(\mathbb{R}^d) \rightarrow \mathbb{R}$  is a Lipschitz functional which has the GCP with respect to  $x$ . Fix  $\rho > 0$ . There is a constant  $C(F, \rho)$  such that given  $R > 0$ ,  $r \in (0, 1)$ , and  $u, v \in \mathcal{K}_\rho$ , then*

$$|F(u) - F(v)| \leq C(F, \rho) r^{-\beta} \left( \|u - v\|_{C^\beta(B_{R+r}(x))} + \|u - v\|_{L^\infty(\mathbb{R}^d \setminus B_R(x))} \right).$$

**Remark 3.3.** It is worth comparing [Proposition 3.2](#) with [Assumption 1.3](#). In the latter, one is interested in how  $I(u, x)$  depends very little on the values of  $u$  far away from  $x$  (so, as  $r \rightarrow \infty$ ), whereas the former deals with a weak version of this property that holds only for  $r \in (0, 1)$  but which follows alone from the GCP without the need for further assumptions on  $F$ .

**Proof.** Take  $\phi \in C_b^2(\mathbb{R}^d)$ , such that  $0 \leq \phi \leq 1$  and  $\phi(x) = 0$ . Then, for any  $y$  we have

$$u(y) \leq w(y) := u(y) + \phi(y) (\|u - v\|_{L^\infty(\text{spt}(\phi))} - (u(y) - v(y))),$$

with the above being an equality for  $y = x$ . Now, let  $\rho_0$  be chosen so that

$$2\|\phi\|_{C^2(\mathbb{R}^d)}\rho_0 \leq \rho.$$

Then, let us suppose that  $u, v \in \mathcal{K}_\rho$  are such that  $\|u - v\|_{C_b^\beta(\mathbb{R}^d)} \leq \rho_0$ . In this case, we have  $w \in \mathcal{K}$  since  $u \in \mathcal{K}_\rho$  and in this case the GCP says that

$$F(u) \leq F(w).$$

Moreover,  $F(w) \leq F(v) + L\|w - v\|_{C^\beta}$  and  $w - v = (1 - \phi)(u - v) + \phi\|u - v\|_{L^\infty(\text{spt}(\phi))}$ , thus

$$F(u) - F(v) \leq L\|(1 - \phi)(u - v)\|_{C^\beta} + L\|u - v\|_{L^\infty(\text{spt}(\phi))}\|\phi\|_{C^\beta}.$$

Consider the function  $\phi(y) = \phi_{r,R}(y - x)$ . Thanks to  $r \in (0, 1)$ , the following estimates hold

$$\begin{aligned} \|\phi\|_{C^\beta} &\leq Cr^{-\beta}, \\ \|(1 - \phi)(u - v)\|_{C^\beta} &\leq Cr^{-\beta}\|u - v\|_{C^\beta(B_{R+r})}. \end{aligned}$$

Substituting these in the inequality for  $F(u) - F(v)$ , the desired inequality follows when  $\|u - v\|_{C^\beta}$  is no larger than  $\rho_0$ . Otherwise,  $\|u - v\|_{C^\beta} \geq \rho_0$  and iterating the inequality in the previous case one obtains that

$$|F(u) - F(v)| \leq C(F, \rho) r^{-\beta} \left( \|u - v\|_{C^\beta(B_{R+r}(x))} + \|u - v\|_{L^\infty(\mathbb{R}^d \setminus B_R(x))} \right). \quad \square$$

**Lemma 3.4.** *Let  $F : \mathcal{K} \subset C_b^\beta(\mathbb{R}^d) \rightarrow \mathbb{R}$  be a Lipschitz functional which has the GCP with respect to  $x$ . Then, for every  $\ell \in \partial F$  we have*

$$\langle \ell, v \rangle \leq 0 \text{ if } v \leq 0 \text{ everywhere and } v(x) = 0.$$

In other words, if  $F$  has the GCP with respect to  $x$ , then any  $\ell$  arising as a generalized gradient of  $F$  also has the GCP with respect to  $x$ . Furthermore, for any such  $\ell$  and  $r \in (0, 1)$  we have

$$|\langle \ell, v \rangle| \leq Cr^{-\beta} \left( \|v\|_{C^\beta(B_r)} + \|v\|_{L^\infty(\mathbb{R}^d)} \right).$$

**Proof.** Let  $u \in \mathcal{K}$ , and let  $v \in C_b^\beta(\mathbb{R}^d)$  be such that

$$v \leq 0 \text{ in } \mathbb{R}^d, v(x) = 0.$$

Then,  $u_t = u + tv$  touches  $u$  from below at  $x$  for each small  $t$ , therefore  $F(u_t) \leq F(u)$  for every  $t$ , and

$$F^0(u, v) = \limsup_{t \rightarrow 0} \frac{F(u + tv) - F(u)}{t} \leq 0.$$

Since,

$$\max_{\ell \in \partial F(u)} \langle \ell, v \rangle = F^0(u, v),$$

it follows that  $\langle \ell, v \rangle \leq 0$  for any  $\ell \in \partial F(u)$ , and the first part of the Lemma is proved. For the second part, one argues similarly, except that instead of invoking the GCP, one applies [Proposition 3.2](#) in order to pass the same estimate for any  $\ell \in \partial F$ .  $\square$

Fix a functional  $\ell$  having the GCP with respect to  $x$ . Then, define  $C_\ell$  by

$$C_\ell := \langle \ell, 1 \rangle. \quad (3.3)$$

This associates a constant  $C_\ell$  to any  $\ell$  having the GCP. Likewise, we shall associate a vector  $B_\ell$  and positive semi-definite matrix  $A_\ell$ . First, let us introduce some notation,

$$\mathcal{S} := \{\phi \in C_c^2(B_2(0)) \mid \phi \equiv 1 \text{ in a neighborhood of } 0, 0 \leq \phi \leq 1 \text{ in all of } \mathbb{R}^d\}. \quad (3.4)$$

Given  $\phi, \eta \in \mathcal{S}$ , define the function

$$P_{\phi, \eta, u, x}(\cdot) = \begin{cases} u(x) + \phi(\cdot - x)(\nabla u(x), \cdot - x) + \frac{1}{2}\eta(\cdot - x)(D^2u(x)(\cdot - x), \cdot - x) & \text{if } \beta \in [2, 3), \\ u(x) + \phi(\cdot - x)(\nabla u(x), \cdot - x) & \text{if } \beta \in [1, 2), \\ u(x) & \text{if } \beta \in (0, 1). \end{cases} \quad (3.5)$$

For  $x = 0$  we will simply write  $P_{\phi, \beta, u}$ . Observe that, for example, if  $\beta = 2$  then  $P_{\phi, \eta, u, x}$  is a smooth function which, in a neighborhood of  $x$ , coincides with the second order Taylor polynomial of the function  $u$  at the point  $x$ .

**Definition 3.5.** Given any  $\phi \in \mathcal{S}$  let  $B_{\ell, \phi}$  be the vector defined by

$$(B_{\ell, \phi}, e) = \langle \ell, \phi(\cdot)(\cdot, e) \rangle, \quad \forall \text{ vectors } e.$$

At the same time, given  $\eta \in \mathcal{S}$  let  $A_{\ell, \eta}$  be the symmetric matrix defined by

$$\text{tr}(A_{\ell, \eta} M) = \langle \ell, \eta(\cdot) \frac{1}{2}(M(\cdot), \cdot) \rangle, \quad \forall \text{ symmetric matrices } M.$$

The following lemmas will characterize all of functionals having the GCP with respect to 0 (compare with Courrèges original proof [\[19\]](#), see also [\[28\]](#)).

**Lemma 3.6.** Let  $\ell : C_b^\beta(\mathbb{R}^d) \rightarrow \mathbb{R}$  be a bounded linear functional which has the GCP with respect to 0, and  $\phi, \eta \in \mathcal{S}$  (defined in (3.4)). There is a positive measure  $\mu_\ell$  on  $\mathbb{R}^d \setminus \{0\}$  with

$$\int_{\mathbb{R}^d \setminus \{0\}} \min\{1, |y|^\beta\} \mu_\ell(dy) \leq C\|\ell\|,$$

such that for any  $u \in C_b^\beta(\mathbb{R}^d)$  we have the following representation,

for  $\beta \geq 2$ , and  $u \in C_b^\beta(\mathbb{R}^d) \cap C^2(0)$ ,

$$\langle \ell, u \rangle = C_\ell u(0) + (B_{\ell, \phi}, \nabla u(0)) + \text{tr}(A_{\ell, \eta} D^2 u(0)) + \int_{\mathbb{R}^d} u(y) - P_{\phi, \eta, u}(y) \mu_\ell(dy),$$

for  $\beta \in [1, 2)$ , and  $u \in C_b^\beta(\mathbb{R}^d) \cap C^1(0)$ ,

$$\langle \ell, u \rangle = C_\ell u(0) + (B_{\ell, \phi}, \nabla u(0)) + \int_{\mathbb{R}^d} u(y) - P_{\phi, \eta, u}(y) \mu_\ell(dy),$$

for  $\beta \in (0, 1)$ , and  $u \in C_b^\beta(\mathbb{R}^d)$ ,

$$\langle \ell, u \rangle = C_\ell u(0) + \int_{\mathbb{R}^d} u(y) - u(0) \mu_\ell(dy).$$

(The notation,  $C^2(0)$  and  $C^1(0)$ , appears in [Definition 1.8](#).)

**Remark 3.7.** We want to note that the dependence of  $\mu$  only on  $\ell$  is not a typo. Even though the vector  $B_{\ell, \phi}$  and matrix  $A_{\ell, \eta}$  clearly depend on the functions  $\phi$  and  $\eta$ , the reader can see in the proof in (3.6) that  $\mu_\ell$  does not depend on  $\phi$  or  $\eta$ .

**Proof.** It suffices to prove the representation formula for  $u \in C_b^2(\mathbb{R}^d)$  (even if  $\beta \neq 2$ ), as it trivially extends to all of  $C_b^\beta(\mathbb{R}^d)$  by approximation. We fix  $u \in C_b^2(\mathbb{R}^d) \cap C^2(0)$ . We recall  $P_{\phi, \eta, u}$  is defined in (3.5). Since  $P_{\phi, \eta, u} \in C_b^\beta(\mathbb{R}^d)$  for each fixed  $\phi, \eta$ , we may write

$$u = u - P_{\phi, \eta, u} + P_{\phi, \eta, u},$$

and linearity gives

$$\langle \ell, u \rangle = \langle \ell, P_{\phi, \eta, u} \rangle + \langle \ell, u - P_{\phi, \eta, u} \rangle$$

Let us study each of these two terms. Using the definition of  $C_\ell$ ,  $B_{\ell, \phi}$ , and  $A_{\ell, \eta}$ , we have for  $\beta \geq 2$

$$\begin{aligned} \langle \ell, P_{\phi, \eta, u} \rangle &= u(0)\langle \ell, 1 \rangle + \sum_{i=1}^d \partial_i u(0)\langle \ell, x_i \phi(x) \rangle + \frac{1}{2} \sum_{i,j=1}^d \partial_{ij}^2 u(0)\langle \ell, \eta(x) x_i x_j \rangle \\ &= C_\ell u(0) + (B_{\ell, \phi}, \nabla u(0)) + \frac{1}{2} \text{tr}(A_{\ell, \eta} D^2 u(0)), \end{aligned}$$

as well as the corresponding expressions in the other cases when  $\beta < 2$ . Next, we analyze the second term in the expression for  $\langle \ell, u \rangle$  above, that is

$$\langle \ell, u - P_{\phi, \eta, u} \rangle.$$

First take the case  $\beta \neq 1$ . Given  $w \in C_b^\beta(\mathbb{R}^d)$ , define  $\tilde{w}$  by

$$\tilde{w}(x) := w(x) \frac{|x|^\beta}{1 + |x|^\beta}.$$

Observe that since  $\beta \neq 1$ , the function  $\tilde{1} = |x|^\beta(1 + |x|^\beta)^{-1}$  belongs to  $C_b^\beta(\mathbb{R}^d)$ . The linear transformation  $w \mapsto \tilde{w}$  defines a linear functional  $\tilde{\ell}$  via the relation

$$\langle \tilde{\ell}, w \rangle := \langle \ell, \tilde{w} \rangle.$$

This clearly defines a bounded functional on  $C_b^\beta(\mathbb{R}^d)$ . In fact, however, this functional extends uniquely to a bounded functional in  $C_b^0(\mathbb{R}^d)$ : since  $\tilde{w}$  is touched from above at 0 by the function  $\|w\|_{L^\infty} \tilde{1}$ , the GCP guarantees that

$$|\langle \tilde{\ell}, w \rangle| \leq \|w\|_{L^\infty} \langle \ell, \frac{|x|^\beta}{1+|x|^\beta} \rangle.$$

This shows  $\tilde{\ell}$  is a uniquely defined continuous functional on  $C_b^0(\mathbb{R}^d)$  whose norm as a functional on  $C_b^0(\mathbb{R}^d)$  is no larger than  $\|\ell\| \|\frac{|x|^\beta}{1+|x|^\beta}\|_{C^\beta}$ . It follows there is a measure  $\tilde{\mu}$  such that

$$\langle \tilde{\ell}, w \rangle = \int_{\mathbb{R}^d} w(y) \tilde{\mu}(dy). \quad (3.6)$$

Moreover, since  $\langle \tilde{\ell}, w \rangle \geq 0$  whenever  $w \geq 0$ ,  $\tilde{\mu}(dy)$  is a non-negative measure. Now, since  $u \in C_b^2(\mathbb{R}^d)$ , we have that the function

$$w(x) := (u(x) - P_{\phi, \eta, u}(x)) \frac{1+|x|^\beta}{|x|^\beta},$$

remains continuous as  $x \rightarrow 0$ , so  $w \in C_b^0(\mathbb{R}^d)$  and thus  $\langle \tilde{\ell}, w \rangle$  is well defined. In this case, we have

$$\langle \ell, u - P_{\phi, \eta, u} \rangle = \langle \tilde{\ell}, w \rangle,$$

and we obtain the formula

$$\langle \ell, u - P_{\phi, \eta, u} \rangle = \int_{\mathbb{R}^d} (u(y) - P_{\phi, \eta, u}(y)) \frac{1+|y|^\beta}{|y|^\beta} \tilde{\mu}(dy).$$

In particular, taking  $\mu(dy) := \frac{1+|y|^\beta}{|y|^\beta} \tilde{\mu}(dy)$ , it follows that

$$\int_{\mathbb{R}^d \setminus \{0\}} \min\{1, |y|^\beta\} \mu(dy) \lesssim \|\ell\| \|\frac{|x|^\beta}{1+|x|^\beta}\|_{C^\beta} < \infty,$$

and

$$\langle \ell, u - P_{\phi, \eta, u} \rangle = \int_{\mathbb{R}^d \setminus \{0\}} u(y) - P_{\phi, \eta, u}(y) \mu(dy).$$

Revisiting the expression of  $\ell$ , we have when  $\beta \geq 2$

$$\langle \ell, u \rangle = C_\ell u(0) + (B_{\ell, \phi}, \nabla u(0)) + \frac{1}{2} \text{tr}(A_{\ell, \eta} D^2 u(0)) + \int_{\mathbb{R}^d \setminus \{0\}} u(y) - P_{\phi, \eta, u}(y) \mu(dy),$$

and the analogous formulas follow for the other cases where  $\beta \neq 1$ , per the change in definition of the function  $P_{\phi, \eta, u}$  in (3.5). It remains to consider the case  $\beta = 1$ .

Since  $|x|$  is not a  $C^1$  function, we are going to approximate it by a more regular function. For every small  $\varepsilon > 0$  we repeat the argument above with  $\beta = 1 + \varepsilon$  and conclude that for some  $\mu_\varepsilon$  we have the formula

$$\langle \ell, u \rangle = C_\ell u(0) + (B_{\ell, \phi}, \nabla u(0)) + \int_{\mathbb{R}^d \setminus \{0\}} u(y) - P_{\phi, \eta, u}(y) \mu_\varepsilon(dy),$$

and this measure  $\mu_\varepsilon$  is positive and satisfies the bound

$$\int_{\mathbb{R}^d \setminus \{0\}} \min\{1, |y|^\beta\} \mu(dy) \lesssim \|\ell\| \|\frac{|x|^{1+\varepsilon}}{1+|x|^{1+\varepsilon}}\|_{C^1}.$$

Since

$$\sup_{\varepsilon \in (0, 1)} \|\frac{|x|^{1+\varepsilon}}{1+|x|^{1+\varepsilon}}\|_{C^1} < \infty,$$

it follows that the respective finite measures  $\{\tilde{\mu}_\varepsilon\}_{\varepsilon \in (0,1)}$  have uniformly bounded mass. Therefore, it is not difficult to show (using  $\ell$  to get tightness for the  $\tilde{\mu}_\varepsilon$ ) that along a subsequence  $\varepsilon \rightarrow 0$  we can find a limit  $\tilde{\mu}$ , and if we let  $\mu := (1 + |y|)|y|^{-1}\tilde{\mu}$  then

$$\int_{\mathbb{R}^d \setminus \{0\}} \min\{1, |y|\} \mu(dy) < \infty,$$

and again, for any  $u \in C_b^2(\mathbb{R}^d)$ ,

$$\langle \ell, u \rangle = C_\ell u(0) + (B_{\ell,\phi}, \nabla u(0)) + \int_{\mathbb{R}^d \setminus \{0\}} u(y) - P_{\phi,\eta,u}(y) \mu(dy), \quad \square$$

We consider the following special functions. For  $\delta > 0$ , define (see (3.2) for definition of  $\psi_{r,R}$ )

$$\phi_\delta(x) := \psi_{\delta,1-2\delta}, \quad (3.7)$$

$$\eta_\delta(x) := \psi_{\delta,\delta}(x). \quad (3.8)$$

Note that  $\phi_\delta \equiv 1$  inside  $B_{1-2\delta}$  and  $\phi_\delta \equiv 0$  outside  $B_{1-\delta}$ , while  $\eta_\delta \equiv 1$  inside  $B_\delta$  and  $\eta_\delta \equiv 0$  outside  $B_{2\delta}$ . Furthermore, we note that  $\delta \leq \delta'$  implies that  $\eta_\delta \leq \eta_{\delta'}$ .

**Lemma 3.8.** *Assume that  $\beta \in [0, 3)$ ,  $l : C_b^\beta(\mathbb{R}^d) \rightarrow \mathbb{R}$  is a bounded linear functional with the GCP with respect to 0, and that  $A_{\ell,\eta}$ ,  $B_{\ell,\phi}$  are as in Definition 3.5. Taking  $\eta_\delta$  as in (3.8), the limit*

$$A_\ell := \lim_{\delta \searrow 0} A_{\ell,\eta_\delta},$$

*exists for all  $\beta \in [0, 3)$ , and  $A_\ell \equiv 0$  if  $\beta < 2$ . Moreover, if  $\phi_\delta$  is as in (3.7), there is a sequence  $\delta_k \searrow 0$  such that the following limit exists*

$$B_\ell := \lim_{k \rightarrow \infty} B_{\ell,\phi_{\delta_k}}.$$

**Proof.** Let  $\eta_1, \eta_2 \in \mathcal{S}$  and such that  $\eta_1 \leq \eta_2$ . Then for any positive semi-definite  $M$  we have

$$\frac{1}{2}\eta_1(x)(Mx, x) \leq \frac{1}{2}\eta_2(x)(Mx, x), \quad \text{with equality at } x = 0.$$

Since  $\ell$  has the GCP with respect to 0, it follows that

$$\langle \ell, \frac{1}{2}\eta_1(x)(Mx, x) \rangle \leq \langle \ell, \frac{1}{2}\eta_2(x)(Mx, x) \rangle.$$

From this monotonicity and the elementary inequality  $|\langle \ell, \frac{1}{2}\eta(x)(Mx, x) \rangle| \leq C|M| \max_{ij} \|\eta x_i x_j\|_{C^\beta}$  we conclude that the following limit exists for every positive semi-definite  $M$

$$\lim_{\delta \searrow 0} \langle \ell, \frac{1}{2}\eta_\delta(x)(Mx, x) \rangle.$$

At the same time, when  $\beta < 2$  we have  $\|\eta_\delta x_i x_j\|_{C^\beta} \rightarrow 0$  as  $\delta \searrow 0$  for all  $i, j$ , so in this case the limit is zero. Now, given a symmetric matrix  $M$ , write  $M = M^+ - M^-$ , where both  $M^+$  and  $M^-$  are positive semi-definite. Then, we also have that the limit

$$\lim_{\phi \in \mathcal{S}, \eta \searrow 0} \langle \ell, \frac{1}{2}\eta(x)(Mx, x) \rangle$$

exists for any symmetric matrix  $M$ . It is clear then that this limit is linear as a function of  $M$ , and therefore, there is a unique symmetric matrix  $A_\ell$  such that

$$\text{tr}(A_\ell M) = \lim_{\eta \searrow 0} \langle \ell, \frac{1}{2}\eta(x)(Mx, x) \rangle. \quad (3.9)$$

Moreover, this matrix  $A_\ell$  is positive semi-definite and  $A_{\ell,\eta_\delta} \rightarrow A_\ell$  as  $\delta \searrow 0$ , and  $A_\ell = 0$  when  $\beta < 2$ . It remains to analyze the limit of  $B_{\ell,\phi_\delta}$  along a subsequence. For every  $\delta \in (0, 1)$

$$(B_{\phi_\delta})_i = \langle \ell, \phi_\delta x_i \rangle.$$

Now, recall the estimate from [Lemma 3.4](#), which implies

$$|\langle \ell, \phi_\delta x_i \rangle| \leq C(\|\phi_\delta x_i\|_{C^\beta(B_{1/2})} + \|\phi_\delta x_i\|_{L^\infty(\mathbb{R}^d)}).$$

A direct computation shows that

$$\sup_{0 < \delta < 1} \|\phi_\delta x_i\|_{C^\beta(B_{1/2})} < \infty.$$

It follows that

$$\sup_{0 < \delta < 1} |B_{\phi_\delta}| < \infty,$$

and by compactness, there must be a subsequence  $\delta_k \rightarrow 0$  for which  $\{B_{\ell,\phi_{\delta_k}}\}_k$  converges.  $\square$

**Lemma 3.9.** *Assume that  $\beta \in [0, 3)$ . Let  $\ell : C_b^\beta(\mathbb{R}^d) \rightarrow \mathbb{R}$  be a bounded linear functional which has the GCP with respect to 0. For  $\beta \geq 2$  and any  $u \in C_b^\beta(\mathbb{R}^d) \cap C^2(0)$ , we have the representation*

$$\langle \ell, u \rangle = C_\ell u(0) + (B_\ell, \nabla u(0)) + \text{tr}(A_\ell D^2 u(0)) + \int_{\mathbb{R}^d} u(y) - u(0) - \chi_{B_1(0)}(\nabla u(0), y) \mu_\ell(dy).$$

This representation is unique. This means that if there were  $\tilde{C}$ ,  $\tilde{B}$ ,  $\tilde{A}$  and  $\tilde{\mu}$  a measure in  $\mathbb{R}^d \setminus \{0\}$  all such that

$$\langle \ell, u \rangle = \tilde{C} u(0) + (\tilde{B}, \nabla u(0)) + \text{tr}(\tilde{A} D^2 u(0)) + \int_{\mathbb{R}^d} u(y) - u(0) - \chi_{B_1(0)}(\nabla u(0), y) \tilde{\mu}(dy).$$

for all  $u$ , then  $\tilde{C} = C_\ell$ ,  $\tilde{B} = B_\ell$ ,  $\tilde{A} = A_\ell$ , and  $\tilde{\mu} = \mu_\ell$ . Furthermore, if  $\beta < 2$  and  $u \in C^\beta(\mathbb{R}^d) \cap C^1(0)$ , then  $A_\ell = 0$ , and if  $\beta < 1$ , then  $B_\ell = 0$  and the integrand on the right can be replaced with just  $u(y) - u(0)$ .

**Proof.** Let  $\delta, \delta' \in (0, 1)$ . Applying [Lemma 3.6](#) with the functions  $\phi_\delta$  and  $\eta_{\delta'}$ ,

$$\langle \ell, u \rangle = C_\ell u(0) + (B_{\ell,\phi_\delta}, \nabla u(0)) + \text{tr}(A_{\ell,\eta_{\delta'}}, D^2 u(0)) + \int_{\mathbb{R}^d} u(y) - P_{\phi_\delta, \eta_{\delta'}, u}(y) \mu_\ell(dy).$$

Since  $\min\{1, |y|^\beta\}$  is integrable against  $\mu_\ell$ , it follows that

$$\lim_{\delta' \searrow 0} \int_{\mathbb{R}^d \setminus \{0\}} \eta_{\delta'}(y) (D^2 u(0) y, y) \mu_\ell(dy) = 0.$$

Therefore,

$$\lim_{\delta' \searrow 0} \int_{\mathbb{R}^d \setminus \{0\}} u(y) - P_{\phi_\delta, \eta_{\delta'}, u}(y) \mu_\ell(dy) = \int_{\mathbb{R}^d \setminus \{0\}} u(y) - u(0) - \phi_\delta(y) (\nabla u(0), y) \mu_\ell(dy).$$

Then, thanks to [Lemma 3.8](#), the formula for  $\langle \ell, u \rangle$  becomes (for every fixed  $\delta \in (0, 1)$ )

$$\langle \ell, u \rangle = C_\ell u(0) + (B_{\ell,\phi_\delta}, \nabla u(0)) + \text{tr}(A_\ell D^2 u(0)) + \int_{\mathbb{R}^d \setminus \{0\}} u(y) - u(0) - \phi(y) (\nabla u(0), y) \mu_\ell(dy).$$

Now, let  $\delta_k \searrow 0$  be chosen so that  $B_{\ell,\phi_{\delta_k}} \rightarrow B_\ell$  (which can be done thanks to [Lemma 3.8](#)). From the definition of  $\phi_\delta$ , we have that

$$u(y) - u(0) - \phi_{\delta_k}(y) (\nabla u(0), y) \text{ is monotone in } k.$$

At the same time, for every  $y \in \mathbb{R}^d$  we have

$$\lim_{k \rightarrow \infty} \phi_{\delta_k}(y) = \chi_{B_1(0)}.$$

Therefore, by monotone convergence we have

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^d \setminus \{0\}} u(y) - u(0) - \phi_{\delta_k}(y)(\nabla u(0), y) \mu_\ell(dy) = \int_{\mathbb{R}^d \setminus \{0\}} u - u(0) - \chi_{B_1(0)}(y)(\nabla u(0), y) \mu_\ell(dy).$$

From where it follows that

$$\langle \ell, u \rangle = C_\ell u(0) + (B_\ell, \nabla u(0)) + \text{tr}(A_\ell D^2 u(0)) + \int_{\mathbb{R}^d \setminus \{0\}} u(y) - u(0) - \chi_{B_1(0)}(y)(\nabla u(0), y) \mu_\ell(dy),$$

as claimed. It remains to prove the uniqueness part. For this, it is enough to show that if for all  $u$  we have  $\langle \ell, u \rangle = 0$  and

$$\langle \ell, u \rangle = C_\ell u(0) + (B_\ell, \nabla u(0)) + \text{tr}(A_\ell D^2 u(0)) + \int_{\mathbb{R}^d} u(y) - u(0) - \chi_{B_1(0)}(y)(\nabla u(0), y) \mu_\ell(dy),$$

then  $C_\ell = 0, B_\ell = 0, A_\ell = 0$  and  $\mu_\ell = 0$ . First, consider any  $u$  with compact support which is disjoint from  $\{0\}$ , for such a  $u$  we have

$$\langle \ell, u \rangle = \int_{\mathbb{R}^d} u(y) \mu_\ell(dy),$$

Since  $u$  can be any function with compact support in  $\mathbb{R}^d \setminus \{0\}$ , it follows that  $\mu_\ell = 0$ . Evaluating  $\ell$  at the function  $u(x) \equiv 1$  we obtain  $C_\ell = 0$ . Lastly, evaluating  $\ell$  at all of the functions of the form  $(x, e)$ ,  $e \in \mathbb{R}^d$  and  $(Mx, x)$ ,  $M$  symmetric matrix, we see that  $B_\ell \cdot e = 0$  for any vector  $e$  and  $\text{tr}(AM) = 0$  for any symmetric matrix  $M$ , so that  $B_\ell = 0$  and  $A_\ell = 0$ .  $\square$

By a simple change of variables, [Lemma 3.9](#) implies the following.

**Corollary 3.10.** *Assume that  $x$  is fixed,  $\beta \in [0, 3)$ , and let  $\ell : C_b^\beta(\mathbb{R}^d) \rightarrow \mathbb{R}$  be a bounded linear functional which has the GCP with respect to  $x$ . For  $\beta \geq 2$  any  $u \in C_b^\beta(\mathbb{R}^d) \cap C^2(x)$  we have the representation*

$$\langle \ell, u \rangle = C_\ell u(x) + (B_\ell, \nabla u(x)) + \text{tr}(A_\ell D^2 u(x)) + \int_{\mathbb{R}^d} u(x+y) - u(x) - \chi_{B_1(0)}(\nabla u(x), y) \mu_\ell(dy).$$

As before, this representation is unique, and when  $\beta < 2$  and  $u \in C_b^\beta(\mathbb{R}^d) \cap C^1(x)$ , we have  $A_\ell = 0$ , while for  $\beta < 1$  we have  $B_\ell = 0$  and the integrand can be replaced with just  $u(x+y) - u(x)$ .

### 3.1. Proofs of [Theorems 1.9](#) and [1.10](#)

With [Lemmas 3.4](#) and [3.9](#) and [Corollary 3.10](#) in hand, we can now prove [Theorems 1.10](#) and [1.9](#).

**Proof of Theorem 1.10.** Consider the functional,

$$F(u) := I(u, 0).$$

Now, by [Theorem 2.6](#), we have that

$$F(u) = \min_a \max_b \{f_{ab} + \langle \ell_{ab}, u \rangle\}.$$

By [Lemma 3.4](#), each  $\ell_{ab}$  is a linear operator having the GCP with respect to 0, in which case [Lemma 3.9](#) says that for  $u \in C_b^\beta(\mathbb{R}^d) \cap C^2(0)$ ,

$$\langle \ell_{ab}, u \rangle = \text{tr}(A_{ab} D^2 u(0)) + B_{ab} \cdot \nabla u(0) + C_{ab} u(0) + \int_{\mathbb{R}^d} u(y) - u(0) - \chi_{B_1}(0)(\nabla u(0), y) \mu_{ab}(dy).$$

The translation invariance of  $I$  boils down to the identity

$$I(u, x) = F(\tau_x u).$$

Therefore,

$$I(u, x) = \min_a \max_b \{f_{ab} + \langle \ell_{ab}, \tau_x u \rangle\}$$

However,  $\langle \ell_{ab}, \tau_x u \rangle$  has a simple expression, namely

$$\text{tr}(A_{ab} D^2 u(x)) + B_{ab} \cdot \nabla u(x) + C_{ab} u(x) + \int_{\mathbb{R}^d} u(x+y) - u(x) - \chi_{B_1}(0) \nabla u(x) \cdot y \mu_{ab}(dy),$$

and this proves the theorem.  $\square$

**Proof of Theorem 1.9.** The beginning of the proof is similar to that of the previous one. For each  $x \in \mathbb{R}^d$ , define a functional

$$F_x(u) := I(u, x), \quad \forall u \in C_b^\beta(\mathbb{R}^d).$$

Applying [Theorem 2.6](#), it follows that

$$F_x(u) := \min_{v \in C_b^\beta(\mathbb{R}^d)} \max_{\ell \in \partial F_x} \{F_x(v) + \langle \ell, u - v \rangle\}.$$

Applying [Lemma 3.4](#), it follows that for any  $\ell \in \partial F_x$

$$\langle \ell, u \rangle = C u(x) + (B, \nabla u(x)) + \text{tr}(A D^2 u(x)) + \int_{\mathbb{R}^d} u(x+y) - u(x) - \chi_{B_1}(0)(\nabla u(x), y) \mu(dy).$$

Since  $F_x(v) = I(v, x)$  this proves the Theorem, with  $\mathcal{K}(I)_x = \{L \mid L(u) = \langle \ell, u \rangle \text{ for } \ell \in \partial F_x\}$ .  $\square$

**Remark 3.11.** It is worthwhile to compare the proof of [Theorem 1.9](#) to the much longer and complicated one given in [\[28\]](#). The simplicity here is made possible by the use of a mean value theorem for Lipschitz functionals ([Theorem 2.5](#)) in the infinite dimensional setting, which suffices to prove [Theorem 1.9](#) as it involves a min–max formula in terms of linear functionals in  $C_b^2$  and not linear operators from  $C_b^2(\mathbb{R}^d)$  to  $C_b^0(\mathbb{R}^d)$ . The more complicated method from [\[28\]](#) is however still of value, specially if one is interested in obtaining a min–max representation in terms of a family of linear operators from  $C_b^2$  to  $C_b^0$ . Moreover, it is by adapting the method from [\[28\]](#) that we are able to prove [Theorem 1.11](#), after analyzing the spatial properties of the finite dimensional approximations (see in Section 5).

## 4. Finite dimensional approximations to $C_b^\beta(\mathbb{R}^d)$

### 4.1. Graph approximations

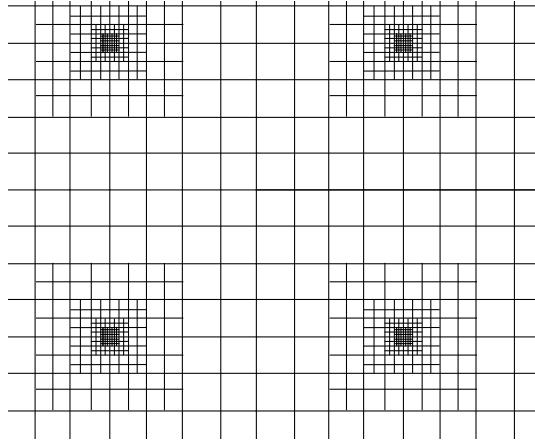
The following nested family of sets will be important in what follows

$$G_n := 2^{-n} \mathbb{Z}^d.$$

It will be convenient to write  $h_n := 2^{-n}$ . Then,  $h_n$  represents the maximum possible distance between  $x \in \mathbb{R}^d$  and  $G_n$ , and in particular  $\text{dist}(x, G_n) \leq h_n$  for all  $x \in \mathbb{R}^d$ . Observe that

$$G_1 \subset G_2 \subset G_3 \dots,$$

and note also the union of the sets  $G_n$  is dense in  $\mathbb{R}^d$ .



**Fig. 1.** A (periodic) cube decomposition of  $\mathbb{R}^d \setminus \mathbb{Z}^d$ .

**Definition 4.1.** We consider the following function spaces

$$\begin{aligned} C(G_n) &:= \{u : G_n \rightarrow \mathbb{R}^d\}, \\ C_*(G_n) &:= \{u \in C(G_n) \mid u(x) = 0 \text{ if } x \notin [-2^n, 2^n]^d\}. \end{aligned}$$

These spaces will be related to  $C_b^\beta(\mathbb{R}^d)$  by restriction, which we think of as a map denoted by  $T_n$  and given by

$$T_n : C_b^\beta(\mathbb{R}^d) \rightarrow C(G_n), \quad T_n u := u|_{G_n}.$$

**Remark 4.2.** The space  $C_*(G_n)$  is a finite dimensional vector space.

#### 4.2. Cube decomposition and partition of unity

In this section we shall apply the Whitney theory to extend functions in a grid  $r\mathbb{Z}^d$  to all of  $\mathbb{R}^d$ . Since it is in our interest for the Whitney construction to be compatible with the grid structure, we shall do the usual cube decomposition making sure the resulting family of cubes is invariant under translations by vectors in  $r\mathbb{Z}^d$ , the resulting construction is illustrated in Fig. 1.

**Lemma 4.3.** For every  $r > 0$ , there exists a collection of cubes  $\{Q_k\}_k$  such that

- (1) The cubes  $\{Q_k\}_k$  have pairwise disjoint interiors.
- (2) The cubes  $\{Q_k\}_k$  cover  $\mathbb{R}^d \setminus r\mathbb{Z}^d$ .
- (3)  $c_1 \text{diam}(Q_k) \leq \text{dist}(Q_k, \mathbb{Z}^d) \leq c_2 \text{diam}(Q_k)$ .
- (4) For every  $h \in r\mathbb{Z}^d$ , there is a bijection  $\sigma_h : \mathbb{N} \rightarrow \mathbb{N}$  such that  $Q_k + h = Q_{\sigma_h k}$  for every  $k \in \mathbb{N}$ .

**Proof.** We consider the case  $r = 1$ , once the collection of cubes is  $\{Q_k\}_k$  obtained in this case, the general case follows via scaling by taking the family  $\{rQ_k\}_k$ .

Consider the cube  $Q_0 = [-1/2, 1/2]^d$ , let  $\mathcal{M}_0$  denote the family of  $2^d$  equal size cubes obtained from  $Q_0$  by bisecting each of its sides. Let  $\mathcal{M}_k$  denote the family of cubes obtained from applying this same procedure to each of the cubes in  $\mathcal{M}_{k-1}$ . Note that the side length of each cube in  $\mathcal{M}_k$  is just  $2^{-k}$ . Now, we construct a family  $\mathcal{F}_0$  as follows, with  $R_k := \{2\sqrt{d}2^{-k} \leq |x| \leq 2\sqrt{d}2^{-(k-1)}\}$  for each  $k \in \mathbb{N}$ , then

$$\mathcal{F}_0 := \bigcup_k \{Q \in \mathcal{M}_k : Q \cap R_k \neq \emptyset\}.$$

Observe that if  $Q \in \mathcal{F}_0$  then  $Q \in \mathcal{M}_k$  for some  $k$  and there is some  $x \in Q$  such that  $2\sqrt{d}2^{-k} \leq |x|$  and  $|x| \leq 2\sqrt{d}2^{-(k-1)}$ . This means,

$$\sqrt{d}2^{-k} = 2\sqrt{d}2^{-k} - \text{diam}(Q) \leq \text{dist}(Q, 0) \leq 2\sqrt{d}2^{-k},$$

and since  $\text{diam}(Q) = \sqrt{d}2^{-k}$ , we conclude that

$$\text{diam}(Q) \leq \text{dist}(Q, 0) \leq 4\text{diam}(Q) \quad \forall Q \in \mathcal{F}_0.$$

On the other hand, we have that

$$\bigcup_{Q \in \mathcal{F}_0} Q = [-1/2, 1/2]^d \setminus \{0\}.$$

If  $\mathcal{F}$  denotes the subfamily of maximal cubes in  $\mathcal{F}_0$ , it follows that: the union of these cubes is still  $[-1/2, 1/2]^d \setminus \{0\}$ , the inequality  $\text{diam}(Q) \leq \text{dist}(Q, 0) \leq 4\text{diam}(Q)$  holds for each  $Q \in \mathcal{F}$ , and the cubes have pairwise disjoint interiors.

Denote by  $\{Q_k\}_k$  an enumeration of the family of cubes of the form  $Q + z$ , where  $Q \in \mathcal{F}$  and  $z \in \mathbb{Z}^d$ . It is clear that  $\{Q_k\}_k$  covers all of  $\mathbb{R}^d \setminus \mathbb{Z}^d$  and that these cubes have pairwise disjoint interiors. Furthermore, for any  $h \in \mathbb{Z}^d$  the map  $Q \rightarrow Q + h$  gives a bijection of the set  $\{Q_k\}_k$  onto itself, therefore one can represent it via a bijection  $\sigma_h : \mathbb{N} \rightarrow \mathbb{N}$  so that  $Q_k + h = Q_{\sigma_h k}$ . Last but not least, as each cube of the form  $Q + z$  is closest to  $z$  than to any other point in  $\mathbb{Z}^d$ , property (3) follows from the respectively inequality for the family  $\mathcal{F}$ .  $\square$

**Remark 4.4.** We apply [Lemma 4.3](#) with  $r = 2^{-n}$ , for some  $n \in \mathbb{N}$ , and for the rest of the section shall refer to the resulting cubes as  $\{Q_{n,k}\}_k$ .

Furthermore, for every  $n$  and  $k$ , we will denote the center of  $Q_{n,k}$  by  $y_{n,k}$ , and for each  $n$  and  $k$  we will denote by  $\hat{y}_{n,k}$  the unique point in  $G_n$  such that

$$\text{dist}(y_{n,k}, G_n) = |y_{n,k} - \hat{y}_{n,k}|,$$

(note that there is only one since by construction not a single center  $y_{n,k}$  lies at equidistance to two different lattice points).

In particular, for each of the bijections  $\sigma_h : \mathbb{N} \rightarrow \mathbb{N}$  from [Lemma 4.3](#) we have

$$y_{n,k} + h = y_{n,\sigma_h k}, \quad \hat{y}_{n,k} + h = \hat{y}_{n,\sigma_h k}, \quad \forall n, k.$$

**Remark 4.5.** In all what follows, given a cube  $Q$ , we shall denote by  $Q^*$  the cube with same center as  $Q$  but whose sides are increased by a factor of  $9/8$ . Observe that for every  $n$  and  $k$ , we have  $Q_{n,k}^* \subset \mathbb{R}^d \setminus 2^{2-n}\mathbb{Z}^d$ , and that any given  $x$  lies in at most some number  $C(d)$  of the cubes  $Q_k^*$ .

**Proposition 4.6.** For every  $n$ , there is a family of functions  $\phi_{n,k}(x)$  such that

- (1)  $0 \leq \phi_{n,k}(x) \leq 1$  for every  $k$  and  $\phi_{n,k} \equiv 0$  outside  $Q_{n,k}^*$  (using the notation in [Remark 4.5](#))
- (2)  $\sum_k \phi_{n,k}(x) = 1$  for every  $x \in \mathbb{R}^d \setminus G_n$ .
- (3) There is a constant  $C$ , independent of  $n$  and  $k$ , such that

$$|\nabla^i \phi_{n,k}(x)| \leq \frac{C}{\text{diam}(Q_{n,k})^i}.$$

- (4) For every  $z \in G_n$ , we have

$$\phi_{n,k}(x - z) = \phi_{n,\sigma_z k}(x), \quad \forall k, x,$$

where  $\sigma_z$  are the bijections introduced above.

**Proof.** Fix a  $C^\infty$  function  $\phi$  such that

$$\begin{aligned} 0 &\leq \phi \leq 1, \\ \phi &\equiv 1 \text{ in } Q_0 = [-1/2, 1/2]^d, \\ \phi &\equiv 0 \text{ outside } Q_0^*. \end{aligned}$$

Let  $\ell(Q)$  denote the common length for the sides of  $Q_{n,k}$ , and with  $y_{n,k}$  as given in [Remark 4.4](#) we define

$$\tilde{\phi}_{n,k} := \phi \left( \frac{x - y_{n,k}}{\ell_{n,k}} \right).$$

Consider the function

$$\Phi(x) = \sum_k \tilde{\phi}_{n,k}(x).$$

It follows from [Remark 4.5](#) that given any  $x$ , at most  $C(d)$  of the terms appearing in the sum are non-zero in a neighborhood of  $x$ , and therefore  $\Phi$  is a smooth function. Then, define

$$\phi_{n,k}(x) := \tilde{\phi}_{n,k}(x) \Phi(x)^{-1}.$$

It is clear that the functions  $\{\phi_{n,k}\}_k$  satisfy properties (1) and (2). Property (3) follows easily from the chain rule, using the differentiability of the function  $\phi$ . It remains to check property (4), let  $z \in G_n$ , then

$$\begin{aligned} \phi_{n,k}(x - z) &= \phi \left( \frac{x - (y_{n,k} + z)}{\ell(Q_{n,k})} \right) \Phi(x - z)^{-1} \\ &= \phi \left( \frac{x - y_{n,\sigma_z k}}{\ell(Q_{n,\sigma_z k})} \right) \Phi(x)^{-1} = \phi_{n,\sigma_z k}(x), \end{aligned}$$

where we used that  $\ell(Q_{n,k}) = \ell(Q_{n,\sigma_z k})$ , which follows clearly from the definition of  $\sigma_z$ .  $\square$

#### 4.3. Discrete derivatives

In what follows, it will be in our interest to approximate the first and second derivatives of a function  $u \in C_b^\beta(\mathbb{R}^d)$  (see [\(1.6\)](#) for our convention regarding the meaning of  $C_b^\beta$ ) at a point  $x \in G_n$  using only information about the values of  $u$  on  $G_n$ . This motivates the following two definitions (we recall that  $h_n = 2^{-n}$ ).

**Definition 4.7.** The vector  $(\nabla_n)^1 u(x)$  is defined via the system of equations ( $k = 1, \dots, d$ )

$$((\nabla_n)^1 u(x), e_k) := (2h_n)^{-1} [u(x + h_n e_k) - u(x - h_n e_k)]$$

**Definition 4.8.** The matrix  $(\nabla_n)^2 u(x)$  is defined via the system of equations ( $k, \ell = 1, \dots, d$ ),

$$((\nabla_n)^2 u(x) e_k, e_\ell) := h_n^{-2} [u(x + h_n e_k + h_n e_\ell) - u(x + h_n e_k) - u(x + h_n e_\ell) + u(x)]$$

**Remark 4.9.** From the definition it is clear that these discrete derivatives commute with translations with respect to a vector  $z \in G_n$ . That is, given a function  $u$  and  $z \in G_n$  then for every  $x \in G_n$  we have

$$((\nabla_n)^1 \tau_z u)(x) = ((\nabla_n)^1 u)(x + z)$$

Depending on how regular the function  $u$  is, these discrete derivative operators enjoy quantitative “continuity estimates” as functions on  $G_n$ . An important point being that these estimates are uniform in  $n$  once  $u$  is fixed.

**Proposition 4.10.** *There is a universal constant  $C$  such that for  $u \in C_b^\beta(\mathbb{R}^d)$  and  $x \in G_n$ ,*

$$\begin{aligned} |(\nabla_n)^1 u(x) - \nabla u(x)| &\leq C \|u\|_{C^\beta} h_n^{\beta-1}, \quad \text{if } \beta \in [1, 2], \\ |(\nabla_n)^2 u(x) - D^2 u(x)| &\leq C \|u\|_{C^\beta} h_n^{\beta-2}, \quad \text{if } \beta \in [2, 3]. \end{aligned}$$

**Proof.** See the [Appendix](#).  $\square$

**Proposition 4.11.** *Fix  $u \in C_b^\beta(\mathbb{R}^d)$ . Then, given  $x_1, x_2 \in G_n$ , we have*

$$\begin{aligned} |u(x_1) - u(x_2)| &\leq C \|u\|_{C^\beta} |x_1 - x_2|^\beta, \quad \text{if } \beta \in [0, 1], \\ |(\nabla_n)^1 u(x_1) - (\nabla_n)^1 u(x_2)| &\leq C \|u\|_{C^\beta} |x_1 - x_2|^{\beta-1}, \quad \text{if } \beta \in [1, 2], \\ |(\nabla_n)^2 u(x_1) - (\nabla_n)^2 u(x_2)| &\leq C \|u\|_{C^\beta} |x_1 - x_2|^{\beta-2}, \quad \text{if } \beta \in [2, 3]. \end{aligned}$$

**Proof.** See the [Appendix](#).  $\square$

#### 4.4. The Whitney extension and projection operators.

**Definition 4.12.**

$$p_{u,k}^\beta(x) := \begin{cases} u(\hat{y}_{n,k}) & \text{if } \beta \in [0, 1) \\ u(\hat{y}_{n,k}) + (\nabla_n^1 u(\hat{y}_{n,k}), x - \hat{y}_{n,k}) & \text{if } \beta \in [1, 2) \\ u(\hat{y}_{n,k}) + (\nabla_n^1 u(\hat{y}_{n,k}), x - \hat{y}_{n,k}) + \frac{1}{2} (\nabla_n^2 u(\hat{y}_{n,k})(x - \hat{y}_{n,k}), (x - \hat{y}_{n,k})) & \text{if } \beta \in [2, 3) \end{cases}$$

We are now ready to define the Whitney extension operator.

$$E_n^\beta(u, x) := \begin{cases} u(x) & \text{if } x \in G_n, \\ \sum_k p_{u,k}^\beta(x) \phi_{n,k}(x) & \text{if } x \notin G_n. \end{cases} \quad (4.1)$$

The projector operator  $\pi_n^\beta : C_b^\beta(\mathbb{R}^d) \rightarrow C_b^\beta(\mathbb{R}^d)$  is given by

$$\pi_n^\beta := E_n^\beta \circ T_n, \quad (4.2)$$

where we recall that  $T_n u = u|_{G_n}$  ([Definition 4.1](#)).

**Theorem 4.13.** *There is a constant  $C$  such that for any  $n$  and any  $u \in C_b^\beta(\mathbb{R}^d)$  we have*

$$\|\pi_n^\beta u\|_{C^\beta(\mathbb{R}^d)} \leq C \|u\|_{C^\beta(\mathbb{R}^d)}.$$

**Proof.** This follows arguing exactly as in [54, Chapter VI, Theorem 3 and 4], making use of the regularity estimates in [Proposition 4.11](#). Since this is a standard argument, we omit the details.  $\square$

**Proposition 4.14.** *Let  $z \in G_n$  and  $u \in C_b^\beta$ , then.*

$$\pi_n^\beta(\tau_z u) = \tau_z \pi_n^\beta(u).$$

**Proof.** Let us show that  $\pi_n^\beta(\tau_z u)(x) = \tau_z \pi_n^\beta(u)(x)$  for every  $x \in \mathbb{R}^d$  and  $z \in G_n$ . Note that if  $x \in G_n$  then the equality is trivial, so let us take  $x \in \mathbb{R}^d \setminus G_n$  and  $z \in G_n$ , then we have

$$\pi_n^\beta(\tau_z u)(x) = \sum_k p_{\tau_z u, k}^\beta(x) \phi_{n,k}(x).$$

Furthermore, it is not difficult to check that (see [Remark 4.9](#))

$$p_{\tau_z u, k}^\beta(x) = p_{u, \sigma_z k}^\beta(x + z),$$

while part (4) of [Proposition 4.6](#) implies that

$$\phi_{n, k}(x) = \phi_{n, \sigma_z k}(x + z).$$

From these two identities we conclude that

$$\pi_n^\beta(\tau_z u)(x) = \sum_k p_{u, \sigma_z k}^\beta(x + z) \phi_{n, \sigma_z k}(x + z) = \sum_k p_{u, k}^\beta(x + z) \phi_{n, k}(x + z) = \tau_z \pi_n^\beta(u)(x),$$

where we used that  $\sigma_z$  is bijective, this proves the proposition.  $\square$

**Remark 4.15.** Given  $\varepsilon \in (0, 1)$  there is a  $C > 1$  such that for every  $n \in \mathbb{N}$ ,  $x_0 \in G_n$ , and unit vector  $x_* \in \mathbb{R}^d$  there is some  $x_1 \in G_n$  and  $s > 0$  such that

$$|sx_* - (x_1 - x_0)| \leq h_n, \quad C^{-1}h_n^\varepsilon \leq |x_1 - x_0| \leq Ch_n^\varepsilon.$$

Indeed, this follows from the fact that  $h_n^\varepsilon x_* \in [-h_n^\varepsilon, h_n^\varepsilon]^d$  and that  $[-h_n^\varepsilon, h_n^\varepsilon]^d \cap (G_n - x_0)$  is a  $h_n$ -net in  $[-h_n^\varepsilon, h_n^\varepsilon]^d$ , so there is  $x_1 \in [-h_n^\varepsilon, h_n^\varepsilon]^d \cap (G_n - x_0)$  such that  $|h_n^\varepsilon x_* - (x_1 - x_0)| \leq h_n$ . Then, the inequalities for  $|x_1 - x_0|$  follow from two applications of the triangle inequality and the fact that  $\varepsilon < 1$  and  $h_n \leq 1/2$  for all  $n \geq 1$ .

**Proposition 4.16.** Let  $w \in C_b^\beta(\mathbb{R}^d)$  be such that  $w(x) \geq 0$  for every  $x \in G_n$  and such that  $w(x_0) = 0$  at some  $x_0 \in G_n$ . Then, there is a universal  $C$  such that

$$\begin{aligned} |\nabla \pi_n^\beta w(x_0)| &\leq C \|w\|_{C^\beta} h_n^{\min\{2, \beta\}-1}, \quad \text{if } \beta \geq 1, \\ |(\nabla^2 \pi_n^\beta w(x_0))_-| &\leq C \|w\|_{C^\beta} h_n^{(\min\{3, \beta\}-2)/2}, \quad \text{if } \beta \geq 2. \end{aligned}$$

Here, for a given symmetric matrix  $D$ ,  $D_-$  denotes its negative part.

**Proof.** Fix any  $x \in G_n$ . Thanks to [Proposition 4.10](#) and the fact that  $|x - x_0| \geq h_n$  we have

$$|w(x) - w(x_0) - (\nabla \pi_n^\beta w(x_0), x - x_0)| \leq C \|w\|_{C^\beta} |x - x_0|^{\min\{2, \beta\}}.$$

Since  $w(x_0) = 0$ , and  $w(x) \geq 0$  by assumption,

$$0 \leq (\nabla \pi_n^\beta w(x_0), x - x_0) + C \|w\|_{C^\beta} |x - x_0|^{\min\{2, \beta\}}.$$

It is easy to see there is some  $x_1 \in G_n$  such that  $|x_0 - x_1| = h_n$  and

$$(\nabla \pi_n^\beta w(x_0), x_1 - x_0) = -|\nabla \pi_n^\beta w(x_0)|_{\ell^\infty} |x_1 - x_0|,$$

and therefore,

$$(\nabla \pi_n^\beta w(x_0), x_1 - x_0) \leq -C_d^{-1} |\nabla \pi_n^\beta w(x_0)| |x_1 - x_0|.$$

Combining these inequalities and recalling [Theorem 4.13](#) it follows that

$$|\nabla \pi_n^\beta w(x_0)| \leq C \|w\|_{C^\beta} h_n^{\min\{2, \beta\}-1}.$$

This proves the estimate for the gradient when  $\beta \geq 1$ . Now assume  $\beta \geq 2$ , the beginning of the argument in this case goes along similar lines. For any  $x \in G_n$  we have that

$$|w(x) - w(x_0) - (\nabla \pi_n^\beta w(x_0), x - x_0) - \frac{1}{2}(\nabla^2 \pi_n^\beta w(x_0)(x - x_0), x - x_0)| \leq C\|w\|_{C^\beta} |x - x_0|^{\min\{3, \beta\}},$$

where we have once again used [Theorem 4.13](#). Thus, since  $w(x_0) = 0$  and  $w(x) \geq 0$  for  $x \in G_n$ ,

$$(\nabla \pi_n^\beta w(x_0), x - x_0) + \frac{1}{2}(\nabla^2 \pi_n^\beta w(x_0)(x - x_0), x - x_0) + C\|w\|_{C^\beta} |x - x_0|^{\min\{3, \beta\}} \geq 0.$$

Now, since we are on a lattice, it is obvious that for any  $x \in G_n$  we have that  $x' := 2x_0 - x \in G_n$ . In this case we can add up the inequalities for  $x$  and  $x'$ , and conclude that

$$\begin{aligned} & (\nabla \pi_n^\beta w(x_0), x - x_0) + \frac{1}{2}(\nabla^2 \pi_n^\beta w(x_0)(x - x_0), x - x_0) \\ & + (\nabla \pi_n^\beta w(x_0), x' - x_0) + \frac{1}{2}(\nabla^2 \pi_n^\beta w(x_0)(x' - x_0), x' - x_0) + 2C\|w\|_{C^\beta} |x - x_0|^{\min\{3, \beta\}} \geq 0. \end{aligned}$$

Since  $x' - x_0 = -(x - x_0)$ , we conclude that

$$(\nabla^2 \pi_n^\beta w(x_0)(x - x_0), x - x_0) + 2C\|w\|_{C^\beta} |x - x_0|^{\min\{3, \beta\}} \geq 0, \quad \forall x \in G_n.$$

Let  $x_* \in \mathbb{R}^d$  be a unit vector such that

$$-(\nabla^2 \pi_n^\beta w(x_0)x_*, x_*) = |(\nabla^2 \pi_n^\beta w(x_0))_-|$$

According to [Remark 4.15](#), there is  $x_1 \in G_n$  and  $s > 0$  such that

$$|sx_* - (x_1 - x_0)| \leq h_n, \quad C^{-1}h_n^\varepsilon \leq |x_1 - x_0| \leq Ch_n^\varepsilon.$$

For this  $x_1$  we have

$$\begin{aligned} |(\nabla^2 \pi_n^\beta w(x_0))_-|s^2 &= -(\nabla^2 \pi_n^\beta w(x_0)x_*, x_*)s^2 \\ &\leq -(\nabla^2 \pi_n^\beta w(x_0)(x_1 - x_0), x_1 - x_0) + C\|w\|_{C^\beta} |sx_* - (x_1 - x_0)|. \end{aligned}$$

This, together with the previous step, shows that

$$C^{-2}|(\nabla^2 \pi_n^\beta w(x_0))_-|(h_n^\varepsilon)^2 \leq 2C\|w\|_{C^\beta} h_n^{\min\{3, \beta\}\varepsilon} + C\|w\|_{C^\beta} h_n,$$

again having used [Theorem 4.13](#). Simplifying, this becomes

$$|(\nabla^2 \pi_n^\beta w(x_0))_-| \leq C\|w\|_{C^\beta} (h_n^{(\min\{3, \beta\}-2)\varepsilon} + h_n^{1-\varepsilon}).$$

Choosing  $\varepsilon = 1/2$ , and noting  $(\min\{3, \beta\} - 2) \leq 1$ , we conclude that

$$|(\nabla^2 \pi_n^\beta w(x_0))_-| \leq C\|w\|_{C^\beta} h_n^{(\min\{3, \beta\}-2)/2}. \quad \square$$

We fix an auxiliary function  $\eta_0 : [0, \infty) \rightarrow \mathbb{R}_+$ , with  $\eta_0 \in C^\infty(\mathbb{R}_+)$ , and

$$0 \leq \eta_0 \leq 1, \quad \eta_0'(t) \geq 0 \text{ for all } t, \quad \eta_0(t) = t \text{ for } t \leq 1/2, \quad \eta_0(t) = 1 \text{ for } t \geq 1. \quad (4.3)$$

The function  $\eta_0$ , as well as the following two estimates, will be useful in the next section. Essentially,  $\eta_0(t)$  should be thought of as a smooth replacement for  $\min\{1, t\}$ .

**Lemma 4.17.** Let  $1 \leq \beta < \beta_0 < 3$ , and consider  $w \in C_b^{\beta_0}(\mathbb{R}^d)$  and  $x_0 \in G_n$  such that

$$w \geq 0 \text{ in } G_n \text{ and } w(x_0) = 0.$$

Then, there is a function  $R_{\beta_0, n, w, x_0}$  such that  $R(x_0) = 0$ , and

$$\begin{aligned} \pi_n^\beta w(x) + R_{\beta_0, n, w, x_0}(x) &\geq 0, \quad \forall x \in \mathbb{R}^d, \\ \|R_{\beta_0, n, w, x_0}\|_{C^\beta(\mathbb{R}^d)} &\leq Ch_n^\gamma \|w\|_{C^{\beta_0}(\mathbb{R}^d)}, \end{aligned}$$

for some constant  $\gamma = \gamma(\beta, \beta_0) \in (0, 1)$ .

**Remark 4.18.** For  $\beta \in (0, 1)$ , it is straightforward that  $w \geq 0$  in  $G_n$  guarantees that  $\pi_n^\beta w \geq 0$  everywhere, that is, the Whitney extension for  $\beta \in (0, 1)$  is order preserving. Accordingly, Lemma 4.17 is only needed for  $\beta > 1$ .

**Proof.** We consider the cases  $1 \leq \beta < 2$  and  $\beta \geq 2$  separately. First suppose  $\beta \in [1, 2)$ . Let  $\phi_0(t)$  be a smooth function such that  $0 \leq \phi_0(t) \leq 1$  for all  $t$ ,  $\phi_0(t) = 1$  for  $t \leq 1/4$  and  $\phi_0(t) = 0$  for  $t \geq 1$ . Then set

$$\tilde{w}(x) = \pi_n^\beta w(x) - (\nabla \pi_n^\beta w(x_0), x - x_0) \phi_0(x - x_0).$$

For each  $x \in \mathbb{R}^d$ , let  $\hat{x}$  denote a point in  $G_n$  such that  $|x - \hat{x}| = \text{dist}(x, G_n) \leq h_n$ . Then, since  $w(\hat{x}) \geq 0$  for any  $\hat{x}$  (from the assumption), we have

$$\begin{aligned} \tilde{w}(x) &= \tilde{w}(\hat{x}) + (\tilde{w}(x) - \tilde{w}(\hat{x})) \\ &\geq -(\nabla \pi_n^\beta w(x_0), x - x_0) \phi_0(x - x_0) - C \|\tilde{w}\|_{C^{\beta_0}} |\hat{x} - x| \\ &\geq -(\nabla \pi_n^\beta w(x_0), x - x_0) \phi_0(x - x_0) - C \|\tilde{w}\|_{C^{\beta_0}} h_n. \end{aligned}$$

By Proposition 4.16, we have  $|\nabla \pi_n^\beta w(x_0)| \leq C \|w\|_{C^{\beta_0}} h_n$  when  $\beta_0 > 1$ , therefore,

$$\tilde{w}(x) \geq -C \|w\|_{C^{\beta_0}} h_n, \quad \forall x \in \mathbb{R}^d,$$

where we have used Theorem 4.13 to bound  $\|\pi_n^\beta w\|_{C_0^\beta}$ . On the other hand, since  $\beta_0 > 1$  and  $\nabla \tilde{w}(x_0) = 0$ , we have

$$\begin{aligned} \tilde{w}(x) &\geq -\|\tilde{w}\|_{C^{\beta_0}} |x - x_0|^{\beta_0}, \\ &\geq -C \|w\|_{C^{\beta_0}} |x - x_0|^{\beta_0} \quad \forall x \in \mathbb{R}^d, \end{aligned}$$

Now, we take  $\eta_0$  as in (4.3) and define the function

$$\tilde{R}(x) := 2C \|w\|_{C^{\beta_0}} h_n \eta_0 \left( \frac{|x - x_0|^{\beta_0}}{h_n} \right).$$

If  $|x - x_0|^{\beta_0} \geq h_n/2$ , then

$$\tilde{w}(x) + \tilde{R}(x) = \tilde{w}(x) + C \|w\|_{C^{\beta_0}} h_n \geq 0.$$

If on the contrary,  $|x - x_0|^{\beta_0} \leq h_n/2$ , then

$$\tilde{w}(x) + \tilde{R}(x) = \tilde{w}(x) + C \|w\|_{C^{\beta_0}} |x - x_0|^{\beta_0} \geq 0.$$

We conclude that

$$\tilde{w}(x) + \tilde{R}(x) \geq 0, \quad \forall x \in \mathbb{R}^d.$$

On the other hand, an elementary computation (see the [Appendix](#)) shows that

$$\|\tilde{R}\|_{C^\beta} \leq Ch_n^\gamma \|w\|_{C^{\beta_0}}.$$

Finally, let

$$R_{\beta_0, n, w, x_0}(x) := \tilde{R}(x) - (\nabla \pi_n^\beta(x_0), x - x_0) \phi_0(x - x_0).$$

We conclude that  $\|R_{\beta_0, n, w, x_0}\|_{C^\beta} \leq Ch_n^\gamma \|w\|_{C^{\beta_0}}$  and

$$\pi_n^\beta w(x) + R_{\beta_0, n, w, x_0}(x) \geq 0, \quad \forall x \in \mathbb{R}^d.$$

This proves the Proposition when  $\beta \in [1, 2)$ . The argument for  $\beta \geq 2$  is similar, we only highlight the main differences. This time, we subtract not just the first order part of  $w$  near  $x_0$ , but also the second order part, namely we consider the function

$$\tilde{w} := \pi_n^\beta w(x) - (\nabla \pi_n^\beta(x_0), x - x_0) \phi_0(x - x_0) - \frac{1}{2} ((\nabla^2 \pi_n^\beta(x_0))_-(x - x_0), x - x_0) \phi_0(x - x_0).$$

Then, one applies again [Proposition 4.16](#) and use the regularity of  $w$  to obtain (in analogy to the previous case)

$$\tilde{w}(x) \geq -C \|w\|_{C^{\beta_0}} \max\{h_n, |x - x_0|^{\beta_0}\}$$

The respective function  $\tilde{R}$  is defined exactly as  $\tilde{R}$  and one argues as in the previous case.  $\square$

**Remark 4.19.** The argument in the proof provides – after small modifications – a closely related result: if instead of  $w \in C_b^\beta(\mathbb{R}^d)$  we assume that  $w \in C_b^0(\mathbb{R}^d)$  and that for some  $M > 0$  and  $\beta_0 > \beta$  we have

$$|w(x)| \leq M|x - x_0|^{\beta_0}, \quad \forall x \in \mathbb{R}^d,$$

then there is as before a function  $\hat{R}_{\beta_0, n, w, x_0}$  such that  $\hat{R}_{\beta_0, n, w, x_0}(x_0) = 0$  and  $\pi_n^\beta w(x) + R_{\beta_0, n, w, x_0}(x) \geq 0$  for all  $x$ , but this time the  $C^\beta$  estimate for  $\hat{R}_{\beta_0, n, w, x_0}$  is

$$\|\hat{R}_{\beta_0, n, w, x_0}\|_{C^\beta} \leq Ch_n^\gamma (\|w\|_{L^\infty} + M).$$

The following proposition will be useful later in the proof of [Proposition 5.8](#).

**Proposition 4.20.** Let  $1 \leq \beta < \beta_0 < 3$  or  $\beta \in (0, 1)$  and  $\beta_0 = \beta$ . Fix  $f \in C_c^\infty(\mathbb{R}^d)$ , and let  $\eta_0$  be as in [\(4.3\)](#). Let  $x_0 \in G_n$  and  $w(x) = f(x - x_0) \eta_0(|x - x_0|^{\beta_0})$ , then

$$\begin{aligned} \pi_n^\beta(w, x) &\leq C \|f\|_{L^\infty} \eta_0(|x - x_0|^{\beta_0}), \quad \forall x \in \mathbb{R}^d, \text{ if } \beta \in (0, 1), \\ \pi_n^\beta(w, x) &\leq C \|f\|_{L^\infty} \eta_0(|x - x_0|^{\beta_0}) + \hat{R}_{\beta_0, n, w, x_0}(x), \quad \forall x \in \mathbb{R}^d, \text{ if } \beta \in [1, 2], \end{aligned}$$

for some function  $\hat{R}_{\beta_0, n, w, x_0}$  such that  $\hat{R}_{\beta_0, n, w, x_0}(x_0) = 0$  and

$$\|\hat{R}_{\beta_0, n, w, x_0}\|_{C^\beta} \leq C \|f\|_{L^\infty} h_n^\gamma,$$

where  $\gamma$  is as in [Lemma 4.17](#).

**Proof.** Define the function  $\tilde{w}(x) := (\|f\|_{L^\infty} - f(x - x_0)) \eta_0(|x - x_0|^{\beta_0})$ . Then  $\tilde{w}(x_0) = 0$  and

$$|\tilde{w}(x)| \leq 2 \|f\|_{L^\infty} \eta_0(|x - x_0|^{\beta_0}), \quad \forall x \in \mathbb{R}^d,$$

while, since  $\eta_0 \geq 0$ , we also have  $\tilde{w}(x) \geq 0$  for every  $x \in G_n$ . If  $\beta \in [1, 2]$ , using [Lemma 4.17](#) and the function  $\hat{R}_{\beta_0, n, w, x_0}$  from [Remark 4.19](#), we have

$$\pi_n^\beta(\tilde{w}, x) + \hat{R}_{\beta_0, n, w, x_0}(x) \geq 0, \quad \forall x,$$

This inequality, after some rearranging, yields (for  $\beta \in [1, 2]$ )

$$\pi_n^\beta(w, x) \leq \|f\|_{L^\infty} \pi_n^\beta(\eta_0(|\cdot - x_0|^{\beta_0}), x) + \hat{R}_{\beta_0, n, w, x_0}(x), \quad \forall x \in \mathbb{R}^d.$$

Since we also have  $\|\tilde{w}\|_{L^\infty} \leq C\|f\|_{L^\infty}$ , we have again by [Remark 4.19](#)

$$\|\hat{R}_{\beta_0, n, w, x_0}\|_{C^\beta} \leq C\|f\|_{L^\infty} h_n^\gamma,$$

and the Proposition is proved in this case. For  $\beta \in (0, 1)$  we argue along similar lines, using [Remark 4.18](#) instead of [Lemma 4.17](#).  $\square$

#### 4.5. Convergence of the projection operators

**Lemma 4.21.** *Let  $0 < \beta < \beta_0 < 3$ , there is a constant  $C$  such that if  $u \in C_b^{\beta_0}(\mathbb{R}^d)$ , then*

$$\|\pi_n^\beta u - u\|_{C^\beta} \leq C h_n^\gamma \|u\|_{C^{\beta_0}}.$$

Here,  $\gamma = \gamma(\beta_0, \beta) \in (0, 1)$ .

**Proof.** For notational simplicity let us write  $f(x) = \pi_n^\beta u(x)$  throughout the proof.

Since  $u = f$  throughout  $G_n$ , for an arbitrary  $x \in G_n$  we have (with  $\hat{x}$  denoting a point in  $G_n$  such that  $\text{dist}(x, G_n) = |x - \hat{x}|$ ), with  $\alpha := \min\{1, \beta_0\}$

$$\begin{aligned} |u(x) - f(x)| &\leq |f(x) - f(\hat{x})| + |u(\hat{x}) - u(x)| \\ &\leq |x - \hat{x}|^\alpha [f]_{C^\alpha} + |x - \hat{x}|^\alpha [u]_{C^\alpha} \\ &\leq C\|u\|_{C^{\beta_0}} h_n^\alpha \leq C\|u\|_{C^{\beta_0}} h_n^\alpha, \end{aligned}$$

where we made use of [Theorem 4.13](#) to obtain  $[f]_{C^\alpha} \leq C\|u\|_{C^\beta}$ . This shows that  $\|u - f\|_{L^\infty}$  goes to zero at some rate determined by  $\beta_0$  and the size of  $\|u\|_{C^{\beta_0}}$ . To prove the lemma we need to also bound the Hölder seminorm of  $u - f$  and its derivatives, according to  $\beta_0$ .

**The case  $\beta, \beta_0 \in [0, 1]$ .** Fix  $x_1, x_2 \in \mathbb{R}^d$ . First, suppose that  $|x_1 - x_2| \leq \max\{|x_1 - \hat{x}_1|, |x_2 - \hat{x}_2|\}$ , then

$$|f(x_1) - u(x_1) - (f(x_2) - u(x_2))| \leq [f - u]_{C^{\beta_0}} |x_1 - x_2|^{\beta_0} \leq C\|u\|_{C^{\beta_0}} |x_1 - x_2|^{\beta_0}.$$

In this case, and since  $0 \leq \beta < \beta_0 < 1$ , we have that  $|x_1 - x_2|^{\beta_0 - \beta} \leq \max\{|x_1 - \hat{x}_1|^{\beta_0 - \beta}, |x_2 - \hat{x}_2|^{\beta_0 - \beta}\} \leq h_n^{\beta_0 - \beta}$ . Then, using [Theorem 4.13](#)

$$|f(x_1) - u(x_1) - (f(x_2) - u(x_2))| \leq [f - u]_{C^\beta} |x_1 - x_2|^\beta \leq C\|u\|_{C^{\beta_0}} h_n^{\beta_0 - \beta} |x_1 - x_2|^\beta.$$

Next, suppose that  $|x_1 - x_2| > \max\{|x_1 - \hat{x}_1|, |x_2 - \hat{x}_2|\}$ . In this case

$$\begin{aligned} |f(x_1) - u(x_1) - (f(x_2) - u(x_2))| &\leq \|f\|_{C^{\beta_0}} |x_1 - \hat{x}_1|^{\beta_0} + \|u\|_{C^{\beta_0}} |x_2 - \hat{x}_2|^{\beta_0} \\ &\leq C\|u\|_{C^{\beta_0}} h_n^{\beta_0 - \beta} |x_1 - x_2|^\beta, \end{aligned}$$

where once again [Theorem 4.13](#) was used. Combining these two estimates, we conclude that

$$[f - u]_{C^\beta} = \sup_{x_1 \neq x_2} \frac{|f(x_1) - u(x_1) - (f(x_2) - u(x_2))|}{|x_1 - x_2|^\beta} \leq C\|u\|_{C^{\beta_0}} h_n^{\beta_0 - \beta}.$$

Then, using that  $h_n \leq 1$  for all  $n \geq 1$ , we have

$$\|f - u\|_{C^\beta} \leq C h_n^\gamma \|u\|_{C^{\beta_0}}.$$

**The case  $\beta, \beta_0 \in [1, 2]$ .** In this case we trivially have the same estimates from the previous case, and only need the bounds for first derivative. This is done as follows, first

$$|\nabla f(x) - \nabla u(x)| \leq |\nabla f(x) - \nabla f(\hat{x})| + |\nabla f(\hat{x}) - \nabla u(\hat{x})| + |\nabla u(x) - \nabla u(\hat{x})|.$$

Then, using [Theorem 4.13](#), we have

$$\begin{aligned} |\nabla f(x) - \nabla u(x)| &\leq [\nabla f]_{C^{\beta_0-1}} h_n^{\beta_0-1} + |\nabla f(\hat{x}) - \nabla u(\hat{x})| + [\nabla u]_{C^{\beta_0-1}} h_n^{\beta_0-1} \\ &\leq C \|u\|_{C^{\beta_0}} h_n^{\beta_0-1} + |\nabla f(\hat{x}) - \nabla u(\hat{x})|. \end{aligned}$$

Recall that  $\nabla f(\hat{x}) = (\nabla_n)^1 u(\hat{x})$ , and use [Proposition 4.10](#) to conclude that

$$|\nabla f(x) - \nabla u(x)| \leq C \|u\|_{C^{\beta_0}} h_n^{\beta_0-1} + C \|u\|_{C^{\beta_0}} h_n^{\beta_0-1}.$$

The Hölder seminorm  $[\nabla f - \nabla u]_{C^\beta}$  is bounded with the same argument used to bound  $[f - u]_{C^\beta}$  in the previous case, we omit the details.

**The case  $\beta = 2, \beta_0 \in (2, 3)$ .** Right as before, we note that

$$|D^2 f(x) - D^2 u(x)| \leq |D^2 f(x) - D^2 f(\hat{x})| + |D^2 f(\hat{x}) - D^2 u(\hat{x})| + |D^2 u(x) - D^2 u(\hat{x})|.$$

Then, applying [Theorem 4.13](#) and [Proposition 4.10](#) as in the previous case, we have

$$\begin{aligned} |D^2 f(x) - D^2 u(x)| &\leq [D^2 f]_{C^{\beta_0-2}} h_n^{\beta_0-2} + |D^2 f(\hat{x}) - D^2 u(\hat{x})| + [D^2 u]_{C^{\beta_0-2}} h_n^{\beta_0-2} \\ &\leq 2C \|u\|_{C^{\beta_0}} h_n^{\beta_0-2} + |\nabla f(\hat{x}) - \nabla u(\hat{x})| \\ &\leq 3C \|u\|_{C^{\beta_0}} h_n^{\beta_0-2}. \end{aligned}$$

For the Hölder seminorm, we repeat the argument used in the case  $\beta \in (0, 1)$ , again we leave the details to the reader.  $\square$

**Remark 4.22.** If  $u \in C_b^0(\mathbb{R}^d)$ , then the same argument from [Lemma 4.21](#) can be used to show

$$\lim_{n \rightarrow \infty} \|u - \pi_n^0(u)\|_{L^\infty(\mathbb{R}^d)} = 0,$$

the rate of convergence being determined by the modulus of continuity of  $u$ .

## 5. Analysis of $I(u, x)$ via the finite dimensional approximations

In this section we introduce a sequence of operators  $I_n$  which approximate  $I$ . The operators  $I_n$  behave like operators in a finite dimensional vector space in the sense that they arise from a composition between linear maps with a Lipschitz map from a finite dimensional space onto itself. This allows us to prove a min–max formula for  $I_n(u, x)$  at least when  $x \in G_n$  by using Clarke’s idea of a generalized gradient [18]. More precisely, we use the fact that  $I_n$  factorizes via a map between finite dimensional vector spaces (which is what the spaces  $C_*(G_n)$  were introduced for), where the generalized gradient can be used, and then lift this to corresponding maps from  $C_b^\beta(\mathbb{R}^d)$  to  $C_b^0(\mathbb{R}^d)$  using the Whitney extension. The majority of the section is concerned with deriving estimates and regularity properties for the linear operators arising in the min–max formula for  $I_n$ , and ultimately concluding such linear operators are pre-compact, which leads to a min–max formula for the original operator.

### 5.1. The operators $I_n$ and their min–max representation

We are going to approximate the operator  $I(\cdot, x)$  via “finite dimensional approximations”, this referring to maps  $I_n : C_b^\beta \rightarrow C_b^0$ , which factorize through a finite dimensional space (see (5.3)).

We introduce a modification of the projection operator  $\pi_n^0$  defined in (4.2). First, we define

$$\text{Pr}_n : C(G_n) \rightarrow C_*(G_n), \quad \text{Pr}_n(u)(x) := u(x)\chi_{[-2^n, 2^n]^d}(x).$$

That is, given  $u \in C(G_n)$ , we define  $\text{Pr}_n(u)$  as the function obtained by restricting  $u$  to  $G_n \cap [-2^n, 2^n]^d$  and then extending it to the rest of  $G_n$  by zero. Then, we define the modified Whitney extension,

$$\hat{E}_n^\beta := E_n^\beta \circ \text{Pr}_n,$$

and the modified projection operator

$$\hat{\pi}_n^\beta := \hat{E}_n^\beta \circ T_n.$$

These are, respectively, bounded linear maps from  $C(G_n)$  to  $C_b^\beta(\mathbb{R}^d)$  and from  $C_b^0(\mathbb{R}^d)$  to  $C_b^\beta(\mathbb{R}^d)$ . Now we are ready to introduce the finite dimensional approximations to the operator  $I$ , define

$$I_n = \hat{\pi}_n^\beta \circ I \circ \hat{\pi}_n^\beta, \quad I_n : C_b^\beta(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d). \quad (5.1)$$

That is, to compute  $I_n(u, x)$ , we first compute the modified projection  $\hat{\pi}_n^\beta u$ , and compute  $I(\hat{\pi}_n^\beta u)$ , to which we later apply the modified projection  $\hat{\pi}_n^\beta$ . In particular,  $I_n$  only depends on the values of  $u$  on  $G_n \cap [-2^n, 2^n]^d$ . Associated to this, we introduce a map,  $i_n$ , defined as follows

$$i_n : C_*(G_n) \rightarrow C_*(G_n), \quad i_n = \text{Pr}_n \circ T_n \circ I \circ E_n^\beta. \quad (5.2)$$

From the definition of  $I_n$ , we have  $I_n = E_n^\beta \circ \text{Pr}_n \circ T_n \circ I \circ E_n^\beta \circ \text{Pr}_n \circ T_n$ , thus we see  $I_n$  and  $i_n$  are themselves related by

$$I_n = E_n^0 \circ i_n \circ \text{Pr}_n \circ T_n. \quad (5.3)$$

The situation for both  $I_n$  and  $i_n$  is represented in the following two diagrams,

$$\begin{array}{ccc} C_b^\beta(\mathbb{R}^d) & \xrightarrow{I_n} & C_b^0(\mathbb{R}^d) \\ \hat{\pi}_n^\beta \downarrow & & \hat{\pi}_n^0 \uparrow \\ C_b^\beta(\mathbb{R}^d) & \xrightarrow{I} & C_b^0(\mathbb{R}^d) \end{array} \quad \begin{array}{ccc} C_*(G_n) & \xrightarrow{i_n} & C_*(G_n) \\ E_n^\beta \downarrow & & \text{Pr}_n \circ T_n \uparrow \\ C_b^\beta(\mathbb{R}^d) & \xrightarrow{I} & C_b^0(\mathbb{R}^d) \end{array}$$

Now, the space  $C_*(G_n)$  is finite dimensional (Remark 4.2), and the map  $i_n : C_*(G_n) \rightarrow C_*(G_n)$  is Lipschitz continuous. Therefore, tools available for Lipschitz functions in the finite dimensional setting can be applied to  $i_n$  and then related to  $I_n$  via (5.3).

We recall the generalized derivative of  $i_n$  in the sense of Clarke [18, Section 2.6].

**Definition 5.1.** Let  $V$  be a Banach space, and  $T : V \rightarrow V$  a Lipschitz continuous function. We define the set of generalized derivatives of  $T$ , by

$$\mathcal{D}T := \text{c.h.}\{L : V \rightarrow V \mid L = \lim_k L_k \text{ where } L_k = DT(x_k), T \text{ is differentiable at } x_k \forall k\}.$$

By Rademacher’s theorem, the set  $\mathcal{D}T$  is not empty when  $V$  is finite dimensional. Applying this to  $i_n : C_*(G_n) \rightarrow C_*(G_n)$ , we have, first, that  $\mathcal{D}i_n$  is non-empty, and secondly that  $\mathcal{D}I_n$  is non-empty as well, this is proved in Lemma 5.3, where we describe the relationship between  $\mathcal{D}i_n$  and  $\mathcal{D}I_n$ . The following Lemma is the mean value theorem for nonsmooth Lipschitz functions between finite dimensional spaces (note the similarity with Theorem 2.5).

**Lemma 5.2.** Assume that  $I : C_b^\beta(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$  is Lipschitz. For any  $u, v \in C_*(G_n)$ , there is a  $L \in \mathcal{D}i_n$  such that

$$i_n(u, x) - i_n(v, x) = L(u - v, x).$$

**Proof.** We refer the reader to [18, Proposition 2.6.5] for a proof of the lemma.  $\square$

The second lemma is basically the chain rule.

**Lemma 5.3.** Assume that  $I : C_b^\beta(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$  is Lipschitz. The set  $\mathcal{D}I_n$  is non-empty, and for any  $L \in \mathcal{D}I_n$  there is a  $\tilde{L} \in \mathcal{D}i_n$  such that

$$L = E_n^0 \circ \tilde{L} \circ T_n,$$

conversely, any  $L$  defined in this way for some  $\tilde{L} \in \mathcal{D}i_n$  belongs to  $\mathcal{D}I_n$ .

**Proof.** Note that  $I_n$  is differentiable at a point  $u$  if and only if  $i_n$  is differentiable at  $\tilde{u} = T_n u$ , a fact which follows applying the chain rule to the identities (5.2) and (5.3). Furthermore, at such  $u$ 's we have

$$DI_n(u) = E_n^* \circ Di_n(\tilde{u}) \circ T_n.$$

If  $u_k$  is a sequence along which  $I_n$  is differentiable, and  $L_k := DI_n(u_k)$  converges to some  $L$ , then the sequence  $\tilde{L}_k := Di_n(\tilde{u}_k)$  has a limit  $\tilde{L}$ , and  $L = E_n^* \circ \tilde{L} \circ T_n$ , taking the convex hull and by the linearity of  $E_n^*$  and  $T_n$ , the lemma follows.  $\square$

The following remark will not be of any relevance until the proof of [Theorem 1.11](#) at the end of this section, but we include it here to illustrate how [Lemmas 5.2](#) and [5.3](#) immediately yield a min–max formula for  $I_n(u, x)$  (for  $x \in G_n$ ).

**Remark 5.4.** Fix  $n$  and let  $x \in G_n$ . Then for any  $u \in C_b^\beta(\mathbb{R}^d)$  we have

$$I_n(u, x) \leq \max_{L \in \mathcal{D}I_n} \{I_n(v, x) + L(u - v, x)\}, \quad \forall x \in G_n, u, v \in C_b^\beta(\mathbb{R}^d). \quad (5.4)$$

Indeed, according to [Lemma 5.2](#) given  $u$  and  $v$  says there is some  $\tilde{L} \in \mathcal{D}i_n$  such that

$$i_n(u) - i_n(v) = \tilde{L}(u - v).$$

In this case, we have  $E_n^0(i_n(u)) - E_n^0(i_n(v)) = E_n^0(\tilde{L}(u - v))$ , and thus setting  $L := E_n^0 \circ \tilde{L} \circ T_n \in \mathcal{D}I_n$ , we have

$$I_n(u) = I_n(v) + L(u - v),$$

and (5.4) immediately follows.

Next we make an elementary observation regarding the nature of the operators  $L \in \mathcal{D}I_n$ . This observation is merely a consequence of the factorization of  $I_n$  through the space  $C(G_n)$ .

**Remark 5.5.** For each  $L \in \mathcal{D}I_n$  there is a function  $K = K_L$ ,  $K : G_n \times G_n \rightarrow \mathbb{R}$  such that

$$Lu(x) = \sum_{y \in G_n} K(x, y)u(x + y), \quad \forall u \in C_b^\beta(\mathbb{R}^d). \quad (5.5)$$

Indeed, simply let us use the basis functions  $\{e_y\}_{y \in G_n} \subset C(G_n)$  given by

$$e_y(x) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{if } x \neq y. \end{cases}$$

Observe that for any  $u \in C_b^\beta(\mathbb{R}^d)$  the function  $T_n u$  has finite support, and in particular  $T_n u = \sum_{y \in G_n} u(y) e_y$  as the sum on the right has at most a finite number of non-zero terms. Thanks to [Lemma 5.3](#), there is some  $\tilde{L} \in \mathcal{D}I_n$  such that  $L = E_n^0 \circ \tilde{L} \circ T_n$  and therefore,

$$Lu(x) = \sum_{y \in G_n} (\tilde{L}e_y)(x)u(y) = \sum_{y \in G_n - x} (\tilde{L}e_{x+y})(x)u(x+y), \quad \forall x \in G_n.$$

Then, defining  $K_L(x, y) = (\tilde{L}e_{x+y})(x)$  for  $x, y \in G_n$  the identity [\(5.5\)](#) follows.

For the rest of this section we analyze the operators  $I_n$  and the sets  $\mathcal{D}I_n$  and obtain in the limit a min–max formula for  $I_n$ . We shall focus on operators satisfying [Assumption 1.4](#). As we see below this property is inherited – to some extent – by the operators  $I_n$ , and by any operator  $L \in \mathcal{D}I_n$ , this fact is covered in the next two propositions. In the subsections that follow, we will use the spatial regularity afforded by [Assumption 1.4](#) to show that the operators in the family  $\mathcal{D}I_n$  have coefficients enjoying some regularity, which in the limit yields regular coefficients.

**Proposition 5.6.** *Let  $I$  be Lipschitz and satisfy [Assumption 1.4](#). Let  $x_1, x_2 \in G_n$  and  $h = x_1 - x_2$ , and  $r \geq 2^{4-n}$ . Then, for any  $u, v \in C_b^\beta(\mathbb{R}^d)$  we have*

$$\begin{aligned} & |I_n(v + \tau_{-h}u, x_1) - I_n(v, x_1) - (I_n(v + u, x_2) - I_n(v, x_2))| \\ & \leq \omega(|h|)C(2r) \left( \|u\|_{C^\beta(B_{4r}(x_2))} + \|u\|_{L^\infty(\mathcal{C}B_r(x_2))} \right). \end{aligned}$$

where  $\omega(\cdot)$  is the modulus of continuity and  $C(\cdot)$  the function given by [Assumption 1.4](#).

**Proof.** Observe that

$$I_n(v + \tau_{-h}u, x_1) - I_n(v, x_1) = I(\pi_n^\beta v + \pi_n^\beta(\tau_{-h}u), x_1) - I_n(\pi_n^\beta, x_1),$$

and recall that [Proposition 4.14](#) says that  $\pi_n^\beta(\tau_{-h}u) = \tau_{-h}\pi_n^\beta(u)$  when  $G_n + h = G_n$ .

Therefore, applying the bound in [Assumption 1.4](#) with  $\frac{3}{2}r$ ,

$$\begin{aligned} & |I_n(v + \tau_{-h}u, x_1) - I_n(v, x_1) - (I_n(v + u, x_2) - I_n(v, x_2))| \\ & = |I(\pi_n^\beta v + \tau_{-h}(\pi_n^\beta u), x_1) - I_n(\pi_n^\beta, x_1) - (I(\pi_n^\beta v + \pi_n^\beta u, x_2) - I(\pi_n^\beta v, x_2))| \\ & \leq \omega(|x_1 - x_2|)C(3r/2) \left( \|\pi_n^\beta u\|_{C^\beta(B_{3r}(x))} + \|\pi_n^\beta u\|_{L^\infty(\mathcal{C}B_{3r/2}(x))} \right). \end{aligned}$$

Now, provided  $r \geq 2^{4-n}$ , we have

$$\begin{aligned} \|\pi_n^\beta u\|_{C^\beta(B_{3r}(x))} & \leq C\|u\|_{C^\beta(B_{4r}(x))}, \\ \|\pi_n^\beta u\|_{L^\infty(\mathcal{C}B_{3r/2}(x))} & \leq C\|u\|_{L^\infty(\mathcal{C}B_r(x))}, \end{aligned}$$

the proposition follows.  $\square$

**Proposition 5.7.** *Let  $I$  be Lipschitz and satisfy [Assumption 1.4](#). Given  $L \in \mathcal{D}I_n$ ,  $x_1, x_2 \in G_n$ ,  $r \geq 2^{4-n}$  and  $u \in C_b^\beta(\mathbb{R}^d)$ , we have the inequality*

$$|L(\tau_{-h}u, x_1) - L(u, x_2)| \leq \omega(|h|)C(2r) \left( \|u\|_{C^\beta(B_{4r}(x_2))} + \|u\|_{L^\infty(\mathcal{C}B_r(x_2))} \right). \quad (5.6)$$

Here,  $h = x_1 - x_2$  and  $\omega(\cdot)$  and  $C(\cdot)$  are given by [Assumption 1.4](#).

**Proof.** Consider any  $v \in C_b^\beta(\mathbb{R}^d)$  such that  $I_n$  is differentiable at  $v$  with derivative  $L$ . Then,

$$\begin{aligned} L(\tau_{-h}u, x_1) &= \lim_{s \rightarrow 0} \frac{1}{s} (I_n(v + s\tau_{-h}u, x_1) - I_n(v, x_1)), \\ L(u, x_2) &= \lim_{s \rightarrow 0} \frac{1}{s} (I_n(v + su, x_2) - I_n(v, x_2)). \end{aligned}$$

By [Proposition 5.6](#), we have

$$\begin{aligned} &|L(\tau_{-h}u, x_1) - L(u, x_2)| \\ &= \limsup_{s \rightarrow 0} \frac{1}{s} |I_n(v + s\tau_{-h}u, x_1) - I_n(v, x_1) - (I_n(v + su, x_2) - I_n(v, x_2))|, \\ &\leq \omega(|h|)C(2r) \limsup_{s \rightarrow 0} \frac{1}{s} \left( \|su\|_{C^\beta(B_{2r}(x))} + \|su\|_{L^\infty(\mathcal{C}B_r(x))} \right), \\ &= \omega(|h|)C(2r) \left( \|u\|_{C^\beta(B_{2r}(x))} + \|u\|_{L^\infty(\mathcal{C}B_r(x))} \right). \end{aligned}$$

This proves the desired inequality for those  $L \in \mathcal{D}I_n$  which happen to be the derivative of  $I_n$  at a point of differentiability. This property is clearly preserved under limits and convex combinations, so it follows any  $L \in \mathcal{D}I_n$  has the desired property.  $\square$

The following proposition is directly related to [Proposition 4.20](#).

**Proposition 5.8.** *Assume that  $I$  is Lipschitz and satisfies [Assumption 1.1](#). For  $f \in C_c^\infty(\mathbb{R}^d)$  let  $w(x) = f(x - x_0)\eta_0(|x - x_0|^\beta)$  with  $\eta_0$  as in [\(4.3\)](#), then*

$$I(\pi_n^\beta u + \pi_n^\beta w, x) - I(\pi_n^\beta u, x) \leq C\|f\|_{L^\infty}.$$

If instead we have  $w(x) = f(x - x_0)\eta_0(|x - x_0|^{\beta_0})$  with  $f$  non-negative and some  $\beta_0 > \beta$ , then

$$I(\pi_n^\beta u + \pi_n^\beta w, x) - I(\pi_n^\beta u, x) \geq -C\|f\|_{L^\infty}h_n^\gamma,$$

for some constant  $\gamma = \gamma(\beta_0, \beta) \in (0, 1)$ .

**Proof.** We apply [Proposition 4.20](#), and we have with  $\hat{R}_{\beta, n, w, x_0}$  from the same proposition,

$$\pi_n^\beta w(x) \leq \hat{w}(x) := C\|f\|_{L^\infty} \left( \eta_0(|x - x_0|^\beta) + \hat{R}_{\beta, n, w, x_0}(x) \right), \quad \forall x \in \mathbb{R}^d,$$

with equality holding for  $x = x_0$ . It follows that  $\pi_n^\beta u + \pi_n^\beta w$  is touched from above at  $x_0$  by  $\pi_n^\beta u + \hat{w}$ . Then, since  $I(\cdot, x)$  has the GCP,

$$I(\pi_n^\beta u + \pi_n^\beta w, x) \leq I(\pi_n^\beta u + \hat{w}, x)$$

This means that

$$I(\pi_n^\beta u + \pi_n^\beta w, x_0) - I(\pi_n^\beta u, x_0) \leq I(\pi_n^\beta u + \hat{w}, x_0) - I(\pi_n^\beta u, x_0) \leq C\|\hat{w}\|_{C^\beta}.$$

Since  $\|\hat{w}\|_{C^\beta} = \|f\|_{L^\infty}\|\eta_0(|\cdot - x_0|^\beta) + \hat{R}_{\beta, n, w, x_0}\|_{C^\beta} \leq C\|f\|_{L^\infty}$  the first inequality is proved. For the second inequality, we apply [Remark 4.19](#) directly, and use that  $I$  has the GCP to conclude that

$$I(\pi_n^\beta u + \pi_n^\beta w + \hat{R}_{\beta, n, w, x_0}, x_0) \geq I(\pi_n^\beta u, x_0).$$

Then, using the Lipschitz property of  $I$  we conclude that

$$I(\pi_n^\beta u + \pi_n^\beta w, x_0) - I(\pi_n^\beta u, x_0) \geq -C\|\hat{R}_{\beta, n, w, x_0}\|_{C^\beta} \geq -Ch_n^\gamma\|f\|_{L^\infty},$$

where we used that  $|w(x)| \leq C\|f\|_{L^\infty} \min\{1, |x - x_0|^{\beta_0}\}$  and [Remark 4.19](#) to obtain the last inequality.  $\square$

**Proposition 5.9.** *Let  $I$  be Lipschitz and satisfy [Assumption 1.3](#). Let  $R \geq 1$  and  $w \in C_b^\beta(\mathbb{R}^d)$  with  $w \equiv 0$  in  $B_{3R}(x_0)$ , then for any  $x \in \cap B_R(x_0)$  we have*

$$|I(\pi_n^\beta u + \pi_n^\beta w, x) - I(\pi_n^\beta u, x)| \leq \rho(R) \|w\|_{L^\infty(\mathbb{R}^d)},$$

where  $\rho$  is the rate coming from [Assumption 1.3](#).

**Proof.** If  $w \equiv 0$  in  $B_{3R}(x_0)$ , then  $\pi_n^\beta w \equiv 0$  in  $B_{2R}(x_0)$ . In other words,  $\pi_n^\beta u$  and  $\pi_n^\beta u + \pi_n^\beta w$  are identically equal in  $B_{2R}(x_0)$ . Therefore, [Assumption 1.3](#) says that

$$|I(\pi_n^\beta u + \pi_n^\beta w, x) - I(\pi_n^\beta u, x)| \leq \rho(R) \|\pi_n^\beta w\|_{L^\infty(\mathbb{R}^d)}, \quad \forall x \in B_R(x_0).$$

By [Proposition 4.20](#),  $\|\pi_n^\beta w\|_{L^\infty(\mathbb{R}^d)} \leq \|w\|_{L^\infty(\mathbb{R}^d)}$ , the proposition is proved.  $\square$

## 5.2. Properties of $\mathcal{DI}_n$

For each  $L \in \mathcal{DI}_n$  and  $x \in G_n$  we define a Borel measure  $\mu_L(x, dy)$  (which is possibly signed) as follows

$$\mu_L(x, dy) := \sum_{y \in G_n \setminus \{0\}} K_L(x, y) \delta_{x+y}. \quad (5.7)$$

where  $K_L(x, y)$  is as in [Remark 5.5](#). From its definition, it is immediate that given  $\phi \in C^\beta$  and  $x \in G_n$  then

$$L(\phi, x) = \int_{\mathbb{R}^d} \phi(x + y) d\mu_L(x, dy).$$

**Proposition 5.10.** *Assume that  $I$  is Lipschitz and satisfies [Assumption 1.1](#). For each  $L \in \mathcal{DI}_n$  and  $x \in G_n$ , and  $\eta_0(t)$  the function in [\(4.3\)](#),*

$$\sup_n \sup_{x \in G_n} \int_{\mathbb{R}^d} f(y) \eta_0(|y|^\beta) \mu_L(x, dy) \leq C \|f\|_{L^\infty}, \quad \forall f \in C_c^\infty(\mathbb{R}^d).$$

**Proof.** Fix  $x_0 \in G_n$ . Let us assume first that  $\beta \neq 1$ . Let  $w(x) = f(x - x_0) \eta_0(|x - x_0|^\beta)$ , then

$$L(w, x_0) = \int_{\mathbb{R}^d} \phi(y) \eta_0(|y|^\beta) \mu_L(x_0, dy).$$

Therefore it suffices to show there is a universal constant such that

$$L(w, x_0) \leq C \|f\|_{L^\infty}, \quad \forall L \in \mathcal{DI}_n.$$

Let us prove this when  $L$  arises as the derivative of  $I_n$  at some  $v \in C_b^\beta$ , namely, that

$$L(\phi, x_0) = \lim_{s \rightarrow 0} (I_n(v + s\phi, x_0) - I_n(v, x_0))/s.$$

In this case, we can apply [Proposition 5.8](#) to the expression on the right and conclude that

$$\lim_{s \rightarrow 0} (I_n(v + sw, x_0) - I_n(v, x_0))/s \leq C \|f\|_{L^\infty},$$

where we used that when  $\beta \neq 1$  the function  $\eta_0(|\cdot - x_0|^\beta)$  belongs to  $C_b^\beta(\mathbb{R}^d)$  and the norm  $\|\eta_0(|\cdot - x_0|^\beta)\|_{C^\beta}$  is bounded in terms of  $\beta, d$ , and the function  $\eta_0$ . This the desired estimate for such  $L$ . Since this property is clearly preserved under limits and convex combinations, it follows that the property holds for all elements of  $\mathcal{DI}_n$ .

The case  $\beta = 1$  proceeds similarly, except one first fixes  $\varepsilon \in (0, 1)$  and considers the function  $\eta_0(|x - x_0|^{\beta+\varepsilon})$  instead. After proceeding as in the previous case, we obtain the estimate

$$\int_{\mathbb{R}^d} f(y) \eta_0(|y|^{\beta+\varepsilon}) \mu_L(x_0, dy) \leq C \|f\|_{L^\infty},$$

for every  $L \in \mathcal{DI}_n$  and  $x_0 \in G_n$ . The constant  $C$  is independent of  $\varepsilon \in (0, 1)$ , since  $\|\eta_0(|\cdot - x_0|^\beta)\|_{C^1}$  is independent of  $\varepsilon$  when  $\varepsilon > 0$ . Letting  $\varepsilon \searrow 0$  for the integral on the left (and using the special form of  $\mu_L(x_0, dy)$ ) one obtains the estimate in the case  $\beta = 1$ .  $\square$

**Proposition 5.11.** *Assume that  $I$  is Lipschitz and satisfies Assumption 1.1. Let  $f \in C_c^\infty(\mathbb{R}^d)$  be a non-negative function. There is a constant  $C = C(I, d, \beta, \beta_0)$  such that given  $\beta_0 > \beta$  then for each  $L \in \mathcal{DI}_n$  and  $x \in G_n$ ,*

$$\inf_n \inf_{x \in G_n} \int_{\mathbb{R}^d} f(y) \eta_0(|y|^{\beta_0}) \mu_L(x, dy) \geq -Ch_n^\gamma \|f\|_{L^\infty}.$$

As before,  $\eta_0$  is the function in (4.3), and  $\gamma = \gamma(\beta, \beta_0)$ .

**Proof.** As in the proof of the previous proposition, we note that if  $x_0 \in G_n$ ,  $w(x) := f(x - x_0) \eta_0(|x - x_0|^{\beta_0})$ , and  $L \in \mathcal{DI}_n$ , then

$$L(w, x_0) = \int_{\mathbb{R}^d} f(y) \eta_0(|y|^\beta) \mu_L(x_0, dy).$$

As in the previous Proposition, it suffices to show that  $L(w, x_0) \geq -C \|f\|_{L^\infty} h_n^\gamma$ , and from  $\mathcal{DI}_n$ 's definition, it suffices to show this for those  $L$ 's in  $\mathcal{DI}_n$  which are the derivative of  $I_n$  at some  $u \in C_b^\beta(\mathbb{R}^d)$ . In this case, given that  $f \geq 0$ , we may apply the second part of Proposition 5.8 to obtain

$$L(w, x_0) = \lim_{s \rightarrow 0} \frac{I(\pi_n^\beta u + \pi_n^\beta(sw), x_0) - I(\pi_n^\beta u, x_0)}{s} \geq \lim_{s \rightarrow 0} -\frac{C \|sf\|_{L^\infty} h_n^\gamma}{s} = -Ch_n^\gamma \|f\|_{L^\infty},$$

and the proposition is proved.  $\square$

Let us recall the function

$$P_{\phi, \eta, u, x}(\cdot) = u(x) + \phi(\cdot - x)(\nabla u(x), \cdot - x) + \frac{1}{2}\eta(\cdot - x)(D^2 u(x)(\cdot - x), (\cdot - x)).$$

In this section we introduce a variation on this function. This modification takes into account the geometry of the grid  $G_n$  as well as the regularity exponent  $\beta$ , and will be used in a way analogous to the previous section.

$$P_{\phi, \eta, u, x}^{(n)}(\cdot) = \begin{cases} u(x) & \text{if } \beta \in (0, 1), \\ u(x) + \phi(\cdot - x)((\nabla_n)^1 u(x), \cdot - x) & \text{if } \beta \in [1, 2), \\ u(x) + \phi(\cdot - x)((\nabla_n)^1 u(x), \cdot - x) + \frac{1}{2}\eta(\cdot - x)((\nabla_n)^2 u(x)(\cdot - x), \cdot - x) & \text{if } \beta \in [2, 3). \end{cases}$$

Associated with this, we introduce functions in  $G_n$  taking (respectively) scalar, vector, and matrix values.

First, some notation. To functions  $\eta, \phi \in \mathcal{S}$  we associate the following family of functions

$$\phi_i(y) = \phi(y)y_i, \quad i = 1, \dots, d, \quad \eta_{ij}(y) = \eta(y)y_i y_j, \quad i, j = 1, \dots, d.$$

Then, for  $L \in \mathcal{DI}_n$  and  $\eta, \phi \in \mathcal{S}$  we define a symmetric matrix  $A_{L, \eta}$ , a vector  $B_{L, \phi}$ , and a scalar  $C_L$ . These are functions in  $G_n$  defined by the formulas,

$$(A_{L, \eta}(x))_{ij} = L(\tau_{-x} \eta_{ij}, x), \quad i, j = 1, \dots, d, \quad (5.8)$$

$$(B_{L, \phi}(x))_i = L(\tau_{-x} \phi_i, x), \quad i = 1, \dots, d, \quad (5.9)$$

$$C_L(x) = L(1, x).. \quad (5.10)$$

The functions  $A_{L, \eta}, B_{L, \phi}, C_L$ , and  $\mu_L$  give us a representation for  $L(u, x)$  for  $x \in G_n$ .

**Proposition 5.12.** Assume that  $I$  is Lipschitz. Let  $L \in \mathcal{DI}_n$ , then for  $\beta \in [2, 3)$  and  $u \in C_b^\beta(\mathbb{R}^d)$  we may write it as

$$L(u, x) = C_L(x)u(x) + B_{L,\phi}(x) \cdot (\nabla_n)^1 u(x) + \text{tr}(A_{L,\eta}(x)(\nabla_n)^2 u(x)) \\ + \int_{\mathbb{R}^d} u(x+y) - P_{\phi,\eta,u,x}^{(n)}(x+y) \mu_L(x, dy).$$

For  $\beta \in [1, 2)$

$$L(u, x) = C_L(x)u(x) + B_{L,\phi}(x) \cdot (\nabla_n)^1 u(x) + \int_{\mathbb{R}^d} u(x+y) - P_{\phi,\eta,u,x}^{(n)}(x+y) \mu_L(x, dy),$$

and for  $\beta \in [0, 1)$

$$L(u, x) = C_L(x)u(x) + \int_{\mathbb{R}^d} u(x+y) - u(x) \mu_L(x, dy).$$

**Proof.** We do the case  $\beta \geq 2$  explicitly, as the others are identical. Let us compute  $L(u, x)$  by adding and subtracting  $L(P_{\phi,\eta,u,x}^{(n)}, x)$ ,

$$L(u, x) = L(u - P_{\phi,\eta,u,x}^{(n)}, x) + L(P_{\phi,\eta,u,x}^{(n)}, x).$$

From Remark 5.5, (5.7), we have that

$$L(u - P_{\phi,\eta,u,x}^{(n)}, x) = \int_{\mathbb{R}^d} u(x+y) - P_{\phi,\eta,u,x}^{(n)}(x+y) \mu_L(x, dy)$$

As for the other term, we observe that

$$L(P_{\phi,\eta,u,x}^{(n)}, x) = u(x)L(1, x) + \sum_{i=1}^d (\nabla_1 u)_i^n(x)L(\tau_{-x}\phi_i, x) + \frac{1}{2} \sum_{i,j=1}^d (\nabla_n)_{ij}^2 u(x)L(\tau_{-x}\eta_{ij}, x).$$

Rewriting the terms on the right and gathering the terms, we conclude that

$$L(P_{\phi,\eta,u,x}^{(n)}, x) = C_L(x)u(x) + (B_{L,\phi}(x), (\nabla_n)^1 u(x)) + \text{tr}(A_{L,\eta}(x)(\nabla_n)^2 u(x)).$$

The remaining cases of  $\beta$  follow from the corresponding definition of  $P_{\phi,\eta,u}^{(n)}$  in those cases.  $\square$

The next two propositions say that the terms appearing Proposition 5.12 satisfy a uniform continuity in  $G_n$ . The first refers to the measure  $\mu_L$ .

**Proposition 5.13.** Assume  $I$  satisfies Assumptions 1.1, 1.3, and 1.4, as stated for  $C_b^\beta(\mathbb{R}^d)$ . Let  $L \in \mathcal{DI}_n$ ,  $x_1, x_2 \in G_n$ , and  $r \geq 2^{4-n}$ . There is a constant  $C(r)$  such that for any  $\zeta \in C_c(\mathbb{R}^d)$  such that  $\zeta \equiv 0$  in  $B_r$ ,

$$\left| \int_{CB_r} \zeta(y) \mu_L(x_1, dy) - \int_{CB_r} \zeta(y) \mu_L(x_2, dy) \right| \leq C(r) \|\zeta\|_{L^\infty} \omega(|x_1 - x_2|),$$

where  $\omega$  is the modulus from Assumption 1.4. In particular,

$$\|\mu_L(x_1, dy) - \mu_L(x_2, dy)\|_{\text{TV}(CB_r)} \leq C(r) \omega(|x_1 - x_2|).$$

On the other hand, if  $\zeta \in C^0(\mathbb{R}^d)$  is such that  $\zeta \equiv 0$  in  $B_{3R}(0)$  for some  $R > 1$ , then for any  $x_0 \in G_n$  we have

$$\int_{\mathbb{R}^d} \zeta(y) \mu_L(x_0, dy) \leq \rho(R) \|\zeta\|_{L^\infty(\mathbb{R}^d)},$$

where  $\rho(\cdot)$  is the function from Assumption 1.3.

**Proof.** From the fact that  $\tau_{-x_1}\zeta$  and  $\tau_{-x_2}\zeta$  vanish in, respectively,  $B_r(x_1)$  and  $B_r(x_2)$ , we have

$$\begin{aligned} L(\tau_{-x_1}\zeta, x_1) - L(\tau_{-x_2}\zeta, x_2) &= \int_{\mathbb{R}^d} \zeta(y) d\mu(x_1, dy) - \int_{\mathbb{R}^d} \zeta(y) d\mu(x_2, dy) \\ &= \int_{\mathcal{C}B_r} \zeta(y) d\mu(x_1, dy) - \int_{\mathcal{C}B_r} \zeta(y) d\mu(x_2, dy). \end{aligned}$$

Since  $\zeta \equiv 0$  in  $B_r$ , [Proposition 5.7](#) says that, as long as  $r \geq 2^{4-n}$

$$\left| \int_{\mathcal{C}B_r} \zeta(y) d\mu(x_1, dy) - \int_{\mathcal{C}B_r} \zeta(y) d\mu(x_2, dy) \right| \leq \omega(|x_1 - x_2|) C(r) \|\zeta\|_{L^\infty(\mathcal{C}B_r)}.$$

This proves the first estimate, for the second one, fix  $\zeta$  and  $x_0 \in G_n$ , and define  $w(x) = \tau_{-x_0}\zeta$ , then

$$L(w, x_0) = \int_{\mathbb{R}^d} \zeta(y) \mu_L(x_0, dy).$$

Therefore, as before, it suffices for us to bound  $L(w, x_0)$  for every  $L \in \mathcal{DI}_n$ , and from the definition of  $\mathcal{DI}_n$  it suffices to prove the bound for those  $L$  such that  $L = \mathcal{DI}_n(v)$  at some  $v$ . In this case, [Proposition 5.9](#) says that

$$L(w, x_0) = \lim_{s \rightarrow 0} \frac{1}{s} (I_n(v + sw, x_0) - I_n(v, x_0)) \leq \rho(R) \|w\|_{L^\infty(\mathbb{R}^d)} = \rho(R) \|\zeta\|_{L^\infty(\mathbb{R}^d)} \quad \square$$

The following notation will be useful in what follows,

$$\begin{aligned} \alpha(r, \eta) &:= C(2r) \left( \max_{1 \leq i, j \leq d} \|\eta_{ij}\|_{C^\beta(B_{4r})} + \max_{1 \leq i, j \leq d} \|\eta_{ij}\|_{L^\infty(\mathcal{C}B_r)} \right), \\ \beta(r, \phi) &:= C(2r) \left( \max_{1 \leq i \leq d} \|\phi_i\|_{C^\beta(B_{4r})} + \max_{1 \leq i \leq d} \|\phi_i\|_{L^\infty(\mathcal{C}B_r)} \right), \end{aligned}$$

where  $C(r)$  is as in [Assumption 1.4](#) (see also [Proposition 5.6](#)).

**Proposition 5.14.** *Assume  $I$  satisfies Assumptions 1.1, 1.3, and 1.4, as stated for  $C_b^\beta(\mathbb{R}^d)$ . Let  $L \in \mathcal{DI}_n$ ,  $r \geq 2^{4-n}$ , and  $x_1, x_2 \in G_n$ , then*

$$\begin{aligned} |A_{L,\eta}(x_1) - A_{L,\eta}(x_2)| &\leq \alpha(r, \eta) \omega(|x_1 - x_2|), \\ |B_{L,\phi}(x_1) - B_{L,\phi}(x_2)| &\leq \beta(r, \phi) \omega(|x_1 - x_2|), \\ |C_L(x_1) - C_L(x_2)| &\leq C(r) \omega(|x_1 - x_2|). \end{aligned}$$

**Proof.** Fix  $x_1, x_2 \in G_n$  and let  $h = x_2 - x_1$ . Applying [Proposition 5.7](#) to  $x = x_1$  and  $h$ , with the functions  $1$ ,  $\phi_i$ , and  $\eta_{ij}$ , we see that for  $r \geq 2^{4-n}$

$$\begin{aligned} |L(\tau_{-x_2}\eta_{ij}, x_2) - L(\tau_{-x_1}\eta_{ij}, x_1)| &\leq \alpha(\eta, r) \omega(|x_1 - x_2|), \\ |L(\tau_{-x_2}\phi_i, x_2) - L(\tau_{-x_1}\phi_i, x_1)| &\leq \beta(\phi, r) \omega(|x_1 - x_2|), \\ |L(1, x_2) - L(1, x_1)| &\leq C \omega(|x_1 - x_2|). \end{aligned}$$

These inequalities respectively amount to the stated estimate for  $A_{L,\eta}$ ,  $B_{L,\phi}$ , and  $C_L$ .  $\square$

### 5.3. Properties of $\mathcal{D}_I$

Now, we define the set  $\mathcal{D}_I$ , which plays the role the Clarke differential played for  $I_n$  (we recall that c.h. stands for “convex hull”).

$$\mathcal{D}_I := \text{c.h.}\{L \mid \exists \{L_{n_k}\}, n_k \rightarrow \infty, L_{n_k} \in \mathcal{DI}_{n_k} \text{ s.t } L(u, \cdot) = \lim_k L_{n_k}(u, \cdot) \forall u\}. \quad (5.11)$$

**Remark 5.15.** We would like to note a point about notation and definitions, namely why above we have  $\mathcal{D}_I$  with  $I$  as a subscript. This is to avoid confusion (or perhaps, to promote it) by distinguishing it from the generalized derivative in the sense of Clarke from [Definition 5.1](#). The objects are closely related, and in fact one would hope that  $\mathcal{D}_I = \mathcal{D}I$ , but we are not concerned with whether this is actually the case as the above definition works for our purposes.

The following is an important Lemma that says – among other things – that  $\mathcal{D}_I$  is non-empty.

**Lemma 5.16.** *Assume  $I$  satisfies [Assumptions 1.1, 1.3, and 1.4](#), as stated for  $C_b^\beta(\mathbb{R}^d)$ . Given a sequence  $n_k \rightarrow \infty$  and operators  $L_{n_k}$  with  $L_{n_k} \in \mathcal{D}I_{n_k}$  for every  $k$ , and  $\phi, \eta \in \mathcal{S}$  we have the following*

(1) *There is a subsequence  $\bar{n}_k$  and functions  $A(x), B(x)$ , and  $C(x)$  defined on  $\mathbb{R}^d$  and taking values respectively in  $\mathbb{S}(d)$ ,  $\mathbb{R}^d$ , and  $\mathbb{R}$ , such that if  $x \in G_n$  for some  $n$  then we have the convergence*

$$A_{L_{\bar{n}_k}, \eta}(x) \rightarrow A(x), \quad B_{L_{\bar{n}_k}, \phi}(x) \rightarrow B(x), \quad C_{L_{\bar{n}_k}}(x) \rightarrow C(x).$$

(2) *There is a function  $\mu(x)$  in  $\mathbb{R}^d$ , taking values on the space of Lévy measures in  $\mathbb{R}^d$ , such that for every  $r > 0$ , and every  $x$  as before we have the convergence*

$$\lim_{k \rightarrow \infty} \|\mu_{L_{\bar{n}_k}}(x) - \mu(x)\|_{\text{TV}(\mathcal{C}B_r)} = 0.$$

(3) *The functions  $A, B, C$ , all have a modulus of continuity  $C\omega(2(\cdot))$ , while for each  $r > 0$  we have the estimate,*

$$\|\mu(x_1) - \mu(x_2)\|_{\text{TV}(\mathcal{C}B_r)} \leq C(r)\omega(2|x_1 - x_2|). \quad (5.12)$$

(4) *If we define  $L$  by*

$$\begin{aligned} L(u, x) := & \text{tr}(A(x)D^2u(x)) + B(x) \cdot \nabla u(x) + C(x)u(x) \\ & + \int_{\mathbb{R}^d} u(x+y) - P_{\phi, \eta, u, x}(x+y) \mu(x, dy) \end{aligned}$$

*Then,  $L \in \mathcal{D}_I$ .*

(5) *Moreover, if  $\beta < 2$ , then we have  $A(x) \equiv 0$ . Furthermore, if  $\beta < 1$  then  $B(x) \equiv 0$  and  $L$  takes the form*

$$L(u, x) = C(x)u(x) + \int_{\mathbb{R}^d} u(x+y) - u(x) \mu(x, dy).$$

**Proof.** Let us fix  $\eta$  and  $\phi$ . First of all, we invoke [Proposition 5.12](#) to obtain the collection of  $A_{L_{n_k}, \eta}, B_{L_{n_k}, \phi}, C_{L_{n_k}}$ , and  $\mu_{L_{n_k}}$ . Furthermore, already as a result of [Proposition 5.12](#), we have item (5) of the lemma.

*Step 1.* (Extension) We have a sequence of functions defined on varying, monotone increasing sets  $G_n$ . One way to show they converge (along a subsequence) to a function in  $\mathbb{R}^d$  is by extending them to all of  $\mathbb{R}^d$  and check whether the resulting sequences are pre-compact.

With this idea in mind, for each  $n \in \mathbb{N}$  we apply the Whitney extension to  $A_{L_n, \eta}, B_{L_n, \eta}, C_{L_n, \eta}$ ,

$$\hat{A}_{L_n, \eta}(x) := E_n^0(A_{L_n, \eta})(x), \quad \hat{B}_{L_n, \phi}(x) := E_n^0(B_{L_n, \phi})(x), \quad \hat{C}_{L_n}(x) := E_n^0(C_{L_n})(x).$$

We repeat the same for  $\mu_{L_n}$ , resulting in a map  $\hat{\mu}_{L_n}$  from  $\mathbb{R}^d$  to the space of Lévy measures, given by the formula

$$\hat{\mu}_{L_n}(x, dy) = \sum_{k=1}^{\infty} \phi_{n,k}(x) \mu(x_k, dy),$$

where  $\{\phi_k\}_k$  is the partition of unity from [Proposition 4.6](#). The functions  $\hat{A}_{L_n, \eta}, \hat{B}_{L_n, \phi}$ , and  $\hat{C}_{L_n}(x)$  all have modulus of continuity  $C\omega(2(\cdot))$ , thanks to [Proposition 5.14](#) and the properties of the Whitney extension

operator, see [54, Chapter VI, Theorem 3]. The same proof from reference [54] can be applied with minor modifications to show that for every  $r > 0$  we have

$$\|\hat{\mu}_{L_n}(x_1) - \hat{\mu}_{L_n}(x_2)\|_{\text{TV}(\mathcal{C}B_r)} \leq C(r)\omega(2|x_1 - x_2|).$$

Furthermore, for every  $x$ , by [Proposition 5.13](#),

$$|\hat{\mu}_{L_n}(x)|(\mathcal{C}B_R) \leq \rho(R),$$

where  $\rho(R) \rightarrow 0$  as  $R \rightarrow \infty$ . This shows that for each  $r > 0$ , the functions  $\{\hat{\mu}_{L_n}|_{\mathcal{C}B_r}\}_n$  are an equicontinuous family of functions taking values inside the space of measures  $\nu$  which are supported in  $\mathcal{C}B_r$  and such that  $\nu(\mathcal{C}B_R) \leq \rho(R)$  for all  $R \geq r$ . This space, equipped with the total variation distance, is a compact metric space.

*Step 2.* (Cantor diagonalization) We now use a standard Cantor diagonalization argument to obtain locally uniform convergence along a subsequence. We construct a family nested sequences  $\tilde{n}_k^m$  in the following recursive manner. First,  $\tilde{n}_k^1$  is a subsequence of  $n_k$  along which the functions converge uniformly in  $B_1$  to functions  $A^1(x), B^1(x)$ , and  $C^1(x)$  defined in  $B_1$ . Next, suppose that for  $m \in \mathbb{N}$  we have build a nested family of sequences  $\tilde{n}_k^1, \dots, \tilde{n}_k^m$  such that the functions  $A_{L_{\tilde{n}_k^m}, \eta}, \dots$ , etc. converge uniformly in  $B_m(0)$  to functions  $A^m(x), \dots$ , etc. In this case, we choose  $\tilde{n}_k^{m+1}$  to be a subsequence of  $\tilde{n}_k^m$  along which  $A_{L_{\tilde{n}_k^{m+1}}, \eta}, \dots$  converge uniformly in  $B_{m+1}$  to functions  $A^{m+1}(x), \dots$  and so on.

Having constructed these  $\tilde{n}_k^m$ , we define the sequence  $\tilde{n}_k$  as  $\tilde{n}_k := n_k^k$ . The resulting sequences converge locally uniformly, respectively, to  $A(x), B(x)$ , and  $C(x)$ .

*Step 3.* (Cantor diagonalization continued)

As noted at the end of Step 1, for every  $r > 0$ , the sequence  $\{\hat{\mu}_{L_{\tilde{n}_k}}\}_k$  is an equicontinuous family of functions taking values in a compact metric space. Therefore, we can apply the Arzela–Ascoli type theorem found in [24, p. 202] to obtain a subsequence  $\tilde{n}_k^1$  of  $\tilde{n}_k$  and a measure  $\mu^1$  such that

$$\lim_{k \rightarrow \infty} \sup_{x \in B_1} \|\hat{\mu}_{L_{\tilde{n}_k^1}}(x) - \mu^1(x)\|_{\text{TV}(\mathcal{C}B_{1/2})} = 0.$$

Now, suppose we have repeated this  $m$  times: we have  $\tilde{n}_k^m$  (a subsequence of  $\tilde{n}_k^{m-1}$ ), as well as a measure  $\mu^m$  such that

$$\lim_{k \rightarrow \infty} \sup_{x \in B_m} \|\hat{\mu}_{L_{\tilde{n}_k^m}}(x) - \mu^m(x)\|_{\text{TV}(\mathcal{C}B_{1/2^m})} = 0.$$

Then, using again the compactness theorem in [24, p. 202] we pick a subsequence  $\tilde{n}_k^{m+1}$  of  $\tilde{n}_k^m$  and a measure  $\mu^{m+1}$  such that

$$\lim_{k \rightarrow \infty} \sup_{x \in B_{m+1}} \|\hat{\mu}_{L_{\tilde{n}_k^{m+1}}}(x) - \mu^{m+1}(x)\|_{\text{TV}(\mathcal{C}B_{1/2^{m+1}})} = 0.$$

Observe that the measures  $\{\mu^m\}$  are such that  $\mu_{\mathcal{C}B_{1/2^m}}^{m+1}(x) = \mu^m(x)$  for all  $x \in B_m$ , which uniquely defines a direct limit measure  $\mu(x)$  for each  $x \in \mathbb{R}^d \setminus \{0\}$ . Letting  $\bar{n}_k := \tilde{n}_k^k$  we see that for every  $R > 0$  and  $r > 0$  we have

$$\lim_{k \rightarrow \infty} \sup_{x \in B_R} \|\hat{\mu}_{L_{\bar{n}_k}}(x) - \mu(x)\|_{\text{TV}(\mathcal{C}B_r)} = 0.$$

Since  $\bar{n}_k$  is a subsequence of  $\tilde{n}_k$ , we still have convergence of  $A_{L_{\bar{n}_k}, \eta}, \dots$  to  $A(x), \dots$  Moreover, the continuity estimates in the previous step all pass to the limit to give respective estimates for  $A(x), B(x), C(x)$ , and  $\mu(x)$  in the respective metrics.

Last but not least, we note that while  $\{\mu_{L_{\bar{n}_k}}\}_k$  are a sequence of signed measures, their limit  $\mu$  will be a measure, which follows at once from [Proposition 5.11](#).

*Step 4.* (Convergence)

First, note that for fixed  $u$ , we have that as  $n \rightarrow \infty$ ,

$$u(x + \cdot) - P_{\phi, \eta, u, x}^{(n)}(x + \cdot) \rightarrow u(x + \cdot) - P_{\phi, \eta, u, x}(x + \cdot) \text{ in } L^\infty(\mathbb{R}^d),$$

which in particular guarantees that, for every fixed  $r > 0$ ,

$$\lim_{k \rightarrow \infty} \int_{CB_r} u(x + y) - P_{\phi, \eta, u, x}^{(n_k)}(x + y) \mu_{L_{n_k}}(x, dy) = \int_{CB_r} u(x + y) - P_{\phi, \eta, u, x}(x + y) \mu(x, dy).$$

Then, by the bound in [Proposition 5.10](#), we conclude that

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} u(x + y) - P_{\phi, \eta, u, x}^{(n_k)}(x + y) \mu_{L_{n_k}}(x, dy) = \int_{\mathbb{R}^d} u(x + y) - P_{\phi, \eta, u, x}(x + y) \mu(x, dy).$$

Therefore, and taking into account the convergence of  $\hat{A}_{L_{\tilde{n}_k}, \eta}$ ,  $\hat{B}_{L_{\tilde{n}_k}, \phi}$ , and  $\hat{C}_{L_{\tilde{n}_k}}$ , and with  $L(u, x)$  defined as in the statement of the Lemma,  $x \in G_n$ , and  $u \in C_b^\beta(\mathbb{R}^d)$ , we have

$$\begin{aligned} \lim_{k \rightarrow \infty} L_{\tilde{n}_k}(x) &= \lim_{k \rightarrow \infty} \{ \text{tr}(\hat{A}_{L_{\tilde{n}_k}, \eta} D^2 u(x)) + \hat{B}_{L_{\tilde{n}_k}, \phi} \cdot \nabla u(x) + \hat{C}_{L_{\tilde{n}_k}}(x) u(x) \} \\ &\quad + \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} u(x + y) - P_{\phi, \eta, u, x}^{(\tilde{n}_k)}(x + y) \hat{\mu}_{L_{n_k}}(x, dy) \\ &= \text{tr}(AD^2 u(x)) + B \cdot \nabla u(x) + C(x) u(x) \\ &\quad + \int_{\mathbb{R}^d} u(x + y) - P_{\phi, \eta, u, x}(x + y) \hat{\mu}(x, dy), \end{aligned}$$

and we conclude that  $L \in \mathcal{D}_I$ .  $\square$

It is to be expected that every  $L \in \mathcal{D}_I$  satisfies the GCP, and thus, it has to be an operator of Lévy type. This is proved in the lemma below, and further, we show that the coefficients in the operator inherit a modulus of continuity from [Assumption 1.4](#).

**Lemma 5.17.** *Assume  $I$  satisfies [Assumptions 1.1](#), [1.3](#), and [1.4](#), as stated for  $C_b^\beta(\mathbb{R}^d)$ . Given  $L \in \mathcal{D}_I$ , and any  $\phi, \eta \in \mathcal{S}$ , the operator  $L$  can be represented as*

$$\begin{aligned} L(u, x) &= C_L(x) u(x) + B_{L, \phi}(x) \cdot \nabla u(x) + \text{tr}(A_{L, \eta}(x) D^2 u(x)) \\ &\quad + \int_{\mathbb{R}^d} u(x + y) - P_{\phi, \eta, u, x}(x + y) \mu(x, dy). \end{aligned}$$

Here,  $\mu_L(x, dy)$  is a Lévy measure satisfying the continuity estimate [\(5.12\)](#), and

$$\begin{aligned} (A_{L, \eta})_{ij}(x) &= L(\tau_{-x} \eta_{ij}, x), \\ (B_{L, \phi})_i(x) &= L(\tau_{-x} \phi_i, x), \\ C_L(x) &= L(1, x), \end{aligned}$$

all have modulus of continuity  $C\omega(2(\cdot))$ .

**Proof.** Fix  $\phi, \eta \in \mathcal{S}$ . Assume first that  $L$  is the limit of a sequence  $L_{n_k}$  with  $L_{n_k} \in \mathcal{D}_{I_{n_k}}$ . Then, by [Lemma 5.16](#) there is a subsequence  $\tilde{n}_k$  as well as (matrix, vector, scalar, measure)-valued functions  $A, B, C$ , and  $\mu$ , all such that

$$C_{L_{\tilde{n}_k}}(x) \rightarrow C(x), \quad B_{L_{\tilde{n}_k}, \phi_k}(x) \rightarrow B(x), \quad A_{L_{\tilde{n}_k}, \eta_k}(x) \rightarrow A(x), \quad \mu_{L_{\tilde{n}_k}}(x, dy) \rightarrow \mu(x, dy).$$

and, as a result, we have

$$\begin{aligned} L(u, x) &= \text{tr}(A(x) D^2 u(x)) + B(x) \cdot \nabla u(x) + C(x) u(x) \\ &\quad + \int_{\mathbb{R}^d} u(x + y) - P_{\phi, \eta, u, x}(y) \mu(x, dy). \end{aligned}$$

The estimate in [Proposition 5.10](#) in the limit as  $n \rightarrow \infty$  implies that

$$\int_{\mathbb{R}^d} \eta_0(|y|^\beta) \mu(x, dy) \leq C,$$

for some constant  $C$  independent of  $x$  and  $L$ . Meanwhile, also the  $n \rightarrow \infty$  limit of the estimate in [Proposition 5.11](#) implies that  $\mu(x, dy)$  is a non-negative measure in  $\mathbb{R}^d \setminus \{0\}$ . The positivity of  $\mu$  means that the previous estimate is equivalent to

$$\int_{\mathbb{R}^d} \min\{1, |y|^\beta\} \mu(x, dy) \leq C.$$

Since  $L_{\tilde{n}_k}(u, x) \rightarrow L(u, x)$ , for every  $u$ , we have in particular, for  $x \in \bigcup G_k$

$$(A_{L, \eta})_{ij}(x) = L_{\tilde{n}_k}(\tau_{-x} \eta_{ij}, x) \rightarrow L(\tau_{-x} \eta_{ij}, x).$$

From where it follows that  $(A_{L, \eta})_{ij}(x) = L(\tau_{-x} \eta_{ij}, x)$  (and thus for all  $x$ , by continuity), the exact same argument yields that  $(B_{L, \phi})_i(x) = L(\tau_{-x} \phi_i, x)$ , and  $C_L(x) = L(1, x)$ , and the lemma is proved.  $\square$

Let us now simplify things by doing away with the auxiliary functions  $\phi$  and  $\eta$ . To accomplish this, we shall make use of the auxiliary functions from [Section 3](#).

$$\phi_\delta(x) = \psi_{\delta, 1-\delta}, \quad \eta_\delta(x) = \psi_{\delta, \delta}(x), \quad (5.13)$$

where we recall the two-parameter of functions  $\psi_{r, R}(x)$  was defined in [\(3.2\)](#). An important property of these one-parameter families is the bound

$$\sup_{\delta \in (0, 1)} \{\|\phi_\delta\|_{C^\beta(B_{1/2})} + \|\phi_\delta\|_{L^\infty(\mathbb{R}^d)} + \max_{ij} \|\eta_\delta x_i x_j\|_{C^\beta(\mathbb{R}^d)}\} < \infty. \quad (5.14)$$

**Corollary 5.18.** *Assume  $I$  satisfies [Assumptions 1.1, 1.3, and 1.4](#), as stated for  $C_b^\beta(\mathbb{R}^d)$ . Then, any  $L \in \mathcal{D}_I$  has the form,*

$$\begin{aligned} L(u, x) &= C(x)u(x) + B(x) \cdot \nabla u(x) + \text{tr}(A(x)D^2u(x)) \\ &\quad + \int_{\mathbb{R}^d} u(x+y) - u(x) - \chi_{B_1(0)}(y) \nabla u(x) \cdot y \mu(x, dy). \end{aligned}$$

Moreover,  $A, B$ , and  $C$  each have modulus of continuity  $C\omega(2(\cdot))$ , and for every  $r > 0$  and any  $x_1, x_2 \in \mathbb{R}^d$  we have

$$\|\mu_L(x_1) - \mu_L(x_2)\|_{\text{TV}(CB_r)} \leq C(r)\omega(2|x_1 - x_2|).$$

If  $\beta < 2$ , then  $A \equiv 0$ , while if  $\beta < 1$  then  $B \equiv 0$  and the integrand with respect to  $\mu(x, dy)$  in the formula above is replaced with  $u(x+y) - u(x)$ .

**Proof.** Take a decreasing sequence  $\delta_k$  such that  $\delta_k \rightarrow 0$ , and let us take the functions  $\phi_{\delta_k}$  and  $\eta_{\delta_k}$ , as defined in [\(5.13\)](#). Then for each  $k$ ,  $L$  has the representation

$$\begin{aligned} L(u, x) &= C_L(x)u(x) + B_{L, \phi_{\delta_k}}(x) \cdot \nabla u(x) + \text{tr}(A_{L, \eta_{\delta_k}}(x)D^2u(x)) \\ &\quad + \int_{\mathbb{R}^d} u(x+y) - P_{\phi_{\delta_k}, \eta_{\delta_k}, u, x}(x+y) \mu(x, dy), \end{aligned}$$

where  $A_{L, \eta_{\delta_k}}$ ,  $B_{L, \phi_{\delta_k}}$ , and  $C_L$  are as in [Lemma 5.17](#). Now,  $L$  satisfies the estimate

$$|L(\tau_{-x_1}(\eta_{\delta_k})_{ij}, x_1) - L(\tau_{-x_2}(\eta_{\delta_k})_{ij}, x_2)| \leq \alpha(1, \eta_{\delta_k})\omega(2|x_1 - x_2|)$$

Thanks to (5.14), it follows that  $\alpha(1, \eta_{\delta_k}) \leq C$  for all  $k$ . It follows that  $\{A_{L, \eta_{\delta_k}}\}_k$  has a uniform modulus of continuity. The same argument yields a modulus of continuity for  $\{B_{L, \phi_{\delta_k}}\}_k$  and for the function  $C(x)$ , all given by  $C\omega(2|x_1 - x_2|)$ , with  $C$  independent of  $k$  and  $\omega$  being the modulus from Assumption 1.4. This equicontinuity means these sequences of functions are pre-compact at least when restricted to any compact subset of  $\mathbb{R}^d$ , by the Arzela–Ascoli theorem. Therefore, after a Cantor diagonalization argument we see that along some subsequence  $m_k \rightarrow \infty$  these functions converge locally uniformly in  $\mathbb{R}^d$  to functions  $A(x)$ ,  $B(x)$ , respectively. Of course, the functions  $A$ ,  $B$ , and  $C$  all inherit the modulus of continuity  $C\omega(2(\cdot))$ . The respective TV-norm continuity estimate for  $\mu_L$  follows by applying Proposition 5.13 and passing to the limit (always recalling that,  $\mathcal{D}_I$  is the convex hull of such limit points).

With the convergence established, we have

$$\lim_{k \rightarrow \infty} (B_{L, \phi_{\delta_{m_k}}}(x) \cdot \nabla u(x) + \text{tr}(A_{L, \eta_{\delta_{m_k}}}(x) D^2 u(x))) = B(x) \cdot \nabla u(x) + \text{tr}(A(x) D^2 u(x)),$$

and so, for every  $u$  we have the formula

$$\begin{aligned} L(u, x) &= C(x)u(x) + B(x) \cdot \nabla u(x) + \text{tr}(A(x) D^2 u(x)) \\ &\quad + \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} u(x+y) - P_{\phi_{\delta_k}, \eta_{\delta_k}, u, x}(x+y) \mu(x, dy), \end{aligned}$$

It remains to compute the limit of the integral, observe that

$$\int_{\mathbb{R}^d} \eta_{\delta_k}(y) (D^2 u(x)y, y) \mu(x, dy) = \int_{B_{\delta_k}} \eta_{\delta_k}(y) (D^2 u(x)y, y) \mu(x, dy),$$

which means that

$$\left| \int_{\mathbb{R}^d} \eta_{\delta_k}(y) (D^2 u(x)y, y) \mu(x, dy) \right| \leq C |D^2 u(x)| \int_{B_{\delta_k}} |y|^2 d\mu(x, dy).$$

Therefore,

$$\lim_{k \rightarrow 0} \int_{\mathbb{R}^d} \eta_{\delta_k}(y) (D^2 u(x)y, y) \mu(x, dy) = 0.$$

On the other hand, for every  $y$  we have

$$\lim_{k \rightarrow \infty} (u(x+y) - P_{\phi_{\delta_k}, \eta_{\delta_k}, u, x}(y)) = u(x+y) - u(x) - \chi_{B_1}(y) \nabla u(x) \cdot y,$$

and the limit is monotone. Therefore, by monotone convergence we conclude that

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} u(x+y) - P_{\phi_{\delta_k}, \eta_{\delta_k}, u, x}(y) \mu(x, dy) = \int_{\mathbb{R}^d} u(x+y) - u(x) - \chi_{B_1}(y) \nabla u(x) \cdot y \mu(x, dy).$$

and with this the Corollary is proved.  $\square$

#### 5.4. Limits of $I_n$

**Lemma 5.19.** *Assume that  $I : C_b^\beta(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$  is Lipschitz. Let  $K > 0$  and  $0 < \beta < \beta_0 < 3$ . If  $u \in C_b^{\beta_0}(\mathbb{R}^d)$  is supported in  $B_K$ , and  $2^{n-2} \geq K$ , then*

$$\|I_n u - Iu\|_{L^\infty(B_K \cap G_n)} \leq C 2^{-n\gamma} \|u\|_{C^{\beta_0}(\mathbb{R}^d)},$$

for a universal constant  $C$  and  $\gamma = \gamma(\beta_0, \beta) \in (0, 1)$ . Furthermore, we have

$$\lim_{n \rightarrow \infty} \|I(u) - I_n(u)\|_{L^\infty(B_K)} = 0.$$

**Proof.** Let  $u$  be compactly supported in  $B_K$ , and be such that  $\|u\|_{C^{\beta_0}} \leq M$ . First, note that since  $2^{n-2} \geq K$ , then we have

$$\hat{\pi}_n^\beta u = \pi_n^\beta u,$$

thus,  $I_n(u) = \hat{\pi}_n^0 \circ I \circ \pi_n^\beta(u)$ . Keeping this in mind, using the Lipschitz property of  $I$ , we have

$$\|I(u) - I(\hat{\pi}_n^\beta u)\|_{L^\infty(\mathbb{R}^d)} \leq C\|u - \pi_n^\beta u\|_{C^\beta(\mathbb{R}^d)}.$$

Since  $2^{n-2} \geq K$  we have that  $I(\hat{\pi}_n^\beta u) = \hat{\pi}_n^0 I(\hat{\pi}_n^\beta u) = I_n(u)$  when restricted to  $B_K \cap G_n$ , which thanks to [Lemma 4.21](#) implies the first estimate. Next, [Theorem 4.13](#) guarantees that

$$\|\hat{\pi}_n^0 I(u) - \hat{\pi}_n^0 I(\hat{\pi}_n^\beta u)\|_{L^\infty(K)} \leq C\|I(u) - I(\hat{\pi}_n^\beta u)\|_{L^\infty(\mathbb{R}^d)} \leq C\|u - \pi_n^\beta u\|_{L^\infty(\mathbb{R}^d)}.$$

Thus,

$$\begin{aligned} \|I_n(u) - I(u)\|_{L^\infty(K)} &\leq \|\hat{\pi}_n^0 I(u) - I_n(u)\|_{L^\infty(K)} + \|\hat{\pi}_n^0(I(u)) - I(u)\|_{L^\infty(K)} \\ &\leq C\|u - \pi_n^\beta u\|_{C^\beta(\mathbb{R}^d)} + \|\hat{\pi}_n^0(I(u)) - I(u)\|_{L^\infty(K)}. \end{aligned}$$

Applying [Lemma 4.21](#) to the first term and [Remark 4.22](#) to the second, we conclude that

$$\lim_{n \rightarrow \infty} \|I_n u - I u\|_{L^\infty(K)} = 0. \quad \square$$

**Corollary 5.20.** *Assume  $I$  satisfies [Assumptions 1.1, 1.3, and 1.4](#), as stated for  $C_b^\beta(\mathbb{R}^d)$ . Then for every  $u \in C_b^\beta(\mathbb{R}^d)$  and every  $R > 0$ ,*

$$\lim_{n \rightarrow \infty} \|I_n u - I u\|_{L^\infty(B_R)} = 0.$$

**Proof.** Fix  $u \in C_b^\beta(\mathbb{R}^d)$  and  $R, \varepsilon > 0$ . For  $K > 0$  (to be determined later), we may decompose  $u$  as  $u = u_0 + u_1$ , where  $u_0$  is compactly supported in  $B_{2K+1}$  and  $u_1 \equiv 0$  in  $B_{2K}$ , all such that

$$\|u_i\|_{C^\beta(\mathbb{R}^d)} \leq C\|u\|_{C^\beta(\mathbb{R}^d)}, \quad i = 1, 2.$$

The constant  $C > 1$  being independent of  $K$ . Now, by [Assumption 1.3](#) and since  $u \equiv u_0$  in  $B_{2K}$ , we have

$$|I(u_0) - I(u)| \leq \rho(K)\|u - u_0\|_{L^\infty(\mathbb{R}^d)} \leq 2C\rho(K)\|u\|_{C^\beta(\mathbb{R}^d)}.$$

Choose  $K$  large enough so that  $K \geq 2R$  and  $2C\rho(K)\|u\|_{C^\beta(\mathbb{R}^d)} \leq \varepsilon/2$ . Then, with this  $K$ , we apply [Lemma 5.19](#) two times, and conclude that there is some  $n_0 > 0$  such that

$$|I_n(u_0) - I(u_0)| + |I_n(u_0) - I_n(u)| \leq \varepsilon/2 \text{ whenever } n \geq n_0.$$

On the other hand, in all  $\mathbb{R}^d$  we have the pointwise inequality,

$$|I_n(u) - I(u)| \leq |I_n(u_0) - I(u_0)| + |I_n(u_0) - I_n(u)| + |I(u_0) - I(u)|,$$

and it follows that, for  $x \in B_R$  and  $n \geq n_0$ , that

$$|I_n(u, x) - I(u, x)| \leq \varepsilon,$$

and the corollary is proved.  $\square$

### 5.5. Proofs of Theorems 1.11 and 1.14

We conclude this section with the proofs of the remaining theorems.

**Proof of Theorem 1.11.** Consider the set  $\mathcal{D}_I$ . The proof will boil down to showing that for any  $u, v \in C_c^{\beta_0}(\mathbb{R}^d)$  and any  $x \in \mathbb{R}^d$  there is some  $L \in \mathcal{D}_I$  such that

$$I(u, x) \leq I(v, x) + L(u - v, x).$$

Fix  $u, v$  and  $x$ . Then, by Remark 5.4, for every  $n$  we have

$$I_n(u, x) \leq \max_{L_n \in \mathcal{D}I_n} \{I_n(v, x) + L_n(u - v, x)\}.$$

In particular, for every  $n$ , there is some  $L_n \in \mathcal{D}I_n$  such that (with this same  $u, v$  and  $x$ )

$$I_n(u, x) \leq I_n(v, x) + L_n(u - v, x).$$

Let us obtain an inequality as we let  $n \rightarrow \infty$  along some subsequence. Thanks to Corollary 5.20, for every  $x \in \mathbb{R}^d$  we have

$$\lim_{n \rightarrow \infty} I_n(u, x) = I(u, x), \quad \lim_{n \rightarrow \infty} I_n(v, x) = I(v, x).$$

On the other hand, Lemma 5.16 says there is a subsequence  $n_k$  and an operator  $L$  such that  $L_{n_k}(u - v, x)$  converges to  $L(u - v, x)$ , and moreover  $L \in \mathcal{D}_I$ , by the definition of  $\mathcal{D}_I$ . Then, we conclude that

$$I(u, x) \leq I(v, x) + L(u - v, x) \leq \sup_{L \in \mathcal{D}_I} \{I(v, x) + L(u - v, x)\}.$$

The above holds for any pair of functions  $u$  and  $v$  and any point  $x \in \mathbb{R}^d$ . Taking the minimum over all  $v$ , we obtain for any  $u$  and  $x$ ,

$$I(u, x) = \min_{v \in C_b^{\beta}(\mathbb{R}^d)} \max_{L \in \mathcal{D}_I} \{I(v, x) - L(v, x) + L(u, x)\}.$$

Using  $v \in C_b^{\beta}(\mathbb{R}^d)$  and  $L \in \mathcal{D}_I$  as the set of labels, which we rename  $ab$ , and letting  $f_{ab}(x)$  correspond to the functions  $I(v, x) - L(v, x)$ , we obtain the desired min–max representation.

The  $L^\infty$  bounds for the coefficients follow from the construction of  $A_{\eta_k}$ , etc... in (5.8), (5.9), (5.10). The continuity of the coefficients and the Lévy measures follows from Lemma 5.16.  $\square$

**Proof of Theorem 1.14.** For the versions of Theorems 1.9 and 1.10 with  $\beta < 2$  we apply the last part of Lemma 3.9 to conclude the functionals (or translation invariant operators) appearing in the min–max all have the corresponding simpler form. As for Theorem 1.11, we use instead the last part of Corollary 5.18 to obtain the simpler expression for the Lévy operators in the cases where  $\beta < 2$ .  $\square$

## 6. Some examples

In this section we list some examples to which our results apply, yet the integro-differential structure given in either (1.2) or (1.3) is not readily apparent from the definition of the operator itself. We emphasize that most cases of the *linear* examples that we list were already contained in the classic work of Courrèges [19], but we include them here for the sake of illustration. In all of these examples, the operators satisfy the GCP and the other technical requirements to apply the results presented above. We do not intend to give all details, but rather just make a list, with some appropriate references. At the end of the section, we list how these examples relate to Assumptions 1.1–1.4.

### 6.1. The statement of the examples.

**Example 6.1.** The generator of a Markov process. Assume that  $X_t$  is a Markov process taking values in  $\mathbb{R}^d$ , and that  $\mathbb{E}_x$  is the expectation of the process, having started from  $x$  at  $t = 0$ . The generator is defined as the operator

$$L(u, x) = \lim_{t \rightarrow 0} \frac{\mathbb{E}(u(X_t)) - \mathbb{E}(u(X_0))}{t},$$

over all  $u$  for which the limit exists. (See Liggett [42, Chapter 3].)

Thanks to the fact that  $\mathbb{E}$  preserves ordering, one can immediately see that  $L$  enjoys the GCP. When  $X_t$  is such that  $L : C_b^2 \rightarrow C_b^2$ , this example is covered by Courrègue [19]; but if  $X_t$  is such that  $L : C_b^\beta \rightarrow C_b^0$  (in a Lipschitz fashion) for some  $0 < \beta < 2$ , then by Theorem 1.14, there are fewer terms (see the list just above Theorem 1.14 for our use of the notation  $C_b^\beta(\mathbb{R}^d)$ ). In this context, the result of Courrègue can be seen as a version of the Lévy–Khintchine formula for a process whose increments need not be stationary.

**Example 6.2.** The Dirichlet-to-Neumann map for linear, elliptic operators on half-space. Assume that  $L$  is an operator that admits unique bounded solutions on  $\mathbb{R}_+^{d+1}$  and that has a comparison principle. What we mean by this is the following: we can take  $u \in C_b^{1,\alpha}(\mathbb{R}^d)$  and associate to it the unique bounded solution,  $U_u$  of

$$L(U_u, X) = 0 \text{ in } \mathbb{R}_+^{d+1}, \text{ and } U_u = u \text{ on } \mathbb{R}^d \times 0.$$

A couple of reasonable examples would be

$$L(U, X) = \text{tr}(A(X)D^2U(X)) \text{ or } L(U, X) = \text{div}(A(X)\nabla U),$$

where  $A$  is uniformly elliptic and Hölder continuous. The Dirichlet-to-Neumann map is then defined as

$$I(u, x) := \partial_n U_u(x).$$

First of all, the assumptions on  $A$  are such that for some  $\alpha'$ ,  $U_u \in C_b^{1,\alpha'}(\overline{\mathbb{R}_+^{d+1}})$  and hence the normal derivative is well defined (see, e.g. [23, Chapters 8, 9]). It is not hard to check that this operator satisfies the GCP, and this fact comes entirely from the property that the solution operator, by the assumed comparison principle, preserves ordering of solutions whenever the boundary data are ordered (it has nothing to do with linearity of the solution operator). This is, again, within the context of Courrègue's result, but we can invoke Theorem 1.14 to remove extra terms of order higher than 1. Ellipticity and scaling show that this is always an operator of order 1 (and will map  $C^{1,\alpha} \rightarrow C^{\alpha'}$ ). We note that in this example, via linear equations with nice coefficients, one can derive lots of information about the operator  $\partial_n U_u$  by directly using the Poisson kernel that represents the solution  $U_u$ .

In the context of periodic equations, one can use the results in Sections 4 and 5 to show that the coefficients in the resulting Lévy operators will share the same periodicity. In fact, this is very straightforward if  $I$  is linear. If instead one looks at almost periodic coefficients, it seems reasonable to hope that the coefficients will also be almost periodic, but we have not checked this claim. If it is the case, there could be an application to some boundary homogenization problems with irrationally oriented half-spaces inside a periodic medium, related to [31]. Operators related to the Dirichlet-to-Neumann mapping of this example are also of interest in conformal geometry, see Chang–Gonzalez [13]. It is also possible to consider an elliptic equation with weights in order to obtain some operators of order different than 1, e.g. Caffarelli–Silvestre [9].

**Example 6.3.** The boundary process of a reflected diffusion. (See Hsu [32], or [33, Chp. IV, Sec. 7] and/or [45, Sec. 8].)

In this context, one starts with a diffusion in  $\mathbb{R}_+^{d+1}$ , say  $X_t$ , so that  $X_t$  reflects off of the bottom boundary whenever it reaches it. Under a time rescaling of  $X_t$  (because it spends zero time on the boundary), the resulting process can be viewed at times only when it hits  $\mathbb{R}^d \times \{0\}$ , and induces a pure jump process on  $\mathbb{R}^d \times \{0\}$ . This process is generated by an operator of the form (1.2) with  $A \equiv 0$ . It turns out that this generator for the boundary process is exactly the Dirichlet-to-Neumann mapping from the previous example. This process was studied in a smooth domain for Brownian motion by Hsu [32].

**Example 6.4.** Subordinated diffusions and Bernstein functions. (See Schilling–Song–Vondraček [46].)

The time-rescaling of the reflected diffusion in the previous example is just one choice of a rescaling, and in general one can time-rescale diffusions on  $\mathbb{R}^d$  (so no boundary space here) in a myriad of fashions to create new stochastic processes from one reference Brownian motion. This is a process known as subordination, and it can be used to create operators with generators in the class (1.2), starting with one that may simply only contain the second order term. The generator for the subordinated process will enjoy the GCP because the generator of the original diffusion also enjoys the GCP. This technique has played a large and fundamental role in the study of Lévy processes, and one can see it in use in e.g., the book of Schilling–Song–Vondraček [46], especially [46, Chapter 13]. The subordination formula is closely related to an extension into plus one space variables, and this extension was used to create operators of fractional order that enjoy the GCP in the work of Stinga–Torrea [55] and also provide other properties of the fractional operators.

**Example 6.5.** The Monge–Ampère operator,  $\text{MA}(u, x) = \det(D^2u)$ .

When one restricts this operator to the subset of  $C^2$  of convex functions, then  $\text{MA}$  is in fact (degenerate) elliptic and locally Lipschitz. Specifically for each  $\delta > 0$ ,  $\text{MA}$  is uniformly elliptic (depending upon  $\delta$ ), Lipschitz, and translation invariant as a mapping,

$$\text{MA} : \{u \in C_b^2(\mathbb{R}^d) : \frac{1}{\delta} > D^2u > \delta\} \rightarrow C_b^0(\mathbb{R}^d).$$

Thus,  $\text{MA}$ , must enjoy a min–max structure. Experts have known and utilized this min–max property of  $\text{MA}$  in the study of fully nonlinear elliptic equations for a long time, and one can show that

$$(\text{MA}(u, x))^{1/d} = \frac{1}{d} \inf\{\text{tr}(AD^2u(x)) : A \geq 0, \text{ and } \det(A) = 1\}.$$

In fact, this formula is intimately connected with various investigations into nonlocal operators that should be an analog of  $\text{MA}$  in the fractional setting (as of yet, there is not one that is considered better than others). Some works that address nonlocal analogs of  $\text{MA}$  are: [8, 12], and [29].

**Example 6.6.** General nonlocal operators as treated in Caffarelli–Silvestre [10, 11]. These are simply operators that are assumed to satisfy the GCP, are defined for all functions in  $C^{1,1}(\mathbb{R}^d)$ , map  $C_b^2(\mathbb{R}^d) \rightarrow C_b^0(\mathbb{R}^d)$ , and satisfy a form of uniform ellipticity that is given by the existence of concave respectively convex operators,  $\mathcal{M}_{\mathcal{L}}^-$  and  $\mathcal{M}_{\mathcal{L}}^+$  so that

$$\text{for all } u, v \in C^{1,1}(\mathbb{R}^d), \quad \mathcal{M}_{\mathcal{L}}^-(u - v, x) \leq I(u, x) - I(v, x) \leq \mathcal{M}_{\mathcal{L}}^+(u - v, x). \quad (6.1)$$

Here,  $\mathcal{L}$  is a class of linear operators that is usually a particular subset of those that satisfy the Lévy type condition (1.2).

This context for nonlocal operators was given in [10, Definition 3.1], and it played an important role in many of the results — especially when  $\mathcal{L}$  is chosen to contain certain classes of operators. These operators, in

cases in which they are Lipschitz fall into the scope of our results, and furthermore, the role of the extremal operators gives extra information about the min–max formula. In particular, as shown in [28, Section 4.6], when ellipticity occurs with respect to  $\mathcal{M}_{\mathcal{L}}^{\pm}$ , then the min–max may be restricted to only utilize linear functionals (or linear operators) that also satisfy the extremal inequality in (6.1). This also appeared in a homogenization result by one of the authors in which they were unable to show that the limit operator had an explicit integro-differential formula, but rather was only integro-differential and uniformly elliptic in the sense of [10, Definition 3.1] (see the homogenization in [47]).

**Example 6.7.** The Dirichlet to Neumann map for fully nonlinear elliptic equations. In Example 6.2, the linearity of  $L$  is not necessary, and the function  $U_u$  can also be taken to solve a fully nonlinear, uniformly elliptic equation in  $\mathbb{R}_+^{d+1}$ . These equations always possess a comparison principle (by definition), and under most reasonable assumptions, the solution  $U_u$  will be globally  $C^{1,\alpha'}$ , allowing for the normal derivative to be defined classically (see [52] for this regularity).

This was a main topic in the recent paper by the authors and Kitagawa [26]. It turns out that the extremal operators (as in Example 6.6) for the nonlinear D-to-N not only play a crucial role in investigating the Lévy measures in the min–max, but they also take a refreshingly simple form. The extremal operators in this case,  $\mathcal{M}_{\mathcal{L}}^{\pm}$  of Example 6.6, are simply the Dirichlet-to-Neumann operators for the solutions of the corresponding extremal operators for the elliptic second order equation in  $\mathbb{R}_+^{d+1}$ . These are usually called the Pucci extremal operators (see [7]), and solutions to their equations are generally very well behaved. In [26], the properties of the Lévy measures in the min–max are linked to the harmonic measures for linear equations with bounded measurable coefficients (e.g. [39]), but there is still more to learn about them before they can be connected with existing integro-differential theory.

**Example 6.8.** An operator that drives surface evolution in one and two phase free boundary problems related to a type of Hele-Shaw flow. Given  $f \in C^{1,\alpha}(\mathbb{R}^d)$ , such that  $0 < \inf f \leq \sup f < \infty$ , we can define the unique solution,  $U_f$ , of the elliptic equation,

$$\begin{aligned} \Delta U_f &= 0 \text{ in } \{(x, x_{d+1}) : 0 < x_{d+1} < f(x)\}, \\ U_f &= 1 \text{ on } \{x_{d+1} = 0\}, \quad U_f = 0 \text{ on } \{(x, x_{d+1}) : x_{d+1} = f(x)\}. \end{aligned}$$

This allows to define a (fully nonlinear) operator on  $f$  as

$$I(f, x) := \partial_n U_f(x, f(x)),$$

that is, the normal derivative of the solution on the upper boundary given by the graph of  $f$ .

For Hele-Shaw flow in the simplified setting that the free boundary is parametrized by the graph of  $f(\cdot, t)$ , it can be shown that the free boundary evolves by a normal velocity that at each time is given by  $I(f, x)$ . The interpretation here is that fluid flows into the domain under a pressure at the bottom boundary,  $x_{d+1} = 0$ , and the top edge of the fluid exists at  $x_{d+1} = f(x)$ , with  $U_f$  representing the pressure of the fluid. This pressure induces a force on the fluid, which is given by  $\partial_n U_f(x, f(x))$  at the top boundary. This operator, and its implications for rewriting a class of free boundary problems that are similar to Hele-Shaw were studied by the authors and Chang Lara in [16]. In particular, the min–max formula makes it straightforward to convert the free boundary flow into a nonlocal parabolic equation for  $f$ , and this parabolic equation is very similar to ones that have already been studied in the nonlocal literature (e.g. [51]). When  $U_f$  is defined to be harmonic in the domain determined by  $f$ , standard regularity theory immediately gives estimates that show there is some  $\alpha'$  so that the mapping from  $f$  to  $I(f)$  is Lipschitz from  $C^{1,\alpha}(\mathbb{R}^d)$  to  $C^{\alpha'}(\mathbb{R}^d)$ . In [16] it

was also shown that the same Lipschitz property can be obtained when  $U_f$  is defined as the solution of a nonlinear uniformly elliptic second order equation instead of just the Laplacian. This operator gives a good example of what can be said in the translation invariant case of the min–max, and its properties are studied initially in [16]. Even in the simplest case of defining  $U_f$  to be harmonic, the resulting operator  $I$  will always be inherently nonlinear and nonlocal.

## 6.2. Relationship to Assumptions 1.1–1.4

Here we list how each of the above examples fits within the context of [Assumptions 1.1–1.4](#).

([Example 6.1](#)). By construction, this  $L$  is always linear. Thus, [Assumption 1.1](#) follows from simply saying that  $L$  is a bounded operator on  $C^\beta$ , which of course requires assumptions on the process,  $X_t$ , or more specifically the transition probability measure for  $X_t$ . Again, via linearity, [Assumption 1.2](#) follows whenever the process,  $X_t$ , has stationary and independent increments. [Assumptions 1.3](#) and [1.4](#) will be an extra requirement on the transition probability measure for  $X_t$ . In particular (although a bit circular), [Assumption 1.4](#), in view of linearity, is equivalent to the martingale problem for  $X_t$  having a solution and the generator having uniformly continuous coefficients.

([Example 6.2](#)). (The interested reader can see [26] for more details.) [Assumption 1.1](#) holds for  $C^{1,\alpha} \rightarrow C^\alpha'$  when  $A$  is  $\alpha$ -Hölder continuous. [Assumption 1.2](#) holds if  $A$  is a constant. [Assumption 1.3](#) holds in both of the above settings, by using a barrier argument (which is easier implemented for the non-divergence equation). Since  $I$  is linear, [Assumption 1.4](#) holds when  $A$  is Hölder continuous. Indeed, by linearity, checking [Assumption 1.4](#) is equivalent to estimating

$$I(\tau_{-z}u, x+z) - I(u, x).$$

In the case of divergence equations, one can write down the equations satisfied for  $V = \tau_{-z}U_u$ , and then also the equation satisfied by  $W := U_{\tau_{-z}u} - V$ . The desired estimate is then equivalent to estimating  $|\partial_n W(x+z)|$ , i.e. a global Lipschitz estimate for  $W$ . Since  $W$  satisfies

$$\operatorname{div}(A(X)\nabla W(X)) = -\operatorname{div}((A(X) - A(x-z))\nabla V),$$

we see that by global Lipschitz estimates,

$$|\nabla W| \leq C\|(A - A(\cdot - z))\nabla V\|_{L^\infty} \leq C|z|^\alpha,$$

by the original assumption that  $A$  is Hölder continuous. (Note, the Lipschitz estimates here are a standard modification to, e.g. [25, Lemma 3.2] to allow for a right hand side of the form  $\operatorname{div}(f)$  with  $f \in L^\infty$ .)

([Example 6.3](#)). In most reasonable situations in which the diffusion has regular coefficients, this is contained in the previous example.

([Example 6.4](#)). This, of course, depends heavily on the original Markov process and the choice of subordinator. However, one of the most classical situations starts with a Brownian motion and then uses a Lévy stable subordinator. In this case, the resulting operator is translation invariant, and [Assumptions 1.1](#) and [1.2](#) follow more or less by construction.

([Example 6.5](#)). This is a translation invariant operator, and as mentioned already satisfies the Lipschitz property on the specified convex subsets of  $C^2$ . So, [Assumptions 1.1](#) and [1.2](#) hold.

([Example 6.6](#)). As this is a general example, the operators only satisfy the given assumptions when explicitly required to do so. However, the interesting part of this example arises from the fact that the

knowledge of the extremal inequalities in (6.1) in fact gives more detailed information about the linear operators that will appear in the min–max of [Theorems 1.9–1.14](#). This is discussed in [28, Section 4.6].

([Example 6.7](#)). This operator satisfies [Assumption 1.1](#) as a mapping of  $C^{1,\alpha} \rightarrow C^{\alpha'}$  (for some  $0 < \alpha' < \alpha$ ) under standard assumptions about  $F$ . The relevant regularity theory comes from Silvestre–Sirakov [52]. It can also be checked by using the same type of barrier argument that works for [Example 6.2](#) will show [Assumption 1.3](#) is also satisfied. Due to the nonlinear nature of the D-to-N in this setting, it is not obvious how to show that [Assumption 1.4](#) is satisfied — we do not know if it is satisfied or not. Thus, the best one can say about this operator when it is not translation invariant is the outcome of [Theorem 1.9](#). We simply note to the interested reader that because of the lack of exact cancellation from the fact that the mapping is not linear, one probably needs more detailed information about  $F$ . Indeed, using the extremal operators would not help because it would produce

$$\begin{aligned} I(v + \tau_{-z}u, x + z) - I(v, x + z) - (I(v + u, x) - I(v, x)) &\leq M^+(\tau_{-z}u, x + z) - M^-(u, x) \\ &= M^+(u, x) - M^-(u, x). \end{aligned}$$

Here we use  $M^\pm$  as the extremal operators for  $I$ , and also that these are translation invariant. This estimate completely neglects the influence of the shift,  $\tau_z$ , and so it would not be useful (furthermore, one expects that  $M^+(u, x) > M^-(u, x)$ ).

([Example 6.8](#)). As it is stated above, this operator,  $I$ , is actually translation invariant, and so it is straightforward to check that [Assumptions 1.1](#) and [1.2](#) hold. In the case that the equation for  $U$  (i.e.  $\Delta U = 0$ ) is replaced by either a fully nonlinear operator and/or operator that is not translation invariant, it is harder to check all of the applicable assumptions. Again, for fully nonlinear equations that define  $U$ , in [16]  $I$  was checked to be Lipschitz as a map of  $C^{1,\alpha} \rightarrow C^{\alpha'}$  (which took a reasonably non-trivial amount of work).

## Appendix. Additional proofs and computations

**Proof of [Proposition 4.10](#).** Fix  $u \in C_b^\beta(\mathbb{R}^d)$ , and let  $x \in G_n$ , then by the regularity of  $u$ ,

$$|u(x \pm h_n e_k) - (u(x) \pm h_n \nabla u(x_0) \cdot e_k)| \leq C \|u\|_{C^\beta} h_n^{\min\{\beta-1, 1\}}.$$

Therefore,

$$|u(x + h_n e_k) - u(x + h_n e_k) - 2h_n \nabla u(x_0) \cdot e_k| \leq C \|u\|_{C^\beta} h_n^{\min\{\beta-1, 1\}}$$

For the second estimate, we shall make use of

$$|u(x + h_n e_k) - (u(x) + h_n \nabla u(x_0) \cdot e + h_n^2 \frac{1}{2} (D^2 u(x) e, e))| \leq C \|u\|_{C^\beta} h_n^{\min\{\beta-2, 1\}}.$$

Therefore,

$$\begin{aligned} &u(x + h_n e_k + h_n e_\ell) - u(x + h_n e_k) - u(x + h_n e_\ell) + u(x) \\ &= u(x) + h_n \nabla u(x_0) \cdot (e_k + e_\ell) + h_n^2 \frac{1}{2} (D^2 u(x)(e_k + e_\ell, e_k + e_\ell)) \\ &\quad - (u(x) + h_n \nabla u(x_0) \cdot e_k + h_n^2 \frac{1}{2} (D^2 u(x)e_k, e_k)) \\ &\quad - (u(x) + h_n \nabla u(x_0) \cdot e_\ell + h_n^2 \frac{1}{2} (D^2 u(x)e_\ell, e_\ell)) + u(x) \\ &= h_n^2 \frac{1}{2} ((D^2 u(x)(e_k + e_\ell, e_k + e_\ell)) - (D^2 u(x)e_k, e_k) - (D^2 u(x)e_\ell, e_\ell)) \\ &= h_n^2 (D^2 u(x)e_k, e_\ell) \end{aligned}$$

It follows that

$$|u(x + h_n e_k + h_n e_\ell) - u(x + h_n e_k) - u(x + h_n e_\ell) + u(x) - h_n^2 (D^2 u(x)e_k, e_\ell)| \leq C \|u\|_{C^\beta} h_n^{\min\{\beta-2, 1\}},$$

and the proposition is proved.  $\square$

**Proof of Proposition 4.11.** Fix  $u \in C_b^\beta(\mathbb{R}^d)$ .

Step 1. Let  $x \in G_n$ , then

$$\begin{aligned} |(\nabla_n)^1 u(x) - \nabla u(x)| &\leq C \|u\|_{C^\beta} h_n^{\beta-1}, \quad \text{if } \beta \in [1, 2], \\ |(\nabla_n)^2 u(x) - D^2 u(x)| &\leq C \|u\|_{C^\beta} h_n^{\beta-2}, \quad \text{if } \beta \in [2, 3]. \end{aligned}$$

*Proof of Step 1.* By the regularity of  $u$ ,

$$|u(x \pm h_n e_k) - (u(x) \pm h_n \nabla u(x_0) \cdot e_k)| \leq C \|u\|_{C^\beta} h_n^{\min\{\beta-1, 1\}}.$$

Therefore,

$$|u(x + h_n e_k) - u(x + h_n e_k) - 2h_n \nabla u(x_0) \cdot e_k| \leq C \|u\|_{C^\beta} h_n^{\min\{\beta-1, 1\}}$$

Step 2. Given  $x \in G_n$ , we have

$$|(\nabla_n)^1 u(x)| \leq C \|u\|_{C^1}, \quad |(\nabla_n)^2 u(x)| \leq C \|u\|_{C^2}.$$

Step 3.

$$\begin{aligned} |(\nabla_n)^1 u(\hat{x}) - (\nabla_n)^1 u(\hat{y})| &\leq C \|u\|_{C^\beta} d(\hat{x}, \hat{y})^{\beta-1}, \quad \text{if } \beta \in [1, 2], \\ |(\nabla_n)^2 u(\hat{x}) - (\nabla_n)^2 u(\hat{y})| &\leq C \|u\|_{C^\beta} d(\hat{x}, \hat{y})^{\beta-2}, \quad \text{if } \beta \in [2, 3]. \quad \square \end{aligned}$$

Computation for Lemma 4.17

$$\nabla \tilde{R}(x) = 2C \|w\|_{C^{\beta_0}} \eta' \left( \frac{|x - x_0|^{\beta_0}}{h_n} \right) \beta_0 |x - x_0|^{\beta_0-1} \frac{(x - x_0)}{|x - x_0|}$$

If  $|x - x_0|^{\beta_0} \leq h_n$ , then

$$\nabla \tilde{R}(x) = 2C \|w\|_{C^{\beta_0}} \beta_0 |x - x_0|^{\beta_0-1} \frac{(x - x_0)}{|x - x_0|}$$

This expression is zero except when  $|x - x_0| \leq h_n^{1/\beta_0}$ , so

$$|\nabla \tilde{R}(x)| \leq 2C \|w\|_{C^{\beta_0}} \beta_0 h_n^{1-1/\beta_0}.$$

Furthermore, for  $x, x'$  such that  $|x - x_0|^{\beta_0} \leq h_n$ , we have

$$\begin{aligned} |\nabla \tilde{R}(x) - \nabla \tilde{R}(x')| &\leq 2C \beta_0 \|w\|_{C_0^\beta} \left| |x - x_0|^{\beta_0-1} \frac{(x - x_0)}{|x - x_0|} - |x' - x_0|^{\beta_0-1} \frac{(x' - x_0)}{|x' - x_0|} \right| \\ &\leq C \|w\|_{C^{\beta_0}} h_n^{\beta_0-\beta} |x - x'|^\beta. \end{aligned}$$

In conclusion,

$$\|\tilde{R}\|_{L^\infty} + \|\nabla \tilde{R}\|_{L^\infty} + [\nabla \tilde{R}]_{C^{\beta-1}} \leq C \|w\|_{C^{\beta_0}} (h_n + h_n^{1-1/\beta_0} + h_n^{\beta_0-\beta}) \leq C \|w\|_{C^{\beta_0}} h_n^\gamma. \quad \square$$

## References

- [1] Luis Alvarez, Frédéric Guichard, Pierre-Louis Lions, Jean-Michel Morel, Axioms and fundamental equations of image processing, *Arch. Ration. Mech. Anal.* 123 (3) (1993) 199–257.
- [2] Guy Barles, Emmanuel Chasseigne, Adina Ciomaga, Cyril Imbert, Lipschitz regularity of solutions for mixed integro-differential equations, *J. Differential Equations* 252 (11) (2012) 6012–6060.
- [3] G. Barles, E. Chasseigne, C. Imbert, On the Dirichlet problem for second-order elliptic integro-differential equations, *Indiana Univ. Math. J.* 57 (1) (2008) 213–246.
- [4] G. Barles, E. Chasseigne, C. Imbert, Hölder continuity of solutions of second-order elliptic integro-differential equations, *J. Eur. Math. Soc.* 13 (1) (2011) 1–26.

- [5] Guy Barles, Cyril Imbert, Second-order elliptic integro-differential equations: viscosity solutions' theory revisited, *Ann. Inst. H. Poincaré Anal. Non Linéaire* 25 (3) (2008) 567–585.
- [6] Guy Barles, Panagiotis E. Souganidis, A new approach to front propagation problems: theory and applications, *Arch. Ration. Mech. Anal.* 141 (3) (1998) 237–296.
- [7] Luis A. Caffarelli, Xavier Cabré, Fully Nonlinear Elliptic Equations, in: American Mathematical Society Colloquium Publications, vol. 43, American Mathematical Society, Providence, RI, 1995.
- [8] Luis Caffarelli, Fernando Charro, On a fractional Monge-Ampère operator, *Ann. PDE* 1 (1) (2015) 4, 47.
- [9] Luis Caffarelli, Luis Silvestre, An extension problem related to the fractional Laplacian, *Comm. Partial Differential Equations* 32 (7–9) (2007) 1245–1260.
- [10] Luis Caffarelli, Luis Silvestre, Regularity theory for fully nonlinear integro-differential equations, *Comm. Pure Appl. Math.* 62 (5) (2009) 597–638.
- [11] Luis Caffarelli, Luis Silvestre, Regularity results for nonlocal equations by approximation, *Arch. Ration. Mech. Anal.* 200 (1) (2011) 59–88.
- [12] Luis Caffarelli, Luis Silvestre, A nonlocal Monge-Ampère equation, *Comm. Anal. Geom.* 24 (2) (2016) 307–335.
- [13] Sun-Yung Alice Chang, Maria del Mar Gonzalez, Fractional Laplacian in conformal geometry, *Adv. Math.* 226 (2) (2011) 1410–1432.
- [14] Héctor Chang Lara, Regularity for fully non linear equations with non local drift. arXiv:1210.4242 [math.AP], 2012.
- [15] Héctor Chang Lara, Gonzalo Dávila, Regularity for solutions of nonlocal, nonsymmetric equations, *Ann. Inst. H. Poincaré Anal. Non Linéaire* 29 (6) (2012) 833–859.
- [16] Hector A. Chang-Lara, Nestor Guillen, Russell W. Schwab, Some free boundary problems recast as nonlocal parabolic equations. arXiv preprint arXiv:1807.02714, 2018.
- [17] Héctor A. Chang-Lara, Dennis Kriventsov, Further time regularity for nonlocal, fully nonlinear parabolic equations, *Comm. Pure Appl. Math.* 70 (5) (2017) 950–977.
- [18] Frank H. Clarke, Optimization and Nonsmooth Analysis, Vol. 5, SIAM, 1990.
- [19] Philippe Courrège, Sur la forme intégrale-différentielle des opérateurs de  $C_k^\infty$  dans  $C$  satisfaisant au principe du maximum. Séminaire Brelot-Choquet-Deny, Théor. Potentiel 10 (1) (1965) 1–38.
- [20] Hongjie Dong, Tianling Jin, Hong Zhang, Dini and Schauder estimates for nonlocal fully nonlinear parabolic equations with drifts, *Anal. PDE* 11 (6) (2018) 1487–1534.
- [21] Lawrence C. Evans, Some min–max methods for the Hamilton-Jacobi equation, *Indiana Univ. Math. J.* 33 (1) (1984) 31–50.
- [22] L.C. Evans, Panagiotis E. Souganidis, Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations, *Indiana Univ. Math. J.* 33 (5) (1984) 773–797.
- [23] David Gilbarg, Neil S. Trudinger, Elliptic Partial Differential Equations of Second Order, in: Classics in Mathematics, Springer-Verlag, Berlin, 2001, Reprint of the 1998 edition.
- [24] J.W. Green, F.A. Valentine, On the Arzela-Ascoli theorem, *Math. Mag.* 34 (4) (1961) 199–202.
- [25] Michael Grüter, Kjell-Ove Widman, The green function for uniformly elliptic equations, *Manuscripta Math.* 37 (3) (1982) 303–342.
- [26] Nestor Guillen, Jun Kitagawa, Russell W. Schwab, Estimates for dirichlet-to-neumann maps as integro-differential operators. arXiv preprint arXiv:1710.03152, 2017.
- [27] Nestor Guillen, Chenchen Mou, Andrzej Świech, Coupling Lévy measures and comparison principles for viscosity solutions. arXiv preprint arXiv:1805.06955, 2018.
- [28] Nestor Guillen, Russell W. Schwab, Min–max formulas for nonlocal elliptic operators. arXiv preprint arXiv:1606.08417, 2016.
- [29] Nestor Guillen, Russell W. Schwab, Aleksandrov–bakelman–pucci type estimates for integro-differential equations, *Arch. Ration. Mech. Anal.* 206 (1) (2012) 111–157.
- [30] Nestor Guillen, Russell W. Schwab, Neumann homogenization via integro-differential operators, *Discrete Contin. Dyn. Syst.* 36 (7) (2016) 3677–3703.
- [31] Nestor Guillen, Russell W. Schwab, Neumann homogenization via integro-differential operators. Part 2: singular gradient dependence, *SIAM J. Math. Anal.* 50 (2) (2018) 1679–1719.
- [32] Pei Hsu, On excursions of reflecting Brownian motion, *Trans. Amer. Math. Soc.* 296 (1) (1986) 239–264.
- [33] Nobuyuki Ikeda, Shinzo Watanabe, Stochastic Differential Equations and Diffusion Processes, Elsevier, 1981.
- [34] H. Ishii, P.-L. Lions, Viscosity solutions of fully nonlinear second-order elliptic partial differential equations, *J. Differential Equations* 83 (1) (1990) 26–78.
- [35] Espen R. Jakobsen, Kenneth H. Karlsen, A maximum principle for semicontinuous functions applicable to integro-partial differential equations, *NoDEA Nonlinear Differential Equations Appl.* 13 (2) (2006) 137–165.
- [36] Tianling Jin, Jingang Xiong, Schauder estimates for solutions of linear parabolic integro-differential equations, *Discrete Contin. Dyn. Syst.* 35 (12) (2015) 5977–5998.
- [37] Moritz Kassmann, Marcus Rang, Russell W. Schwab, Integro-differential equations with nonlinear directional dependence, *Indiana Univ. Math. J.* 63 (5) (2014) 1467–1498.
- [38] Markos A. Katsoulakis, A representation formula and regularizing properties for viscosity solutions of second-order fully nonlinear degenerate parabolic equations, *Nonlinear Anal.* 24 (2) (1995) 147–158.
- [39] Carlos E. Kenig, Potential theory of non-divergence form elliptic equations, in: Dirichlet Forms, Springer, 1993, pp. 89–128.
- [40] Shigeaki Koike, Andrzej Świech, Representation formulas for solutions of Isaacs integro-PDE, *Indiana Univ. Math. J.* 62 (5) (2013) 1473–1502.

- [41] Dennis Kriventsov,  $C^{1,\alpha}$  interior regularity for nonlinear nonlocal elliptic equations with rough kernels, *Comm. Partial Differential Equations* 38 (12) (2013) 2081–2106.
- [42] Thomas M. Liggett, *Continuous Time Markov Processes*, in: *Graduate Studies in Mathematics*, vol. 113, American Mathematical Society, Providence, RI, 2010.
- [43] Chenchen Mou, Yuming Zhang, Regularity theory for second order integro-PDEs. arXiv preprint [arXiv:1809.05589](https://arxiv.org/abs/1809.05589), 2018.
- [44] Xavier Ros-Oton, Joaquim Serra, Regularity theory for general stable operators, *J. Differential Equations* 260 (12) (2016) 8675–8715.
- [45] Ken-iti Sato, Tadashi Ueno, Multi-dimensional diffusion and the markov process on the boundary, *J. Math. Kyoto Univ.* 4 (3) (1965) 529–605.
- [46] René L. Schilling, Renming Song, Zoran Vondraček, *Bernstein Functions*, second ed., in: *De Gruyter Studies in Mathematics*, vol. 37, Walter de Gruyter & Co. Berlin, 2012, Theory and applications.
- [47] Russell W. Schwab, Periodic homogenization for nonlinear integro-differential equations, *SIAM J. Math. Anal.* 42 (6) (2010) 2652–2680.
- [48] Russell W. Schwab, Stochastic homogenization for some nonlinear integro-differential equations, *Comm. Partial Differential Equations* 38 (2) (2012) 171–198.
- [49] Russell W. Schwab, Luis Silvestre, Regularity for parabolic integro-differential equations with very irregular kernels, *Anal. PDE* 9 (3) (2016) 727–772.
- [50] Joaquim Serra, Regularity for fully nonlinear nonlocal parabolic equations with rough kernels, *Calc. Var. Partial Differential Equations* 54 (1) (2015) 615–629.
- [51] Luis Silvestre, On the differentiability of the solution to the Hamilton-Jacobi equation with critical fractional diffusion, *Adv. Math.* 226 (2) (2011) 2020–2039.
- [52] Luis Silvestre, Boyan Sirakov, Boundary regularity for viscosity solutions of fully nonlinear elliptic equations, *Comm. Partial Differential Equations* 39 (9) (2014) 1694–1717.
- [53] Panagiotis E. Souganidis, Max–min representations and product formulas for the viscosity solutions of Hamilton–Jacobi equations with applications to differential games, *Nonlinear Anal.* 9 (3) (1985) 217–257.
- [54] E.M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1971.
- [55] Pablo Raúl Stinga, José Luis Torrea, Extension problem and Harnack’s inequality for some fractional operators, *Comm. Partial Differential Equations* 35 (11) (2010) 2092–2122.