Regularity results for nonlocal equations by approximation

Luis Caffarelli and Luis Silvestre

March 31, 2010

Abstract

We obtain $C^{1,\alpha}$ regularity estimates for nonlocal elliptic equations that are not necessarily translation invariant using compactness and perturbative methods and our previous regularity results for the translation invariant case.

1 Introduction

We continue with the program started in [2] to extend the existing theory for (fully non linear) second order elliptic equations to nonlocal equations. In this paper we obtain nonlocal Cordes-Nirenberg type estimates. In [2] we obtained interior $C^{1,\alpha}$ estimates for nonlocal equations that are translation invariant. In this paper we extend those results by perturbative methods to non-translation invariant equations and other situations that were not covered in the results in [2].

The main theorem in this paper (Theorem 5.2) is presented in a somewhat abstract setting in order to make it applicable to as many situations as possible. It is a general perturbative argument that says that if an equation is uniformly close to another one which has $C^{1,\alpha}$ solutions, then the solutions of the former are also $C^{1,\alpha}$. In the last section of the paper we present concrete applications. The equations we have in mind are the purely nonlocal Isaacs equations, which have the form

$$\mathrm{I}(u,x) := \inf_{\alpha \in \mathcal{A}} \sup_{\beta \in \mathcal{B}} \int_{\mathbb{R}^n} (u(x+y) + u(x-y) - 2u(x))(2-\sigma) \frac{a_{\alpha\beta}(x,y)}{|y|^{n+\sigma}} \, \mathrm{d}y = f(x) \qquad \text{in } \Omega. \quad (1.1)$$

where $a_{\alpha\beta}$ is a family of nonnegative functions with indexes α and β ranging in arbitrarily sets $\mathcal{A}$ and $\mathcal{B}$. This type of integro-differential equations appear in stochastic games with jump processes. For different assumptions on the kernels $a_{\alpha\beta}$ we obtain different regularity results. Some of the most interesting applications in the last section of this paper include the following.

- if $0 < \lambda \leq a_{\alpha\beta} \leq \Lambda$ and $|\nabla_y a_{\alpha\beta}| \leq C|y|^{-1}$, and $a_{\alpha\beta}(x,y)$ is continuous in x for a modulus of continuity independent of y, then a solution u of (1.1) is $C^{1,\varepsilon}$ for some $\varepsilon > 0$ in the interior of Ω (with an estimate depending on $\|u\|_{L^\infty}$).

- if $0 < \lambda \leq a_{\alpha\beta} \leq \Lambda$ and $|\nabla_y a_{\alpha\beta}| \leq C|y|^{-1}$, $a_{\alpha\beta}$ is constant in x, and $\Lambda - \eta \leq 1 \leq \lambda + \eta$ for a small enough η, then the solution u of (1.1) is $C^{2,\varepsilon}$ for some $\varepsilon > 0$ in the interior of Ω (with an estimate depending on $\|u\|_{L^\infty}$).

The equations we are most interested in arise from stochastic control problems and stochastic games (See [2]). However we develop a general theory that applies as soon as some abstract assumptions are satisfied by a nonlocal equation. Note that in the expression (1.1), it only makes sense to consider symmetric kernels $a_{\alpha\beta}(x,y) = a_{\alpha\beta}(x,-y)$. Indeed, the antisymmetric part of the

kernel would have no effect when we integrate it against the second order incremental quotients of u (which are inherently symmetric).

In [1] the regularity of viscosity solutions of nonlinear integro-differential equations that are not translation invariant is also studied. They obtain Hölder continuity of the solutions. Their hypothesis are quite different from ours. They require certain modulus of continuity of the x dependence in the kernels. Their equations also involve second and first order terms, and they are able to include some degenerate equations into their hypothesis. In [7] a parabolic and convex integro-differential equation is considered and it is proved that the solutions are Lipschitz in space and $C^{1/2}$ Hölder continuous in time under appropriate conditions.

In the translation invariant case, the $C^{1,\alpha}$ estimate is a somewhat general fact for equations satisfying a comparison principle due to the fact that the directional derivatives of solutions satisfy some (a priori discontinuous) elliptic equation. If we try to apply the same idea for x-dependent elliptic equations we usually need to impose an excessive amount of regularity on the dependence. An alternative is to use compactness properties of equations and solutions (see [3] and [4]). In the local (classical PDEs) case the approximation takes place in a bounded domain (say the unit ball) and simply consists in showing (by compactness) that if the oscillation of the dependence on x of the variable equation is small enough, any bounded solution is close to the solution to the translation invariant equation inheriting its regularity up to a small error.

A good portion of the theory depends very little on the actual integral form of the operator $I(u, x)$ and it gives rise to interesting applications. Therefore, more generally, by a nonlocal operator I we will mean a rule that assigns a value $I(u, x)$ for any function u that is $C^2(x)$ and does not grow too much at infinity. A safe assumption would be that u is bounded since the kernels decay at infinity, but for particular cases it is enough to assume that $u \in L^1(\mathbb{R}^n, \omega)$ for a suitable weight ω related to the size of the nonlocal interactions (more concretely, a bound for the tails in the kernels $K_{\alpha\beta}$). We also assume that $I(u, x)$ is continuous if $u \in C^2$. Degenerate ellipticity is understood in the following way, if $v \geq u$ and $v(x) = u(x)$ for some x then $I(v, x) \geq I(u, x)$. Note that the usual elliptic second order operators can be thought of as a special case of this nonlocal elliptic operators, and also as a limiting case of integral operators of the form written above.

For every fixed x_0, the operator $I_{x_0} = I(-, x_0)$ given by $I_{x_0}(u, x) = I(\tau_{x-x_0}u, x_0)$, where $\tau_x u(y) = u(x + y)$, is translation invariant. This is just notation, which corresponds to *freezing* the coefficients in the classical linear theory.

As mentioned above, we need compactness of the family of solutions to equations with dependence on x, like De Giorgi or Krylov-Safonov theory for local operators. This we are able to do in the case of *uniform ellipticity* with respect to a family of linear operators. We say that I is uniformly elliptic with respect to a family of linear operators $\mathcal{L}$ if

$$M_{\mathcal{L}}^- v(x) \leq I(u + v, x) - I(u, x) \leq M_{\mathcal{L}}^+ v(x) \tag{1.2}$$

where

$$M_{\mathcal{L}}^- v(x) := \inf_{L \in \mathcal{L}} Lv(x)$$
$$M_{\mathcal{L}}^+ v(x) := \sup_{L \in \mathcal{L}} Lv(x)$$

This is the same as the definition in [2], but now we do not take the operator I necessarily translation invariant, and we want to include the case that I may be a second order partial differential operator (local operator). For second order equations, this is in fact the definition in [3] with respect to the Pucci extremal operators.

The class $\mathcal{L}$ represents the set of kernels that are allowed in an Isaacs equation like (1.1). Depending on the type of regularity results that we want to obtain, we need to impose less or more conditions on these kernels. That is the reason why we need to keep a flexible definition

of ellipticity with allows us to consider different classes $\mathcal{L}$. A common class of operators is the class $\mathcal{L}_0$ of uniformly elliptic operators of order σ defined in [2] and also later in this paper. These are the linear operators whose kernels are comparable to that of the fractional Laplacian of order σ but no regularity is imposed respect to y. It is the class for which Hölder estimates and the Harnack inequality is proved in [2]. However, in order to prove $C^{1,\alpha}$ estimates for a nonlinear problem like (1.1), a more restrictive class is needed which imposes some regularity of the kernels respect to y.

For the class $\mathcal{L}_0$, or any subclass of it, we can allow the functions u to have certain growth at infinity. The exact condition would be that $\mathrm{I}(u, x)$ is well defined as long as $u \in C^{1,1}(x)$ and $u \in L^1(\mathbb{R}^n, \frac{1}{1+|y|^{n+\sigma}})$ meaning that

$$\int_{\mathbb{R}^n} |u(y)| \frac{1}{1 + |y|^{n+\sigma}} \, \mathrm{d}y < +\infty.$$

Clearly, assuming that u is bounded is a sufficient condition but it is not necessary. In particular, for $\sigma > 1$, we can still compute $\mathrm{I}(u, x)$ if u has linear growth at infinity.

The **main result** of this paper is that if some fixed equation $\mathrm{I}^{(0)}u(x) = 0$ has interior $C^{1,\bar{\alpha}}$ estimates, and $\mathrm{I}(-, x)$ stays close to $\mathrm{I}^{(0)}$ in every scale as x varies in a given domain, then the equation $\mathrm{I}(u, x) = f(x)$ has interior $C^{1,\alpha}$ estimates for some $\alpha < \bar{\alpha}$. This is a nonlinear fractional-order version of the classical Cordes-Nirenberg estimate for second order elliptic equations.

For linear elliptic PDEs, the classical Cordes-Nirenberg estimate ([5], [8]) says that an equation

$$a_{ij}(x)\partial_{ij}u(x) = f(x)$$

has interior $C^{1,\alpha}$ estimates, provided that $f \in L^\infty$ and $|a_{ij}(x) - \delta_{ij}| \leq \varepsilon$ for a small enough ε. In [3], this regularity result was generalized to fully nonlinear equations. Our results in this paper recover the second order case since our estimates are uniform with respect to the order of the equation.

The main theorem in this paper is given in section 5. The proof follows more or less the idea of the proof in [4] chapter 8, but taking care of the nonlocal effects of the values of the function away from the domain of the equation. It is an approximation type argument, based on Lemma 4.7 in section 4 that says essentially that if two equations are very close, then so are their solutions.

It is necessary to give a meaning for two nonlocal equations to be close that will not impose restrictive conditions in the long range interactions and will still provide the necessary compactness. For this purpose, a suitable norm is defined in a preliminary way in section 2 for nonlocal operators. In section 5, the definition of the norm is refined so that it represents that two operators are close even when we rescale them.

The proof of Lemma 4.7 is by compactness, so it requires that the equations that we deal with have continuous solutions up to the boundary. This is why we have section 3 where we construct appropriate barriers to show that solutions to uniformly elliptic nonlocal equations are continuous up to the boundary.

Our main theorem is presented in a very general form in section 5. So in section 6 we provide a list of more concrete applications. The study of the regularity properties of nonlinear integro-differential equations is a relatively new subject. This paper is part of a program trying to extend the fundamental regularity results of fully nonlinear elliptic PDEs to nonlocal equations. The family of integro-differential equations is very rich in the sense that we can come up with a very large variety of equations. We try to provide theorems that apply to a general class of equations. There is always a balance between generality of results and simplicity of exposition and at this moment the choice is not always easy to make.

3

2 Preliminaries: Convergence of operators, uniform ellipticity and viscosity solutions.

In this section we introduce and discuss closeness of operators, uniform ellipticity and the notion of viscosity solutions. We also recall the regularity results of [2] and [9] for uniformly elliptic integro-differential equations. We also recall the interior $C^{1,\alpha}$ estimates from [2] with the hypothesis slightly relaxed.

We start with a definition of nonlocal operators (Definition 2.1). We next describe in what sense we will say that two operators are close. Since our approach is based on the viscosity method, closeness is measured by testing against smooth functions (Definition 2.2). Next we define and discuss in what context our operators are uniformly elliptic. At that point we are ready to define what a viscosity solution to our operators is (Definition 2.5). Then we review the regularity theorems proved in [2] and [9] about viscosity solutions.

In [9], there is a very simple proof that solutions of integro-differential equations with variable order are Hölder continuous, as long as the order of the equation stays strictly away from 0 and 2. The more complicated regularity estimates in [2] hold even when the order of the equation gets arbitrarily close to 2, however, only equations with constant order are considered. To a great extent, this was a choice for simplicity in the presentation, since many of the results in [2] would hold for equations of variable order, but the proofs would be somewhat messier. The results we present in this section are only a small variation of the results in [2]. For our main theorem in section 5, we will make a perturbation argument around an equation with interior $C^{1,\alpha}$ estimates. The applications in section 6 would extend with any extension of our theorems 2.6 and 2.7 below.

We will consider an absolutely continuous weight ω depending on the type of operators that we want to deal with. The purpose of this weight is to measure the contribution of the tails of the integral in the computation of an integro-differential operator Iu. Typically we have in mind $\omega(y) = 1/(1 + |y|^{n+\sigma_0})$ for some $\sigma_0 \in (1, 2)$, so that $u \in L^1(\mathbb{R}^n, \omega)$ means that the integrals in (1.1) are not singular at infinity for any $\sigma \geq \sigma_0$. Throughout this paper we always make the following assumptions on ω.

$$1 + |y| \in L^1(\mathbb{R}^n, \omega) \tag{2.1}$$

$$\sup_{B_r(y)} w \leq C(r)w(y) \tag{2.2}$$

These are the only things we need to assume on the weight ω for the general theorems of section 4. The results in this paper do not apply if the order σ of the equation is smaller than one, in that respect assumption (2.1) is reasonable. Assumption (2.2) is more technical and rules out singular measures ω.

We use the following convenient definition for a nonlocal operator.

Definition 2.1. *A nonlocal operator I is a rule that assigns to a function u a value* I(u, x) *at every point x, satisfying the following assumptions.*

- I(u, x) *is well defined as long as $u \in C^2(x)$ and $u \in L^1(\mathbb{R}^n, \omega)$.*

- *If $u \in C^2(\Omega) \cap L^1(\mathbb{R}^n, \omega)$, then $I(u, x)$ is continuous in Ω as a function of x.*

By $u \in C^2(x)$ we mean that there is a quadratic polynomial q such that $u(y) = q(y) + o(|x-y|^2)$ for y close to x.

Definition 2.1 is very general. It becomes useful when complemented with the concept of (nonlocal) uniform ellipticity (1.2). We do not include (1.2) in the definition above because we want to define the norm of a general operator (defined below) and compute it for differences of elliptic operators.

4

The results in this paper are based on approximation techniques. In order to use that type of argument, we need to understand when two nonlocal operators are close to each other. Since we use the Crandall-Lions viscosity concept, it is natural to define closeness of operators by testing them against the appropriate family of functions. We define the following norm.

Definition 2.2. *Given a nonlocal operator* I, *we define* $\|I\|$ *in some domain* Ω *with respect to some weight* ω *as*

$$\|I\| := \sup \big\{ |I(u,x)|/(1+M) : \ x \in \Omega$$
$$u \in C^2(x)$$
$$\|u\|_{L^1(\mathbb{R}^n,\omega)} \le M$$
$$|u(y) - u(x) - (y-x) \cdot \nabla u(x)| \le M|x-y|^2 \text{ for any } y \in B_1(x) \big\}.$$

The idea of this definition of $\|I\|$ is that $\|I_k - I\| \to 0$ if $I_k(u,x) \to I(u,x)$ in a somewhat uniform way. The condition $\|u\|_{L^1(\mathbb{R}^n,\omega)} \le M$ could be replaced by $|u| \le M$ in $\mathbb{R}^n$ which would make convergence weaker.

Note that in the *local* case when $I(u,x) = F(D^2u, x)$, we have $\omega = 0$ and $\|I\| = \tilde{\beta}(F)$, where $\tilde{\beta}$ is the one from [4].

Our uniform ellipticity assumption (1.2) depends on the class $\mathcal{L}$ of integro-differential operators. A class $\mathcal{L}$ is a set of operators L of the form

$$Lu(x) = \int_{\mathbb{R}^n} (u(x+y) + u(x-y) - 2u(x)) K(y) \, \mathrm{d}y,$$

for a family of nonnegative kernels K satisfying $\int \min(1, |y|^2) K(y) \, \mathrm{d}y < +\infty$.

We will make some minimal assumptions on every class $\mathcal{L}$. In this whole paper, every class $\mathcal{L}$ will satisfy the following two assumptions, that connect the class with the choice of the weight ω.

Assumption 2.3. *If* $K = \sup K_\alpha$ *is the supremum of all kernels in the class, then for every* $r > 0$

$$K(y) \le C_r \omega(y) \qquad \text{if } |y| \ge r.$$

for some constant C_r *(depending on* r*).*

Assumption 2.4. *For some constant* C *and every* $L \in \mathcal{L}$, *we have* $\|L\| \le C$ *(which clearly implies* $\|M_\mathcal{L}^+\| \le C$ *and also* $\|M_\mathcal{L}^-\| \le C$*).*

Assumption 2.3 says that the weight ω controls the tails of the kernels at every scale. Assumption 2.4 says that near the origin the kernels are uniformly controlled by C^2 functions. These are very weak assumptions and are all that is needed for our stability results of section 4. Our regularity results, however, require stronger assumptions in the class $\mathcal{L}$.

In order to understand the assumptions better, let us consider $\mathcal{L}$ to be the set of all fractional Laplacians $-(-\triangle)^{\sigma/2}$ for $\sigma \in [\sigma_0, 2]$, it would satisfy both assumptions with $\omega(y) = (2-\sigma_0)/(1+ |y|^{n+\sigma_0})$. On the other hand if we take the set $\mathcal{L}$ of all operators with kernels of the form $1/|y|^{n+\sigma}$ (where we are omitting the normalizing factor $\approx (2-\sigma)$), then this class would **not** satisfy Assumption 2.4.

In the case of local equations, the class $\mathcal{L}$ would consist of linear second order partial differential operators, and the weight ω is irrelevant (it can be taken as $\omega = 0$ and also $K = 0$). Our results in this case reduce to the usual results for uniformly elliptic partial differential equations [4].

We recall the definition of viscosity solutions given in [2].

5

Definition 2.5 (Viscosity Solutions). *Let I be a uniformly elliptic operator in the sense of (1.2) with respect to some class $\mathcal{L}$. A function $u : \mathbb{R}^n \to \mathbb{R}$, upper (lower) semi continuous in $\overline{\Omega}$, is said to be a subsolution (supersolution) to $\mathrm{I}(u, x) = f(x)$, and we write $\mathrm{I}(u, x) \geq f(x)$ $(\mathrm{I}(u, x) \leq f(x))$ in Ω, if every time all the following happens*

- x_0 *is any point in Ω.*

- N *is a neighborhood of x_0 in Ω.*

- φ *is some C^2 function in $\overline{N}$.*

- $\varphi(x_0) = u(x_0)$.

- $\varphi(x) > u(x)$ $(\varphi(x) < u(x))$ *for every $x \in N \setminus \{x_0\}$.*

Then if we choose as test function v,

$$v := \begin{cases} \varphi & in \ N \\ u & in \ \mathbb{R}^n \setminus N \ , \end{cases}$$

we have $\mathrm{I}(v, x_0) \geq f(x_0)$ $(\mathrm{I}(v, x_0) \leq f(x_0))$.
 A solution is a function u which is both a subsolution and a supersolution.

The first theorem we borrow from [2] is the interior Hölder estimates. We change the assumption of u being bounded for $u \in L^1(\mathbb{R}^n, 1/(1 + |y|^{n+\sigma_0}))$. Let us first recall the class $\mathcal{L}_0$ of integro differential operators.

An integro differential operator L belongs to $\mathcal{L}_0$ if its kernel K satisfies

$$(2 - \sigma)\frac{\lambda}{|y|^{n+\sigma}} \leq K(y) \leq (2 - \sigma)\frac{\Lambda}{|y|^{n+\sigma}} \ . \tag{2.3}$$

In this case $\mathrm{M}_{\mathcal{L}_0}^+$ and $\mathrm{M}_{\mathcal{L}_0}^-$ take a very simple form:

$$\mathrm{M}_{\mathcal{L}_0}^+ v(x) = (2 - \sigma) \int_{\mathbb{R}^n} \frac{\Lambda \delta v(x, y)^+ - \lambda \delta v(x, y)^-}{|y|^{n+\sigma}} \ \mathrm{d}y \tag{2.4}$$

$$\mathrm{M}_{\mathcal{L}_0}^- v(x) = (2 - \sigma) \int_{\mathbb{R}^n} \frac{\lambda \delta v(x, y)^+ - \Lambda \delta v(x, y)^-}{|y|^{n+\sigma}} \ \mathrm{d}y \ . \tag{2.5}$$

where δ represents the second order incremental quotients $\delta u(x, y) = u(x + y) + u(x - y) - 2u(x)$.

The class $\mathcal{L}_0$ depends on the parameter $\sigma \in (0, 2)$. Whenever it is necessary, we will stress that dependence by writing $\mathcal{L}_0(\sigma)$.

The next theorem concerns equations with kernels that can be arbitrarily discontinuous in terms of x. In particular, it applies to the difference of two solutions of a nonlinear integral operator.

Theorem 2.6. *Let $\sigma > \sigma_0$ for some $\sigma_0 > 0$. Let u be function in $L^1(\mathbb{R}^n, \omega)$ for $\omega = 1/(1 + |y|^{n+\sigma_0})$, such that u is continuous in $\overline{B}_1$ and*

$$\mathrm{M}_{\mathcal{L}_0}^+ u \geq -C_0 \qquad in \ B_1$$
$$\mathrm{M}_{\mathcal{L}_0}^- u \leq C_0 \qquad in \ B_1$$

then there is an $\alpha > 0$ (depending only on λ, Λ, n and σ_0) such that $u \in C^\alpha(B_{1/2})$ and

$$u_{C^\alpha(B_{1/2})} \leq C\big(\sup_{B_1} |u| + \|u\|_{L^1(\mathbb{R}^n, \omega)} + C_0 \big)$$

for some constant $C > 0$ which depends on σ_0, λ, Λ and dimension, but not on σ.

Proof. Consider $\tilde{u}(x) = \chi_{B_1}(x)u(x)$. A direct computation shows that $\mathrm{M}^+\tilde{u} \geq -C_0 - c\,\|u\|_{L^1(\mathbb{R}^n,\omega)}$ and $\mathrm{M}^-\tilde{u} \leq C_0 + c\,\|u\|_{L^1(\mathbb{R}^n,\omega)}$ in $B_{3/4}$. We can now apply theorem 12.1 from [2] to $\tilde{u}$. $\square$

Theorem 2.6 is a nonlocal version of Krylov-Safonov theorem for elliptic PDEs in nondivergence form. A nonlocal version of De Giorgi-Nash-Moser theorem can be found in [6].

We will also mention, as a consequence of the previous theorem, the interior $C^{1,\alpha}$ estimates from [2] for functions with some growth at infinity.

Given $\rho_0 > 0$, we define the class $\mathcal{L}_*$ by the operators L with kernels K such that

$$(2-\sigma)\frac{\lambda}{|y|^{n+\sigma}} \leq K(y) \leq (2-\sigma)\frac{\Lambda}{|y|^{n+\sigma}} \tag{2.6}$$

$$|\nabla K(y)| \leq C\omega(y) \qquad \text{in } \mathbb{R}^n \setminus B_{\rho_0} \tag{2.7}$$

Theorem 2.7. *Assume $\sigma \geq \sigma_0 > 0$. There is a $\rho_0 > 0$ (depending on λ, Λ, σ_0 and n) so that if I is a nonlocal translation invariant uniformly elliptic operator with respect to $\mathcal{L}_*$ and u is a continuous function in $\overline{B}_1$ such that $u \in L^1(\mathbb{R}^n, \omega)$ for $\omega = 1/(1+|y|^{n+\sigma_0})$, $\mathrm{I}u = 0$ in B_1, then there is a universal (depends only on λ, Λ, n and σ_0) $\alpha > 0$ such that $u \in C^{1+\alpha}(B_{1/2})$ and*

$$u_{C^{1+\alpha}(B_{1/2})} \leq C\left(\sup_{B_1}|u| + \|u\|_{L^1(\mathbb{R}^n,\omega)} + |\mathrm{I}0|\right)$$

for some constant $C > 0$ (where by $\mathrm{I}0$ we mean the value we obtain when we apply I to the function that is constant equal to zero). The constant C depends on λ, Λ, σ_0, n and the constant in (2.7).

Sketch of the proof. The proof of this theorem goes along the lines of the proof of Theorem 12.1 in [2] by applying Theorem 2.6 to the differential quotients $w^h = (u(x+h) - u(x))/|h|^\beta$ for $\beta = \alpha, 2\alpha, \ldots, 1$. In each step, the quotient w^h is written as $w^h = w_1^h + w_2^h$ where w_2^h vanishes in a ball B_r (with $1/2 < r < 1$) and w_1^h is bounded in $\mathbb{R}^n$, $\mathrm{M}^- w_1^h \leq C$ and $\mathrm{M}^+ w_1^h \geq -C$ in B_r. This constant C, like in the proof of theorem 12.1 in [2], comes from the estimate

$$\int\limits_{\mathbb{R}^n \setminus B_\rho} |u(x+y)| \frac{|K(y) - K(y-h)|}{|h|}\, \mathrm{d}y < C.$$

In [2] the assumption was $u \in L^\infty$ whereas now we have $u \in L^1(\mathbb{R}^n, \omega)$ and so is why the assumption (12.2) in [2] necessary for the bound of the constant C is different from our assumption (2.7) in this paper. $\square$

Note that $\sup_{B_1}|u| + \|u\|_{L^1(\mathbb{R}^n,\omega)} \leq C \sup |u(x)|/(1+|x|^{1+\alpha})$ for any $\alpha < \sigma_0$.

In comparison with $\mathcal{L}_0$, the class $\mathcal{L}_*$ adds the extra regularity assumption (2.7) to the kernels with respect to y. This condition is needed in order to prove the interior $C^{1,\alpha}$ estimates in the ball of radius one. However, the hypothesis is not scale invariant, which makes it not suitable to the type perturbation arguments used later in this paper. Therefore, in section 5 we will further restrict this class in order to make it scale invariant.

Remark 2.8. Assumption (2.7) is used to control the effect on the regularity of u at x of the points y that are at least ρ_0 away from x. This condition depends on the type of control we have of u far from x. In [2] we were assuming the stronger condition $u \in L^\infty(\mathbb{R}^n)$, so we had the weaker condition for the kernels

$$\int\limits_{\mathbb{R}^n \setminus B_{\rho_0}} \frac{|K(y) - K(y-h)|}{|h|}\, \mathrm{d}y \leq C \qquad \text{every time } |h| < \frac{\rho_0}{2}$$

7

Now we have the weaker condition $u \in L^1(\mathbb{R}^n, \omega)$ instead. If we want our estimate to depend on different norms of u, our condition (2.7) will change accordingly.

It is also convenient to define M_2^+ and M_2^- as

$$M_2^+ u(x) = \lim_{\sigma \to 2} M_{\mathcal{L}_0(\sigma)}^+ u(x)$$

$$M_2^- u(x) = \lim_{\sigma \to 2} M_{\mathcal{L}_0(\sigma)}^- u(x)$$

A simple computation shows that $M_2^+ u(x)$ is a second order uniformly elliptic operator of the form $M_2^+ u(x) = F(D^2 u)$ with

$$F(A) := \int_{S^{n-1}} \Lambda \langle y, Ay \rangle^+ - \lambda \langle y, Ay \rangle^- \, dS.$$

Note that the ellipticity constants $\tilde{\lambda}$ and $\tilde{\Lambda}$ of M_2^+ as a partial differential operator depend only on λ, Λ and dimension. Also $F(0) = 0$, so $M_2^+ u(x) \leq M^+(D^2 u)$ where $M^+(D^2 u)$ is the usual Pucci operator with ellipticity constants $\tilde{\lambda}$ and $\tilde{\Lambda}$. Naturally, we also have the corresponding relations for M_2^-.

3 Boundary regularity

In this section we show that if we have a modulus of continuity on the boundary of the domain of some equation, then we can find a (possibly different) modulus of continuity inside the domain. This is done, as usual, by controlling the growth of u away from its boundary data through barriers, scaling and interior regularity.

We start by showing that a truncation of an appropriate radial function is a barrier outside the unit ball. Our barrier function is good as a supersolution $M^+\varphi \leq 0$ for all values of σ greater than a given σ_0, where M^+ is the maximal operator $M_{\mathcal{L}_0(\sigma)}^+$. Another way to say this would be to define a larger class $\mathcal{L}$ which is the union of all classes $\mathcal{L}_0(\sigma)$ for $\sigma \in (\sigma_0, 2)$, then $M_{\mathcal{L}}^+\varphi \leq 0$.

Using this barrier function we will prove that the solutions of uniformly elliptic equations are continuous on the boundary, and then using their interior Hölder regularity (Theorem 2.6) we get that they are continuous up to the boundary in the domain of the equation. The result in Theorem 3.3 is restricted to equations of constant order only because Theorem 2.6 is. In fact, the same continuity result holds for operators of variable order as long as the order remains strictly in between 0 and 2 (because of the results in [9]).

Lemma 3.1. *Given any $\sigma_0 \in (0, 2)$, there is an $\alpha > 0$ and $r > 0$ small so that the function $u(x) = ((|x| - 1)^+)^\alpha$ satisfies $M^+ u(x) \leq 0$ in $B_{1+r} \setminus B_1$ for any $\sigma > \sigma_0$.*

We defer the proof of this lemma to the appendix. We use it to obtain the following corollary.

Corollary 3.2. *For any constant C, there is a continuous function φ such that*

- $\varphi = 0$ *in B_1.*

- $\varphi \geq 0$ *in $\mathbb{R}^n$.*

- $\varphi \geq 1$ *in $\mathbb{R}^n \setminus B_2$.*

- $M^+\varphi \leq 0$ *in $\mathbb{R}^n \setminus B_1$ for any $\sigma > \sigma_0$.*

8

Proof. Consider $\varphi = \min(1, C((|x|-1)^+)^\alpha)$ for some large constant C and apply Lemma 3.1. $\square$

We will use the function φ of Corollary 3.2 as a barrier to prove the boundary continuity of solutions to nonlocal equations. We will do it in spherical domains. Note that φ is a supersolution outside of a ball, so the same method could be used in domains with the exterior ball condition.

Theorem 3.3. *Let $\sigma > \sigma_0 > 0$. Let ρ be a modulus of continuity and M^+ and M^- be the maximal operators $M^+_{\mathcal{L}_0(\sigma)}$ and $M^-_{\mathcal{L}_0(\sigma)}$. Let u be a bounded function such that*

$$\begin{aligned}
M^+u(x) &\geq -C & &\text{in } B_1 \\
M^-u(x) &\leq C & &\text{in } B_1 \\
|u(y) - u(x)| &\leq \rho(|x-y|) & &\text{for every } x \in \partial B_1 \text{ and } y \in \mathbb{R}^n \setminus B_1.
\end{aligned}$$

Then there is another modulus of continuity $\tilde\rho$ so that $|u(y) - u(x)| \leq \tilde\rho(|x-y|)$ for every $x \in \overline{B}_1$ and $y \in \mathbb{R}^n$.

The modulus of continuity $\tilde\rho$ depends only on ρ, λ, Λ, σ_0, n, $\|u\|_{L^\infty}$ and C (but not on σ).

The modulus of continuity $\tilde\rho$ in Theorem 3.3 is in some sense the worst between the internal regularity and the boundary regularity.

Before proving Theorem 3.3, we mention a corollary.

Corollary 3.4. *Let $\sigma > \sigma_0 > 0$. Let ρ be a modulus of continuity. Let u be a function such that*

$$\begin{aligned}
\|u\|_{L^1(R^n, 1/(1+|y|^{n+\sigma_0}))} &\leq C \\
M^+u(x) &\geq -C & &\text{in } B_1 \\
M^-u(x) &\leq C & &\text{in } B_1 \\
|u(y) - u(x)| &\leq \rho(|x-y|) & &\text{for every } x \in \partial B_1 \text{ and } y \in \mathbb{R}^n \setminus B_1.
\end{aligned}$$

Then there is another modulus of continuity $\tilde\rho$ so that $|u(y) - u(x)| \leq \tilde\rho(|x-y|)$ for every $x \in \overline{B}_1$ and $y \in \mathbb{R}^n$.

The modulus of continuity $\tilde\rho$ depends only on ρ, λ, Λ, σ_0, n, and C.

Proof of corollary. Since $|u(y) - u(x)| \leq \rho(|x-y|)$ if $x \in \partial B_1$ and $\int_{B_2} |u| \, dx < C$, then $|u| \leq C$ on ∂B_1. Thus, from the modulus of continuity, $u \leq C$ in B_2.

We apply Theorem 3.3 to $\tilde u(x) = \min(C, \max(u, C))$. Note that $M^+\tilde u \leq M^+u + M^+(\tilde u - u)$ and $M^+(\tilde u - u)$ is controlled by $\|u\|_{L^1(R^n, 1/(1+|y|^{n+\sigma_0}))}$. $\square$

In order to prove Theorem 3.3, we will need a couple of preparatory lemmas that mirror the second order approach. Lemma 3.5 reconstructs the modulus of continuity on the boundary, and Lemma 3.6 in the interior.

Lemma 3.5. *Assume $\sigma \geq \sigma_0 > 0$. Let ρ be a modulus of continuity. Let u be a bounded function such that*

$$\begin{aligned}
M^+u(x) &\geq -C & & \\
u(y) - u(x) &\leq \rho(|x-y|) & &\text{for every } x \in \partial B_1 \text{ and } y \in \mathbb{R}^n \setminus B_1.
\end{aligned}$$

Then there is another modulus of continuity $\tilde\rho$ so that $u(y) - u(x) \leq \tilde\rho(|x-y|)$ for every $x \in \partial B_1$ and $y \in \mathbb{R}^n$.

The modulus of continuity $\tilde\rho$ depends only on ρ, λ, Λ, σ_0, n, the constant C above, and $\|u\|_{L^\infty}$.

9

Proof. We first observe that since $\sigma \geq \sigma_0$ the function $p(x) = \max(0, 4 - |x|^2)$ satisfies $M^+ p \geq -c$ in B_1 for some $c > 0$. So we can consider $u - \frac{C}{c} p$ and reduce the problem to the case $C = 0$ (the values of $\|u\|_{L^\infty}$ and ρ are affected depending on C).

Let ρ_0 be the modulus of continuity of the function φ of Corollary 3.2. Let $\tilde{\rho}$ be

$$\tilde{\rho}(r) = \inf_R \left(\rho(3R) + \|u\|_{L^\infty} \rho_0 \left(\frac{r}{R} \right) \right)$$

Let us first show that $\tilde{\rho}$ is indeed a modulus of continuity. The function $\tilde{\rho}$ is clearly monotone increasing because ρ_0 is. We still need to show that for every $\varepsilon > 0$, there is an $r > 0$ such that $\tilde{\rho}(r) < \varepsilon$. We first choose R such that $\rho(3R) < \varepsilon/2$, and then choose r such that $\|u\|_{L^\infty} \rho_0 \left(\frac{r}{R} \right) < \varepsilon/2$.

Now, we show that $\tilde{\rho}(|y - x|) \geq u(y) - u(x)$ for any $y \in \mathbb{R}^n$ and $x \in \partial B_1$. For any $R > 0$, Consider the barrier function

$$B(y) = u(x) + \rho(3R) + \|u\|_{L^\infty} \varphi \left(-x + \frac{y - x}{R} \right).$$

We see that $B(x) \geq u(x) + \rho(3R) \geq u$ in $B_{3R}(x) \cap (\mathbb{R}^n \setminus B_1)$. Also $B(x) \geq \|u\|_{L^\infty}$ in $\mathbb{R}^n \setminus B_{3R}(x)$. Moreover $M^+ B \leq 0$, whereas $M^+ u \geq 0$ in B_1, so $u \leq B \leq u(x) + \tilde{\rho}(|y - x|)$ everywhere. Taking infimum from all possible choices of R we obtain $\tilde{\rho}(|y - x|) \geq u(y) - u(x)$. $\square$

Lemma 3.6. *Assume $\sigma > \sigma_0 > 0$. Let ρ be a modulus of continuity. Let u be a bounded function such that*

$$M^+ u(x) \geq -C \qquad\qquad in \ B_1 \qquad\qquad (3.1)$$
$$M^- u(x) \leq C \qquad\qquad in \ B_1 \qquad\qquad (3.2)$$
$$|u(y) - u(x)| \leq \rho(|x - y|) \qquad\qquad for \ every \ x \in \partial B_1 \ and \ y \in \mathbb{R}^n. \qquad (3.3)$$

Then there is another modulus of continuity $\tilde{\rho}$ so that $|u(y) - u(x)| \leq \tilde{\rho}(|x - y|)$ for every $x \in \overline{B}_1$ and $y \in \mathbb{R}^n$.

The modulus of continuity $\tilde{\rho}$ depends only on ρ, λ, Λ, the constant C above, σ_0, n, and $\|u\|_{L^\infty}$.

Proof. Let $x \in \overline{B}_1$ and $y \in \mathbb{R}^n$. We will consider two cases. If $\mathrm{dist}(x, y) < \mathrm{dist}(x, \partial B_1)/4$ or if $\mathrm{dist}(x, y) \geq \mathrm{dist}(x, \partial B_1)/4$.

If $\mathrm{dist}(x, y) \geq \mathrm{dist}(x, \partial B_1)/4$ then let x_0 be a point in ∂B_1 where $\mathrm{dist}(x, x_0) = \mathrm{dist}(x, \partial B_1)$,

$$|u(x) - u(y)| \leq |u(x) - u(x_0)| + |u(x_0) - u(y)|$$
$$\leq \rho(|x - x_0|) + \rho(|x_0 - y|)$$
$$\leq \rho(|x - y|/4) + \rho(5|x - y|/4)$$

Let us consider now the case $\mathrm{dist}(x, y) < \mathrm{dist}(x, \partial B_1)/4$. Let $r = \mathrm{dist}(x, \partial B_1)/2$. Let us truncate the function u at $u(x_0) \pm \rho(4r)$ by defining $\overline{u} := \min(u(x_0) + \rho(4r), \max(u, u(x_0) - \rho(4r)))$. So we have

$$|u(z) - \overline{u}(z)| \leq \min((\rho(r + |z - x|) - \rho(4r))^+, M)$$

where $M = 2\|u\|_{L^\infty}$.

We now scale the functions so that $B_r(x)$ becomes B_1. We write

$$\overline{v}(z) = \overline{u}(x + rz)$$
$$v(z) = u(x + rz)$$

10

and we have
$$|v(z) - \overline{v}(z)| \leq \min((\rho(r + r|z|) - \rho(4r))^+, M).$$

We want to estimate $M^+\overline{v}$ in B_1. We know that $M^+v \geq -Cr^\sigma$ in B_1 by rescaling (3.1). Also $M^+\overline{v} \geq M^+v - M^+(v - \overline{v})$. We now estimate $M^+(v - \overline{v})$ in B_1 by above. We have

$$M^+(v(z) - \overline{v}(z)) \leq (2 - \sigma)\Lambda \int_{\mathbb{R}^n \setminus B_1} \frac{2\min((\rho(r + r|y|) - \rho(4r))^+, M)}{|y|^{n+\sigma}} \, dy$$

$$=: I_r$$

We observe that $I_r \to 0$ as $r \to 0$ uniformly in σ as long as $\sigma \geq \sigma_0$. We thus have $-Cr^\sigma - I_r \leq M^+\overline{v}$ in B_1, and $Cr^\sigma + I_r \to 0$ as $r \to 0$.

We can repeat the same computation to obtain $M^-\overline{v} \leq Cr^\sigma + I_r$ in B_1. Applying interior C^α estimates [2], we obtain

$$\overline{v}(z) - \overline{v}(0) \leq C(Cr^\sigma + I_r + \rho(4r))|z|^\alpha = Cm_r|z|^\alpha$$

where $m_r := Cr^\sigma + I_r + \rho(4r)$ goes to zero as $r \to 0$ uniformly for $\sigma \geq \sigma_0$.

Now we scale back with $y = x + rz$ to apply the above estimate to u. We have

$$|u(y) - u(x)| \leq Cm_r \left(\frac{|x - y|}{r}\right)^\alpha$$

so we have proven that $|u(y) - u(x)| \leq \tilde{\rho}(|x - y|)$ as long as we can chose $\tilde{\rho}(d) \geq C\sup_{r>4d} \frac{d^\alpha m_r}{r^\alpha}$. There will be such modulus of continuity as long as we can prove that

$$\sup_{r>4d} \frac{d^\alpha m_r}{r^\alpha} \to 0 \text{ as } d \to 0 \tag{3.4}$$

We will prove this by contradiction to finish the proof of this lemma. Assume (3.4) was not true. Then there would be an $\varepsilon > 0$ and two sequences r_i and d_i with $r_i > 4d_i$ and $d_i \to 0$ such that $d_i^\alpha m_{r_i}/r_i^\alpha > \varepsilon$. Since m_r is bounded above, d_i^α/r_i^α cannot go to zero. But that would imply that $r_i \to 0$ and $d_i^\alpha m_{r_i}/r_i^\alpha < m_{r_i}/4^\alpha \to 0$, which gives a contradiction. $\qquad\square$

Remark 3.7. If we only considered Hölder modulus of continuity, the proof would not be much shorter but it would be more direct. It is nevertheless interesting that it can be made to work for arbitrary modulus of continuity.

Proof of Theorem 3.3. We apply Lemma 3.5 to both u and $-u$ to obtain a modulus of continuity that applies from any point on the boundary ∂B_1 to any point in space $\mathbb{R}^n$. Then we use Lemma 3.6 to finish the proof. $\qquad\square$

Remark 3.8. The conditions $\mathrm{M}^+u \geq -C$ and $\mathrm{M}^-u \leq C$ are implied by u being a solution of a linear equation like

$$\int_{\mathbb{R}^n} \delta u(x, y)(2 - \sigma)\frac{a(x, y)}{|y|^{n+\sigma}} \, dy = f(x)$$

no continuity whatsoever is required for the coefficient $a(x, y)$, just that $\lambda \leq a(x, y) \leq \Lambda$. The function f is only required to be bounded. The modulus of continuity is independent of σ as long as $\sigma > \sigma_0 > 0$, but the result was proved for constant order σ. We could obtain a corresponding result for a solution u of an equation with variable order $\sigma(x)$ if instead of Theorem 2.6, we used the estimates from [9]. The disadvantage is that those estimates would blow up as $\sigma \to 2$.

11

4 Approximation results.

The purpose of this section is to show that if two equations are very close to each other, then so are their solutions. The proof is by compactness: if a sequence of solutions and corresponding equations converges, the limiting function is the unique viscosity solution of the limiting equation. We stress that, thanks to the concept of viscosity solutions, it suffices that the convergence of the functionals be weak (i.e. against smooth test functions).

We obtain our desired result in great generality in Lemmas 4.3 and 4.7. To prove those lemmas we need some previous technical results that we develop in this section.

In the following lemma, we notice that the test functions used in the definition of viscosity solutions can be reduced to only the polynomials of degree two.

Lemma 4.1. *In Definition 2.5 it is enough to consider test functions φ which are quadratic polynomials and neighborhoods N which are balls centered at the touching point.*

Proof. Assume Definition 2.5 holds every time φ is a quadratic polynomial and N is a ball centered at the touching point. We will show it also holds for any C^2 function φ.

Let φ be any C^2 test function from above in a neighborhood N (similar reasoning would apply for test functions from below). Without loss of generality we assume φ touches u at the origin. Let $p(x)$ be the following quadratic polynomial.

$$p(x) = \langle (D^2\varphi(0) + \varepsilon I)x, x \rangle + x \cdot D\varphi(0) + \varphi(0)$$

So, $p \geq \varphi \geq u$ in a neighborhood of the origin B_r. Let

$$v_\varphi(x) = \begin{cases} \varphi(x) & \text{if } x \in N \\ u(x) & \text{if } x \notin N \end{cases}$$

$$v_p(x) = \begin{cases} p(x) & \text{if } x \in B_r \\ u(x) & \text{if } x \notin B_r. \end{cases}$$

We choose r small so that $v_p(x) \leq v_\varphi(x) + 2\varepsilon|x|^2\chi_{B_r}$.

By assumption, $\mathrm{I}(v_p, 0) \geq f(0)$. On the other hand

$$\begin{aligned}
\mathrm{I}(v_\varphi, 0) - \mathrm{I}(v_p, 0) &\geq \mathrm{M}_{\mathcal{L}}^-(v_\varphi - v_p, 0) \\
&\geq -\mathrm{M}_{\mathcal{L}}^+((v_\varphi - v_p)^-, 0) \\
&\geq -C\mathrm{M}_{\mathcal{L}}^+(2\varepsilon|x|^2\chi_{B_r}, 0) \geq -C\varepsilon
\end{aligned}$$

where in the last inequality we used that $\left\|\mathrm{M}_{\mathcal{L}}^+\right\| \leq C$.

Se we have $\mathrm{I}(v_\varphi, 0) \geq \mathrm{I}(v_p, 0) - C\varepsilon \geq f(0) - C\varepsilon$ for arbitrary $\varepsilon > 0$. Thus $\mathrm{I}(v_\varphi, 0) \geq f(0)$. $\square$

We want to show that if $\mathrm{I}_k(u_k, x) = f_k(x)$ and $\mathrm{I}_k \to \mathrm{I}$, $u_k \to u$ and $f_k \to f$ in some appropriate way, then $\mathrm{I}(u, x) = f(x)$. It is indeed true that the convergence $\mathrm{I}_k \to \mathrm{I}$ can be taken with respect to the norm defined in section 2, however we need a stronger result with respect to a weaker version of convergence that we define below.

Definition 4.2. *We say that $\mathrm{I}_k \to \mathrm{I}$ weakly in Ω if for every $x_0 \in \Omega$ and for every function v of the form*

$$v(x) = \begin{cases} p(x) & \text{if } |x - x_0| \leq \rho \\ u(x) & \text{if } |x - x_0| \geq \rho \end{cases}$$

where p is a polynomial of degree two and $u \in L^1(\mathbb{R}^n, \omega)$, we have $\mathrm{I}_k(v, x) \to \mathrm{I}(v, x)$ uniformly in $B_{\rho/2}(x_0)$.

The following lemma is an improvement of Lemma 4.5 from [2].

Lemma 4.3. *Let* I_k *be a sequence of uniformly elliptic operators with respect to some class* $\mathcal{L}$. *Let us assume that Assumption 2.3 holds. Let* u_k *be a sequence of lower semicontinuous functions in* Ω *such that*

- $I_k(u_k, x) \leq f_k(x)$ *in* Ω.

- $u_k \to u$ *in the* Γ *sense in* Ω.

- $u_k \to u$ *in* $L^1(\mathbb{R}^n, \omega)$.

- $f_k \to f$ *locally uniformly in* Ω.

- $I_k \to I$ *weakly in* Ω *(with respect to* ω)

- $|u_k(x)| \leq C$ *for every* $x \in \Omega$.

Then $Iu \leq f$ *in* Ω.

Proof. Let p be a polynomial of degree 2 touching u from below at a point x in a neighborhood $B_r(x)$.

Since u_k Γ-converges to u in Ω, for large k, we can find x_k and d_k such that $p + d_k$ touches u_k at x_k. Moreover $x_k \to x$ and $d_k \to 0$ as $k \to +\infty$.

Since $I_k(u_k, x) \leq f_k(x)$, if we let

$$v_k = \begin{cases} p + d_k & \text{in } B_r(x) \\ u_k & \text{in } \mathbb{R}^n \setminus B_r(x) \ , \end{cases}$$

we have $I_k(v_k, x_k) \leq f_k(x_k)$. Clearly $v = \lim_{k \to +\infty} v_k$.

Let $z \in B_{r/4}(x)$. We have

$$|I_k(v_k, z) - I(v, z)| \leq |I_k(v_k, z) - I_k(v, z)| + |I_k(v, z) - I(v, z)|$$

$$\leq \max\left(|M_{\mathcal{L}}^+(v_k - v)(z)|, |M_{\mathcal{L}}^+(v - v_k)(z)|\right) + |I_k(v, z) - I(v, z)|$$

$$\leq \sup_{L \in \mathcal{L}} |L(v_k - v)(z)| + |I_k(v, z) - I(v, z)|$$

$$\leq \left(\int_{\mathbb{R}^n \setminus B_{r/2}} |\delta(v_k - v)(z, y)| K(y) \, \mathrm{d}y \right) + |I_k(v, z) - I(v, z)|$$

$$\leq C \left(\int_{\mathbb{R}^n \setminus B_{r/2}} 2|v_k(y) - v(y)| \sup_{z \in B_{r/4}} \omega(y + z) \, \mathrm{d}y \right) + C|v_k(z) - v(z)| + |I_k(v, z) - I(v, z)|$$

$$\leq C \|v_k - v\|_{L^1(\mathbb{R}^n, \omega)} + C|v_k(z) - v(z)| + |I_k(v, z) - I(v, z)|$$

Where we used Assumption 2.3 and (2.2).

The term $|I_k(v, z) - I(v, z)|$ goes to zero uniformly for $z \in B_{r/4}(x)$ since we assumed $I_k \to I$ weakly. The sequence $\|v_k - v\|_{L^1(\omega)} \to 0$ and $u_k \to u$ uniformly in $B_r(x)$, therefore the right hand side in the above inequality goes to zero uniformly for $z \in B_{r/4}$. We obtain $I_k(v_k, x) \to I(v, x)$ uniformly in $B_{r/4}(x)$.

We have that Iv is continuous in $B_r(x)$. We now compute

$$|I_k(v_k, x_k) - I(v, x)| \leq |I_k(v_k, x_k) - I(v, x_k)| + |I(v, x_k) - I(v, x)| \to 0 \ .$$

So $I_k(v_k, x_k)$ converges to $I(v, x)$, as $k \to +\infty$. Since $x_k \to x$ and $f_k \to f$ locally uniformly, we also have $f_k(x_k) \to f(x)$, which finally implies $I(v, x) \leq f(x)$. □

13

Remark 4.4. Note that the weight ω is important. Consider the following example. Let $I_k = -(-\triangle)^{1-1/k}$ and $I = \triangle$. Then I is a local operator, and we may think that we can take $\omega = 0$ and overlook the $L^1(\mathbb{R}^n, \omega)$ convergence. However we can consider the sequence of functions

$$u_k(x) = \chi_{B_1}(x)(|x|^2 - 1) - M_k \chi_{\mathbb{R}^n \setminus B_{2k}} \ .$$

If we choose M_k large enough, then we can make $I_k u_k \leq 0$ in B_1 for every k, but $If = \triangle f = 2n > 0$ in B_1.

The point is that in this example there is no single weight ω for which at the same time $u_k \to u$ in $L^1(\mathbb{R}^n, \omega)$ and $\|I_n - I\| \to 0$ with respect to ω.

Lemma 4.5. *Let v be a function*

$$v(x) = \begin{cases} p(x) & \text{if } x \in B_r \\ u(x) & \text{if } x \notin B_r. \end{cases}$$

and let I_k be a sequence of uniformly elliptic operators respect to some class $\mathcal{L}$ satisfying Assumption 2.3. There is a subsequence I_{k_j} such that $f_{k_j}(x) := I_{k_j}(v, x)$ converges uniformly in $B_{r/2}$.

Proof. All we need to do is find a uniform modulus of continuity for f_k in $B_{r/2}$ so that the Lemma follows by Arzela-Ascoli.

Recall the notation $\tau_z u(x) = u(x+z)$. Given $x, y \in B_{r/2}$ with $|x - y| < r/8$, we have

$$f_k(x) - f_k(y) = I_k(v, x) - I_k(v, y)$$
$$\leq M_{\mathcal{L}}^+(v - \tau_{y-x} v, x)$$

Note that $v - \tau_{y-x} v$ is just linear function in $B_{r/4}(x)$

$$\leq \int_{\mathbb{R}^n \setminus B_{r/4}} \delta(v - \tau_{y-x} v, x, z) K(z) \, \mathrm{d}z$$

where K is the kernel of Assumption 2.3, thus we have

$$\leq C \int_{\mathbb{R}^n \setminus B_{r/4}} (v(x+z) + v(x-z) - 2v(x) - v(y+z) - v(y-z) + 2v(y))\omega(z) \, \mathrm{d}z$$

$$\leq C(|v(x) - v(y)| + \|\tau_x v - \tau_y v\|_{L^1(\mathbb{R}^n, \omega)})$$

$$\leq C \left(\sup_{|\bar{x}-\bar{y}| \leq |x-y|, \ \bar{x}, \bar{y} \in B_{r/2}} |v(\bar{x}) - v(\bar{y})| + \int_{\mathbb{R}^n} |\tau_{(\bar{x}-\bar{y})} v - v|(\sup_{B_r(z)} \omega) \, \mathrm{d}z \right) \leq c(|x - y|)$$

Where we use (2.2) and $c(\rho)$ is defined as

$$c(\rho) := C \left(\sup_{|x-y| \leq \rho, \ x, y \in B_{r/2}} |v(x) - v(y)| + \int_{\mathbb{R}^n} |\tau_{(x-y)} v - v| \omega(z) \, \mathrm{d}z \right).$$

Clearly, $c(\rho)$ is a modulus of continuity that depends on v but not on I_k. So the functions f_k have a uniform modulus of continuity and there is a subsequence that converges uniformly by Arzela-Ascoli. $\square$

Theorem 4.6. *Let I_k be a sequence of operators uniformly elliptic with respect to some class $\mathcal{L}$ satisfying Assumptions 2.4 and 2.3. There is a subsequence I_{k_j} that converges weakly.*

Proof. In order to make this proof, first we will construct a dense subset of the test functions v of the form

$$v(x) = \begin{cases} p(x) & \text{if } x \in B_r \\ u(x) & \text{if } x \notin B_r. \end{cases} \tag{4.1}$$

This dense subset is constructed as follows. We know that $L^1(\mathbb{R}^n, \omega)$ is separable, so it has a countable dense subset $\{u_i\}$. Quadratic polynomials are a finite dimensional space which also has a countable dense subset $\{p_i\}$. Now, for every positive integer k, we construct the functions

$$v_{k,i_1,i_2}(x) = \begin{cases} p_{i_1}(x) & \text{if } x \in B_{2^{-k}} \\ u_{i_2}(x) & \text{if } x \notin B_{2^{-k}}. \end{cases}$$

So, given any function v as in (4.1) and any $\varepsilon > 0$ we first choose k such that $2^{-k} < r < 2^{-k+1}$, then we choose i_1 and i_2 such that $\|u_{i_2} - u\|_{L^1(\mathbb{R}^n, \omega)} < \varepsilon$, $|D^2 p_{i_1} - D^2 p| < \varepsilon$, $|Dp_{i_1} - Dp| < \varepsilon$ in $B_{2^{-k}}$, and $|p_{i_1} - p| < \varepsilon$ in $B_{2^{-k}}$.

Since the set $\{v_{k,i_1,i_2}\}$ is countable, we can arrange it in a sequence v_i of the form

$$v_i(x) = \begin{cases} p_i(x) & \text{if } x \in B_{r_i} \\ u_i(x) & \text{if } x \notin B_{r_i}. \end{cases}$$

such that for every v as in (4.1), there is a v_i such that

$$\begin{aligned} \|v - v_i\|_{L^1(\mathbb{R}^n, \omega)} &< \varepsilon \\ |v - v_i| &< \varepsilon \quad \text{in } B_{r/2} \\ |Dv - Dv_i| &< \varepsilon \quad \text{in } B_{r/2} \\ |D^2 v - D^2 v_i| &< \varepsilon \quad \text{in } B_{r/2}. \end{aligned} \tag{4.2}$$

By Lemma 4.5, for each v_i, we can take a subsequence I_{k_j} such that $I_{k_j}(v_i)$ converges uniformly in $B_{r_i/2}$.

By a classical diagonal argument, there is a subsequence I_{k_j} such that for every v_i, $I_{k_j}(v_i)$ converges uniformly in $B_{r_i/2}$. We call this limit $I_\infty(v_i, x)$.

If v is any test function, there is an i such that v is very close to v_i in the sense of (4.2). For any $x \in B_{r/2}$, we have

$$\begin{aligned} I_k(v, x) - I_k(v_i, x) &\leq M_{\mathcal{L}}^+(v - v_i, x) \\ &\leq M_{\mathcal{L}}^+((v - v_i)\chi_{B_{r/2}}, x) + M_{\mathcal{L}}^+((v - v_i)(1 - \chi_{B_{r/2}}), x) \end{aligned}$$

Note that $v - v_i$ is a quadratic function in $B_{r/2}(x)$

$$\leq \varepsilon \|I_k\| + \int\limits_{\mathbb{R}^n \setminus B_{r/2}} \delta(v - v_i, x, z) K(z) \, dz$$

where K is the kernel of Assumption 2.3, thus we have

$$\begin{aligned} &\leq \varepsilon \|I_k\| + C \|v - v_i\|_{L^1(\mathbb{R}^n, \omega)} \\ &\leq C\varepsilon \end{aligned}$$

So, by choosing v_i, we can make $I_k(v, x) - I_k(v_i, x)$ as small as we wish in $B_{r/2}$, uniformly in k. So, for j large, $|I_{k_j}(v, x) - I_\infty(v_i, x)| < 2\varepsilon$ and thus $I_{k_j}(v, x)$ is a Cauchy sequence in $L^\infty(B_{r/2})$. We define $I_\infty(v, x)$ to be the uniform limit of this sequence in $B_{r/2}$.

Therefore we have shown that $I_k(v, x)$ converges uniformly to $I_\infty(v, x)$ in $B_{r/2}$. To finish the proof, all we must show is that this operator I_∞ can be extended to a uniformly elliptic operator for all test functions φ. The key is to note that for any two test functions v_1 and v_2 of the form (4.1) we have $M_{\mathcal{L}}^-(v_1 - v_2, x) \leq I_k(v_1 - v_2, x) \leq M_{\mathcal{L}}^+(v_1 - v_2, x)$, and so this inequality passes to the limit and becomes $M_{\mathcal{L}}^-(v_1 - v_2, x) \leq I_\infty(v_1 - v_2, x) \leq M_{\mathcal{L}}^+(v_1 - v_2, x)$. Now, approximating an arbitrary test function φ as in the proof of Lemma 4.1 we get that there is a unique way to extend I_∞ to all test functions φ such that I_∞ is uniformly elliptic respect to $\mathcal{L}$. $\qquad\square$

Lemma 4.7. *For some $\sigma \geq \sigma_0 > 1$ and $\alpha < \sigma_0 - 1$ we consider nonlocal operators I_0, I_1 and I_2 uniformly elliptic respect to $\mathcal{L}_0(\sigma)$. Assume also that the boundary problem*

$$I_0 u = 0 \ \text{in} \ B_1$$
$$u = g \ \text{in} \ \mathbb{R}^n \setminus B_1$$

does not have more than one solution u for any g continuous and $|g(x)| \leq M(|x| + 1)^{1+\alpha}$.

Given $M > 0$, a modulus of continuity ρ and $\varepsilon > 0$, there is an $\eta > 0$ (small) and a $R > 0$ (large) so that if u, v, I_0, I_1 and I_2 satisfy

$$
\begin{aligned}
I_0(v, x) &= 0 & &\text{in } B_1, \\
I_1(u, x) &\geq -\eta & &\text{in } B_1, \\
I_2(u, x) &\leq \eta & &\text{in } B_1, \\
u &= v & &\text{in } \mathbb{R}^n \setminus B_1, \\
\|I_1 - I_0\| &\leq \eta & &\text{in } B_1, \\
\|I_2 - I_0\| &\leq \eta & &\text{in } B_1, \\
|u(x) - u(y)| &\leq \rho(|x - y|) & &\text{for every } x \in B_R \setminus B_1 \text{ and } y \in \mathbb{R}^n \setminus B_1, \\
|u(x)| &\leq M(|x| + 1)^{1+\alpha}.
\end{aligned}
$$

then $|u - v| < \varepsilon$ in B_1.

Proof. Assume the result was false. There there would be a sequence R_k, $I_0^{(k)}$, $I_1^{(k)}$, $I_2^{(k)}$, η_k, u_k, v_k, I_k, f_k such that $R_k \to \infty$, $\eta_k \to 0$, all the assumptions of the lemma are valid but nonetheless $\sup |u_k - v_k| \geq \varepsilon$ in B_1.

Since $I_0^{(k)}$ is a sequence of uniformly elliptic operators, by Lemma 4.6 we can take a subsequence that converges weakly to some nonlocal operator I_0 that is also elliptic respect to the same class $\mathcal{L}$. Moreover, since $\left\|I_0^{(k)} - I_1^{(k)}\right\| \to 0$ and $\left\|I_0^{(k)} - I_2^{(k)}\right\| \to 0$, then $I_1^{(k)}$ and $I_0^{(k)}$ also converge weakly to I_0.

Since both u_k and v_k have a modulus of continuity ρ on ∂B_1, then by Lemma 3.6 there is a modulus of continuity $\tilde{\rho}$ that extends to the full unit ball $\overline{B}_1$. So u_k and v_k have a uniform modulus of continuity on B_{R_k} with $R_k \to \infty$. We can find a subsequence that converges uniformly on compact sets, thus converging almost everywhere in R^n. By dominated convergence theorem, u_k and v_k converge in $L^1(\mathbb{R}^n, \omega)$ to some function g.

Let u and v be the limits of u_k and v_k respectively. Since u_k and v_k converge to u and v uniformly in B_1 and $\sup |u_k - v_k| \geq \varepsilon$ in B_1, then u and v must be different. We use Lemma 4.3 and we get that u and v solve the same equation $I_0(u, x) = I_0(v, x) = 0$. But then $u = v$ which is a contradiction. $\qquad\square$

Remark 4.8. We will apply Lemma 4.7 to I_0 translation invariant in which case the uniqueness assumption for the boundary problem was proved in [2].

We also present the following simplified version of Lemma 4.7 which will not be used in this paper but it is interesting by itself. The difference with respect to Lemma 4.7 is that in the following version we fix the boundary value g and in exchange we do not require a modulus of continuity in $B_R \setminus B_1$ but only on ∂B_1.

Lemma 4.9. *For some $\sigma \geq \sigma_0 > 1$ and $\alpha < \sigma_0 - 1$ we consider nonlocal operators I_0, I_1 and I_2 uniformly elliptic respect to $\mathcal{L}_0(\sigma)$. Assume also that the boundary problem*

$$I_0 u = 0 \ in \ B_1$$
$$u = g \ in \ \mathbb{R}^n \setminus B_1$$

does not have more than one solution u for some given boundary condition g.

Assume g is continuous on ∂B_1. Given $M > 0$ and $\varepsilon > 0$, there is an $\eta > 0$ (small) so that if u, v, I_0, I_1 and I_2 satisfy

$$
\begin{aligned}
I_0(v, x) &= 0 & &in \ B_1, \\
I_1(u, x) &\geq -\eta & &in \ B_1, \\
I_2(u, x) &\leq \eta & &in \ B_1, \\
u &= v = g & &in \ \mathbb{R}^n \setminus B_1, \\
\|I_1 - I_0\| &\leq \eta & &in \ B_1, \\
\|I_2 - I_0\| &\leq \eta & &in \ B_1,
\end{aligned}
$$

then $|u - v| < \varepsilon$ in B_1 (in this case η depends on g).

Proof. The proof follows the same line as the one of Lemma 4.7. We do it by contradiction assuming the result was false. There there would be a sequence R_k, $I_0^{(k)}$, $I_1^{(k)}$, $I_2^{(k)}$, η_k, u_k, v_k, I_k, f_k such that $\eta_k \to 0$, all the assumptions of the lemma are valid but nonetheless $\sup |u_k - v_k| \geq \varepsilon$ in B_1.

The functions u_k and v_k have a fixed value g outside B_1. Since g is continuous on ∂B_1, by Lemma 3.6, u_k and v_k are equicontinuous in $\overline{B_1}$. So there is a subsequence that converges uniformly in $\overline{B_1}$.

We continue as in the proof of Lemma 4.7. We can take a subsequence such that $I_1^{(k)}$ and $I_0^{(k)}$ converge weakly to I_0. Let u and v the limits of u_k and v_k respectively, so we know that $\sup |u_k - v_k| \geq \varepsilon$ in B_1. But by Lemma 4.3, u and v must solve the same equation $I_0(u, x) = I_0(v, x) = 0$. But then $u = v$ which is a contradiction. $\square$

5 $C^{1,\alpha}$ regularity for variable coefficient equations

The main purpose of this section is to prove a $C^{1,\alpha}$ estimate for nonlocal equations that are not necessarily translation invariant. We will give the result in its most general form, and in the next section we will provide a variety of applications.

It is important to stress the concept of scaling. Since our proofs will involve successive rescaling, we will need some kind of scale invariance. It is not necessary to require a particular equation to be scale invariant, but instead we will consider our equations within a whole class of equations that is scale invariant for which our regularity result up to the boundary is supposed to apply, and we will jump freely from equation to equation in our rescaling process. We say that the class

$\mathcal{L}$ has scale σ if every time the integro-differential operator with kernel $K(y)$ is in $\mathcal{L}$, then the one with kernel $\lambda^{n+\sigma}K(\lambda y)$ is also in $\mathcal{L}$ for any $\lambda < 1$. For example the class $\mathcal{L}_0$ defined in (2.3) has scale σ, whereas the class $\mathcal{L}_*$ defined in (2.6)-(2.7) does not because of condition (2.7).

It is a straight forward computation to check that if $\mathcal{L}$ has scale σ and u solves an equation $I(u, x) = f(x)$ that is elliptic with respect to $\mathcal{L}$, then the function $w(x) = \mu u(\lambda x)$ solves a uniformly elliptic equation with respect to the same class $\mathcal{L}$, $I_{\mu,\lambda}(w, x) = \lambda^\sigma \mu f(\lambda x)$. For example, if $I(u, x)$ corresponds to an integro-differential operator

$$I(u, x) = \int_{\mathbb{R}^n} (u(x+y) + u(x-y) - 2u(x))K(y)\, dy$$

then $I_{\mu,\lambda}$ corresponds to the integro-differential operator

$$I_{\mu,\lambda}(u, x) = \int_{\mathbb{R}^n} (u(x+y) + u(x-y) - 2u(x))\lambda^{n+\sigma} K(\lambda y)\, dy.$$

Note that the coefficient μ does not have any effect on a linear operator.

We will call $\mathcal{L}_1$ to the largest scale invariant class contained in the class $\mathcal{L}_*$ defined in (2.6)-(2.7). This is the class of integro-differential operators with kernels K such that

$$(2-\sigma)\frac{\lambda}{|y|^{n+\sigma}} \leq K(y) \leq (2-\sigma)\frac{\Lambda}{|y|^{n+\sigma}} \tag{5.1}$$

$$|\nabla K(y)| \leq C|y|^{-n-\sigma-1} \qquad \text{in } \mathbb{R}^n \setminus \{0\}. \tag{5.2}$$

From Theorem 2.7, an equation $I_0(u, x) = 0$ has interior $C^{1,\bar{\alpha}}$ estimates for some $\bar{\alpha} > 0$ if I_0 is uniformly elliptic respect to the class $\mathcal{L}_1$.

Our main theorem in this section states essentially that if an equation $I^{(0)}(u, x) = 0$ is uniformly elliptic respect to a scale invariant class with interior $C^{1,\bar{\alpha}}$ estimates and we have another equation $I(u, x) = f(x)$ for which the operators $I(-, x)$ stay close to $I^{(0)}$, then this equation also has interior $C^{1,\alpha}$ estimates for any $\alpha < \min(\bar{\alpha}, \sigma - 1)$. Since our proof involves successive rescaling, we need to measure *closeness* at every scale, so we define a distance in between operators that takes scaling of order σ into account.

Definition 5.1. *Given $\sigma \in (0, 2)$ and an operator I, we define the rescaled operator as above*

$$I_{\mu,\lambda}(w, x) = \lambda^\sigma \mu I(\mu^{-1}w(\lambda^{-1}-), \lambda x)$$

The norm of scale σ is defined as

$$\left\|I^{(1)} - I^{(2)}\right\|_\sigma = \sup_{\lambda < 1} \left\|I^{(1)}_{1,\lambda} - I^{(2)}_{1,\lambda}\right\|$$

where $\|.\|$ is the norm defined in Definition 2.2.

The purpose of the rescaled operator is to stress that if u solves the equation $I(u, x) = f(x)$ in B_λ, then the rescaled function $w(x) = \mu u(\lambda x)$ solves an equation of the same ellipticity type $I_{\mu,\lambda}(w, x) = \lambda^\sigma f(\lambda x)$ in B_1.

The following is the main theorem of this work.

Theorem 5.2. *Assume $\sigma > \sigma_0 > 1$. Let $I^{(0)}$ be a fixed translation invariant nonlocal operator in a class $\mathcal{L} \subset \mathcal{L}_0(\sigma)$ with scale σ and interior $C^{1,\bar{\alpha}}$ estimates (for example $\mathcal{L} = \mathcal{L}_1$).*

Let $\mathrm{I}^{(1)}$ and $\mathrm{I}^{(2)}$ be two nonlocal operators, elliptic with respect to $\mathcal{L}_0(\sigma)$, and assume that

$$\left\| \mathrm{I}^{(0)} - \mathrm{I}^{(j)} \right\|_\sigma < \eta$$

for some $\eta > 0$ small enough and $j = 1, 2$.

Let u be a bounded function that solves the equation

$$\mathrm{I}^{(1)}(u, x) \geq f_1(x) \qquad in \ B_1$$
$$\mathrm{I}^{(2)}(u, x) \leq f_2(x) \qquad in \ B_1$$

for a couple of bounded functions f_1 and f_2.

Then $u \in C^{1,\alpha}(B_{1/2})$ for any $\alpha < \min(\bar{\alpha}, \sigma_0 - 1)$ and

$$\|u\|_{C^{1,\alpha}(B_{1/2})} \leq C(\|u\|_{L^\infty(\mathbb{R}^n)} + \|f_1\|_{L^\infty(B_1)} + \|f_2\|_{L^\infty(B_1)})$$

and the estimate depends only on σ_0, λ, Λ and dimension, but not on σ.

The proof of this theorem is based on an approximation argument inspired in [3] using Lemma 4.7.

Proof. By scaling the problem, we can assume without loss of generality that

$$\|f_1\|_{L^\infty(B_1)} < \eta,$$
$$\|f_2\|_{L^\infty(B_1)} < \eta,$$
$$\|u\|_{L^\infty(\mathbb{R}^n)} \leq 1,$$

and u solves the equation in some large ball B_{2R}.

We will show that there is a $\lambda > 0$ and a sequence of linear functions

$$l_k(x) = a_k + b_k \cdot x,$$

such that

$$\sup_{B_{\lambda^k}} |u - l_k| \leq \lambda^{k(1+\alpha)} \tag{5.3}$$

$$|a_{k+1} - a_k| \leq \lambda^{k(1+\alpha)} \tag{5.4}$$

$$\lambda^k |b_{k+1} - b_k| \leq C_2 \lambda^{k(1+\alpha)} \tag{5.5}$$

It is a standard computation that these conditions imply that u is $C^{1,\alpha}$ at the origin (see [3]).

Let $l_0 = 0$. We proceed by induction. Assume we have (5.3)-(5.4)-(5.5) up to some value of k. We will now show it for $k + 1$. Consider

$$w_k(x) = \frac{[u - l_k](\lambda^k x)}{\lambda^{k(1+\alpha)}},$$

then since $\mathcal{L}$ has scale σ, w_k solves an equation of the same ellipticity type

$$\mathrm{I}_k^{(1)}(w_k, x) = \mathrm{I}_{\lambda^{-k(1+\alpha)}, \lambda^k}^{(1)}(w_k, x) \geq \lambda^{k(\sigma-1-\alpha)} f_1(\lambda^k x) \qquad (\text{and} \ \leq \ \text{for} \ \mathrm{I}_k^{(2)} \ \text{and} \ f_2 \ )$$

Since $\sigma - 1 - \alpha \geq \sigma_0 - 1 - \alpha > 0$, the right hand side becomes smaller as k increases. Moreover,

$$\left\| \mathrm{I}_k^{(j)} - \mathrm{I}_k^{(0)} \right\| \leq \left\| \mathrm{I}^{(j)} - \mathrm{I}^{(0)} \right\|_\sigma < \eta.$$

19

By the inductive hypothesis, we have that $|w_k| \le 1$ in B_1. Let $\alpha < \alpha_1 < \min(\bar{\alpha}, \sigma_0)$, we will show that we can construct the sequence l_k so that we also have

$$w_k(x) \le |x|^{1+\alpha_1} \qquad \text{if } x \notin B_1$$

We choose R from Lemma 4.7, so that we have that w_k is Hölder continuous in $B_R \setminus B_1$ and we can apply Lemma 4.7 to the function h that solves

$$\begin{aligned} \mathrm{I}_k^{(0)} h &= 0 && \text{in } B_1 \\ h(x) &= w_k(x) && \text{in } \mathbb{R}^n \setminus B_1 \end{aligned}$$

then $|w_k - h| < \varepsilon(\eta, R)$ in B_1.

Since $\mathrm{I}_k^{(0)}$ is translation invariant and elliptic with respect to $\mathcal{L}$, it has interior $C^{1,\bar{\alpha}}$ estimates. Let $\bar{l} = \bar{a} + \bar{b} \cdot x$ be the linear part of h at the origin. Since $|h| \le 1 + \varepsilon(\eta, R)$ in B_1, $|\bar{a}| \le 1 + \varepsilon(\eta, R)$. By the $C^{1,\bar{\alpha}}$ estimates, $\bar{b} \le C_2$. We also have that $|h(x) - \bar{l}(x)| \le C_3|x|^{1+\bar{\alpha}}$ in $B_{1/2}$. So, we have the following estimates:

$$|w_k(x) - \bar{l}(x)| \le \varepsilon(\eta, R) + C_3 \lambda^{1+\bar{\alpha}} \qquad\qquad \text{in } B_{1/2} \tag{5.6}$$

$$|w_k(x) - \bar{l}(x)| \le \varepsilon(\eta, R) + 2 + C_2 \qquad\qquad \text{in } B_1 \setminus B_{1/2} \tag{5.7}$$

$$|w_k(x) - \bar{l}(x)| \le \varepsilon(\eta, R) + 1 + |x|^{1+\alpha_1} + C_2|x| \qquad \text{in } \mathbb{R}^n \setminus B_1 \tag{5.8}$$

We will choose λ small to be determined below, and then η and R so that $\varepsilon(\eta, R) \le \lambda^{1+\bar{\alpha}}$. The question is how to pick λ so that we can complete the inductive step. We have

$$l_{k+1}(x) = l_k(x) + \lambda^{k(1+\alpha)} \bar{l}\left(\frac{x}{\lambda^k}\right),$$

$$w_{k+1} = \frac{(u - l_{k+1})(\lambda^k x)}{\lambda^{k(1+\alpha)}} = \frac{(w_k - \bar{l})(\lambda x)}{\lambda^{1+\alpha}}.$$

Let us analyze the bounds (5.6), (5.7) and (5.8) in the rescaled setting. We get

$$|w_{k+1}(x)| \le \lambda^{\bar{\alpha}-\alpha} + C_3 \lambda^{\bar{\alpha}-\alpha} \qquad\qquad \text{in } B_{\lambda^{-1}/2}$$

$$|w_{k+1}(x)| \le \lambda^{\bar{\alpha}-\alpha} + (2 + C_2)\lambda^{-(1+\alpha)} \qquad\qquad \text{in } B_1 \setminus B_{\lambda^{-1}/2}$$

$$|w_{k+1}(x)| \le \lambda^{\bar{\alpha}-\alpha} + 1 + |\lambda|^{\alpha_1-\alpha}|x|^{1+\alpha_1} + C_2|\lambda|^{-\alpha}|x| \qquad \text{in } \mathbb{R}^n \setminus B_{\lambda^{-1}}$$

We start by choosing λ small enough so that $(1 + C_3)\lambda^{\bar{\alpha}-\alpha} < 1$ and we have $|w_{k+1}| \le 1$ in B_1. If we choose λ small enough, we also have the following estimates in each of the three rings

$$|w_{k+1}(x)| \le \lambda^{\bar{\alpha}-\alpha} + C_3 \lambda^{\bar{\alpha}-\alpha} \le 1 \le |x|^{1+\alpha_1} \qquad\qquad \text{in } B_{\lambda^{-1}/2} \setminus B_1$$

$$|w_{k+1}(x)| \le \lambda^{\bar{\alpha}-\alpha} + (2 + C_2)\lambda^{-(1+\alpha)} \le C\lambda^{\alpha_1-\alpha}\lambda^{-(1+\alpha_1)} \le |x|^{1+\alpha_1} \qquad \text{in } B_{\lambda^{-1}} \setminus B_{\lambda^{-1}/2}$$

$$|w_{k+1}(x)| \le \lambda^{\bar{\alpha}-\alpha} + 1 + |\lambda|^{\alpha_1-\alpha}|x|^{1+\alpha_1} + C_2|\lambda|^{-\alpha}|x| \le |x|^{1+\alpha_1} \qquad \text{in } \mathbb{R}^n \setminus B_{\lambda^{-1}}$$

so that $|w_{k+1}| < |x|^{1+\alpha_1}$ outside B_1. Moreover we observe that

$$|a_{k+1} - a_k| \le \lambda^{k(1+\alpha)}$$

$$\lambda^k |b_{k+1} - b_k| \le C_2 \lambda^{k(1+\alpha)}$$

where C_2 is the same as above and depends only on the $C^{1,\alpha}$ estimates. We also have $u(x) - l_{k+1}(x) = \lambda^{k(1+\alpha)}\left(w(\frac{x}{\lambda^k}) - \bar{l}(\frac{x}{\lambda^k})\right)$, thus

$$|u(x) - l_{k+1}(x)| \le \lambda^{(k+1)(1+\alpha)} \qquad \text{in } B_{\lambda^{k+1}}.$$

This completes the inductive step and the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

6 Applications

Since the main theorem in section 5 was stated in its more general way which is somewhat abstract, we provide concrete applications in this section aimed at showing the strength of this result.

6.1 Linear equations with variable coefficients

We start by stating the result of the theorem in the simplest case when the equation is linear and close to an operator in $\mathcal{L}_1$. This is the integral version of the classical Cordes-Nirenberg theorem.

Theorem 6.1. *Let $\sigma > 1$ and let u be a bounded function that solves the equation*

$$\mathrm{I}(u, x) := \int_{\mathbb{R}^n} \delta u(x, y)(2 - \sigma) \frac{a(x, y)}{|y|^{n+\sigma}} \, \mathrm{d}y = f(x) \qquad \text{in } B_1$$

Assume that $f \in L^\infty(\mathbb{R}^n)$ and $|a(x, y) - a_0(y)| < \eta$ for every $x \in B_1$ and some small $\eta > 0$. Assume a_0 is a bounded function such that $K(y) = (2 - \sigma)a_0(y)/|y|^{n+\sigma}$ satisfies (5.1) and (5.2). Then for any $\alpha < \sigma - 1$, $u \in C^{1,\alpha}(B_{1/2})$ if η is small enough and

$$\|u\|_{C^{1,\alpha}(B_{1/2})} \leq C(\|u\|_{L^\infty(\mathbb{R}^n)} + \|f\|_{L^\infty(B_1)}).$$

Proof. We apply Theorem 5.2. In this case $\mathrm{I}^{(0)}$ is given by

$$\mathrm{I}^{(0)}(u, x) = \int_{\mathbb{R}^n} \delta u(x, y) \frac{(2 - \sigma)a_0(y)}{|y|^{n+\sigma}} \, \mathrm{d}y$$

By assumption, this operator $\mathrm{I}^{(0)}$ belongs to $\mathcal{L}_1$, which is a scale invariant class with interior $C^{1,\alpha}$ estimates. In this case it is easy to see that since the equation is linear and the coefficients do not depend on x, the derivatives of the solution solve the same equation and the solutions are actually $C^{2,\alpha}$. So in particular $\mathrm{I}^{(0)}$ has interior $C^{1,1}$ estimates.

The weight ω must be a weight that controls the tails of the kernel. In this case a natural choice is

$$\omega(y) = \frac{1}{(1 + |y|)^{n+\sigma}}.$$

Now we estimate $\left\|\mathrm{I} - \mathrm{I}^{(0)}\right\|_\sigma$,

$$\left\|\mathrm{I} - \mathrm{I}^{(0)}\right\|_\sigma = \sup_{\lambda < 1} \left\|\mathrm{I}_{1,\lambda} - (\mathrm{I}^{(0)})_{1,\lambda}\right\|.$$

We compute

$$(\mathrm{I}_{1,\lambda} - (\mathrm{I}^{(0)})_{1,\lambda})(u, x) = \int_{\mathbb{R}^n} \delta u(x, y) \frac{(2 - \sigma)(a(\lambda x, \lambda y) - a_0(\lambda y))}{|y|^{n+\sigma}} \, \mathrm{d}y$$

with $|a(\lambda x, \lambda y) - 1| < \eta$. From Definition 2.2

$$\left\|\mathrm{I}_{1,\lambda} - (\mathrm{I}^{(0)})_{1,\lambda}\right\| = \sup_{u \text{ as in Definition 2.2}} \frac{\mathrm{I}_{1,\lambda}(u, x) - (\mathrm{I}^{(0)})_{1,\lambda}(u, x)}{1 + M}$$

$$\leq \sup_{u \text{ as in Definition 2.2}} \frac{1}{1 + M} \int_{\mathbb{R}^n} |\delta u(x, y)| \frac{(2 - \sigma)\eta}{|y|^{n+\sigma}} \, \mathrm{d}y$$

$$(6.1)$$

In Definition 2.2 we take functions u such that $\|u\|_{L^1(\mathbb{R}^n,\omega)}$ and $|u(y)-u(x)-(y-x)\cdot\nabla u(x)| \leq M|x-y|^2$ for $y \in B_1(x)$. Note that these two conditions together imply $|u(x)| \leq CM$ for some universal constant C.

From $|u(y)-u(x)-(y-x)\cdot\nabla u(x)| \leq M|x-y|^2$, we obtain

$$\int_{B_1} |\delta u(x,y)| \frac{(2-\sigma)}{|y|^{n+\sigma}} \, dy \leq \int_{B_1} M|y|^2 \frac{(2-\sigma)}{|y|^{n+\sigma}} \, dy$$
$$\leq CM$$

On the other hand, from $\|u\|_{L^1(\mathbb{R}^n,\omega)}$ and $|u(x)| \leq CM$ we bound the tail of the integral

$$\int_{\mathbb{R}^n\setminus B_1} |\delta u(x,y)| \frac{(2-\sigma)}{|y|^{n+\sigma}} \, dy \leq \int_{\mathbb{R}^n\setminus B_1} |u(x+y)| + |u(x-y)| + |u(x)| \frac{(2-\sigma)}{|y|^{n+\sigma}} \, dy$$
$$\leq CM$$

Replacing in (6.1), $\left\|I_{1,\lambda} - I_{1,\lambda}^{(0)}\right\| \leq C\eta$ for any λ, thus $\left\|I - I^{(0)}\right\|_\sigma \leq C\eta$. So if we choose η small enough we can apply Theorem 5.2 and conclude that the equation $I(u,x) = f$ has interior $C^{1,\alpha}$ estimates for any $\alpha < \sigma - 1$. $\qquad\square$

It is interesting to note in the theorem above that if we fix an $\alpha > 0$, the constant C in Theorem 6.1 does not blow up as $\sigma \to \infty$.

6.2 Nonlinear equations with variable coefficients

By combining Theorem 5.2 with Theorem 2.7 we obtain the following result

Theorem 6.2. *Let $\sigma \geq \sigma_0 > 1$ and let $\mathcal{L}_1(\sigma)$ be the class defined in (5.1)–(5.2).*

Assume $I^{(0)}$ is translation invariant nonlocal operator uniformly elliptic respect to $\mathcal{L}_1(\sigma)$ and I uniformly elliptic respect to $\mathcal{L}_0$. Assume $\left\|I(-,x) - I^{(0)}\right\|_\sigma < \eta$ for every $x \in B_1$ and some small $\eta > 0$. Let $f \in L^\infty(B_1)$ and u be a bounded function that solves the equation

$$I(u,x) = f(x) \qquad in \ B_1.$$

Then $u \in C^{1,\alpha}(B_{1/2})$ for some small $\alpha > 0$ and

$$[u]_{C^{1,\alpha}(B_{1/2})} \leq C(\|u\|_{L^\infty(\mathbb{R}^n)} + \|f\|_{L^\infty(B_1)}).$$

and the estimate depends only on σ_0, λ, Λ, C_1 and dimension, but not on σ.

A concrete example of a nonlinear operator I such that Theorem 6.2 applies would be

$$I(u,x) = \inf_\alpha \sup_\beta \int_{\mathbb{R}^n} (u(x+y) + u(x-y) - 2u(x)) \frac{(2-\sigma)(a_0(y) + a_{\alpha\beta}(x,y))}{|y|^{n+\sigma}} \, dy$$

such that we have for some small η,

$$|a_{\alpha\beta}(x,y)| < \eta \qquad \text{for every } \alpha, \beta$$
$$\lambda \leq a_0(y) \leq \Lambda$$
$$|\nabla a_0(y)| \leq C|y|^{-1}$$

Of course, a regularity result that requires coefficients not to vary much with respect to x implies a regularity estimate every time the coefficients are continuous in x. We state it in the following theorem.

Theorem 6.3. *Given $\sigma \geq \sigma_0 > 1$, let $\mathrm{I}(u, x)$ be given by*

$$\mathrm{I}(u, x) = \inf_\alpha \sup_\beta \int_{\mathbb{R}^n} \delta u(x, y) \frac{(2 - \sigma) a_{\alpha\beta}(x, y)}{|y|^{n+\sigma}} \, \mathrm{d}y$$

such that for every α, β we have $\lambda < a_{\alpha\beta}(x, y) < \Lambda$ and $\nabla_y a_{\alpha\beta}(x, y) \leq C_1/((2 - \sigma)|y|)$ and also $|a_{\alpha\beta}(x_1, y) - a_{\alpha\beta}(x_2, y)| < c(|x_1 - x_2|)$ for some uniform modulus of continuity c. Then solutions u to the equation

$$\mathrm{I}(u, x) = f \qquad in \ \Omega$$

are of class $C^{1,\alpha}(B_{1/2})$ for some small $\alpha > 0$ and

$$[u]_{C^{1,\alpha}(B_{1/2})} \leq C(\|u\|_{L^\infty(\mathbb{R}^n)} + \|f\|_{L^\infty(B_1)}).$$

and the estimate depends only on σ_0, λ, Λ, C_1, the modulus of continuity c and dimension, but not on σ.

Proof. Around each point $x_0 \in B_{1/2}$ we can find a ball $B_r(x_0)$ so that $|a_{\alpha\beta}(x, y) - a_{\alpha\beta}(x_0, y)| < c(r) \leq \eta$. This implies that $\|\mathrm{I}(-, x) - \mathrm{I}(-, x_0)\|_\sigma < C\eta$ as in the proof of Theorem 6.1, and we apply Theorem 6.2 with $\mathrm{I}^{(0)} = \mathrm{I}(-, x_0)$, scaled in $B_r(x)$. □

6.3 Nonlinear equations with constant coefficients but non-differentiable kernels

It is interesting to note that Theorem 5.2 is useful to obtain new results even in the translation invariant case. In Theorem 2.7, it is required that every kernel must be differentiable away from the origin. That assumption can be relaxed in the following way. We can actually obtain a $C^{1,\alpha}$ estimate for nonlocal equations that are uniformly elliptic respect to the class $\mathcal{L}$ given by operators with kernel K such that

$$K(y) = (2 - \sigma) \frac{a_1(y) + a_2(y)}{|y|^{n+\sigma}}$$

$$\lambda \leq a_1(y) \leq \Lambda$$

$$|a_2| \leq \eta$$

$$|\nabla a_1(y)| \leq \frac{C_1}{|y|} \qquad in \ \mathbb{R}^n \setminus \{0\}$$

Theorem 6.4. *Assume $\sigma \geq \sigma_0 > 1$. If η is small enough (depending on λ, Λ, C_1 and dimension) then the nonlocal equation*

$$\mathrm{I}(u, x) = \inf_\alpha \sup_\beta \int_{\mathbb{R}^n} (u(x + y) + u(x - y) - 2u(x)) K_{\alpha\beta}(y) \, \mathrm{d}y = f(x)$$

has interior $C^{1,\alpha}$ estimates if all kernels belong to the class $\mathcal{L}$ above.

Proof. Let $L \in \mathcal{L}$ be an operator with kernel K, we can write it as $K = K^1 + K^2$ where $K^1 = (2 - \sigma)\frac{a_1(y)}{|y|^{n+\sigma}}$ and correspondingly $L = L^1 + L^2$. But then we see that $\left\|L - L^1\right\|_\sigma < c\eta$. So we write

$$\mathrm{I}^{(0)}(u, x) = \inf_\alpha \sup_\beta L^1_{\alpha\beta} u(x)$$

and we have $\left\|\mathrm{I} - \mathrm{I}^{(0)}\right\|_\sigma < c\eta$ so that we can apply Theorem 5.2. □

This theorem applies for a class that is still smaller than $\mathcal{L}_0$. It would be interesting to determine whether the class $\mathcal{L}_0$ has interior $C^{1,\alpha}$ estimates or not. We leave that question open.

6.4 Nonlinear equations close to the fractional Laplacian. Very regular solutions.

We can obtain another new result in the translation invariant case from Theorem 5.2. In this case we look at non linear equations that are sufficiently close to the fractional Laplacian and we obtain $C^{2,\alpha}$ regularity. We provide a theorem on the regularity of fully nonlinear translation invariant nonlocal equations when the ellipticity constants are sufficiently close to each other. This is an improvement on Theorem 2.7 when these conditions are satisfied.

Theorem 6.5. *Assume $\sigma > \sigma_0 > 1$. There is a $\eta > 0$ and $\rho_0 > 0$ so that if $\Lambda < 1 + \eta$, $\lambda > 1 - \eta$, I is a nonlocal translation invariant uniformly elliptic operator with respect to $\mathcal{L}_*$ (given in (2.6)-(2.7)) and u is a continuous function in $\overline{B}_1$ such that $u \in L^1(\mathbb{R}^n, \omega)$, $Iu = 0$ in B_1, then there is a universal (depends only on n and σ_0) $\alpha > 0$ such that $u \in C^{2+\alpha}(B_{1/2})$ and*

$$u_{C^{2+\alpha}(B_{1/2})} \leq C \left(\sup_{B_1} |u| + \|u\|_{L^1(\mathbb{R}^n, \omega)} + |I0| \right)$$

for some constant $C > 0$ (where by $I0$ we mean the value we obtain when we apply I to the function that is constant equal to zero). The constant C depends on σ_0, n and the constant in (2.7).

Proof. From Theorem 2.7, we know that $u \in C^{1,\alpha}(B_{3/4})$. In particular the function is differentiable. Let $w = u_e$ be a directional derivative. We proceed as in the proof of Theorem 2.7 (Theorem 12.1 in [2]) and write $w = w_1 + w_2$ where w_2 vanishes in $B_{5/8}$ and w_1 solves

$$\begin{aligned} \mathrm{M}^+_{\mathcal{L}_*} w_1 &\geq -C & &\text{in } B_{5/8} \\ \mathrm{M}^-_{\mathcal{L}_*} w_1 &\leq C & &\text{in } B_{5/8} \end{aligned}$$

The difference with the proof of Theorem 12.1 in [2] is that now we apply Theorem 5.2 instead of Theorem 2.6. Since $\lambda > 1 - \eta$ and $\Lambda < 1 + \eta$, then $\left\| \mathrm{M}^+_{\mathcal{L}_*} w_1 + (-\triangle)^\sigma \right\| < c\eta$ and $\left\| \mathrm{M}^-_{\mathcal{L}_*} w_1 + (-\triangle)^\sigma \right\| < c\eta$. So Theorem 5.2 tells us that $w = u_e$ is $C^{1,\alpha}$ in $B_{1/2}$, and thus $u \in C^{2,\alpha}(B_{1/2})$. $\qquad \square$

Theorem 5.2 is also useful to reduce the equation to its leading order part. A simple example would be the following.

Theorem 6.6. *Let $\sigma \in (1,2)$ and $r > 0$. Let u be a bounded function that solves the equation*

$$I(u,x) = -(-\triangle)^{\sigma/2} u(x) + \inf_\alpha \sup_\beta \int_{\mathbb{R}^n} \delta u(x,y)(2-\sigma) \frac{|y|^r a_{\alpha\beta}(y)}{|y|^{n+\sigma}} \, \mathrm{d}y = 0 \qquad \text{in } B_1$$

Assume that $f \in L^\infty(\mathbb{R}^n)$, $r > 0$, $|a(y)| < C$, and $|\nabla a(y)| \leq C/|y|$. Then $u \in C^{2,\alpha}(B_{1/2})$ for some $\alpha > 0$ and

$$\|u\|_{C^{2,\alpha}(B_{1/2})} \leq C \|u\|_{L^\infty(\mathbb{R}^n)}.$$

Proof. First of all let us notice that we can place the fractional Laplacian inside the inf sup and rewrite the equation as

$$I(u,x) = \inf_\alpha \sup_\beta \int_{\mathbb{R}^n} \delta u(x,y)(2-\sigma) \frac{c_\sigma + |y|^r a_{\alpha\beta}(y)}{|y|^{n+\sigma}} \, \mathrm{d}y = 0 \qquad \text{in } B_1$$

where c_σ is the normalization constant of the fractional Laplacian which remains bounded below and above for $\sigma \in (1,2)$.

When we make a rescaling of the equation of order σ we obtain

$$\mathrm{I}_{1,\lambda} = \inf_\alpha \sup_\beta \int_{\mathbb{R}^n} \delta u(x,y)(2-\sigma) \frac{1 + \lambda^r |y|^r a_{\alpha\beta}(x,\lambda y)}{|y|^{n+\sigma}} \ dy.$$

Then, for λ small enough, we will have $\left\| \mathrm{I}_{1,\lambda} + (-\triangle)^{\sigma/2} \right\|_\sigma < \eta$ and we proceed as in the proof of Theorem 6.5. $\quad\square$

Appendix

In this appendix, we prove Lemma 3.1. We can give an intuitive proof avoiding lengthy computations by using Lemma 4.3. We notice that we did not use Lemma 3.1 to prove Lemma 4.3.

Proof of Lemma 3.1. For some $r > 0$, we will show that if we consider $u_\alpha(x) = ((|x|-1)^+)^\alpha$ and $x_0 = (1+r)e_1$, then $M^+u(x_0) \le 0$ if α is small enough. By rotation this implies the result for any x such that $|x| = 1+r$. If $1 < |x| < 1+r$, we can still get the result by a scaling centered at e_1 (i.e. by considering $\tilde{u}(x) = ((|x+\lambda e_1| - (1+\lambda))^+)^\alpha$ for λ such that $|x_0| - 1 = r(1+\lambda)$), since that scaling would only decrease the values of the function elsewhere. This observation implies also that if $M^+u(x_0) \ge 0$, then also $M^+u(x) \ge 0$ for all x such that $|x| \le |x_0|$.

Let M be a large number and let v_j be

$$v_\alpha(x) = \max\left(\frac{u_\alpha(x) - 1}{\alpha}, -M \right).$$

The key observation is the following

$$\lim_{\alpha \to 0} v_\alpha(x) = \max(\log(|x|-1), -M) = l_M(x)$$

and this convergence holds uniformly on compact sets.

As $\sigma \to 2$, M_σ^+ converges to a uniformly elliptic second order differential operator that is comparable to the maximal Pucci operator with ellipticity constants depending on λ, Λ and dimension, that we call M_2^+. It is a simple computation that $M^+(D^2 \log(|x|-1)) \to -\infty$ as $|x| \to 1$ (where $M^+(D^2u)$ stands for the standard Pucci operator). So if we pick r small enough and M larger than $-\log r$, then there will be a σ_1 such that $M_\sigma^+ l_M < -1$ if $\sigma > \sigma_1$. Moreover, by choosing M large enough, we can also make sure that $M_\sigma^+ l_M < -1$ for $\sigma \le \sigma_1$. This is how we pick r and M.

Assume that the result of the lemma was not true. So there would be a sequence $\alpha_j \to 0$ and $\sigma_j \in [\sigma_0, 2]$ such that $M_{\sigma_j}^+ u_{\alpha_j}(x_0) \ge 0$, and thus $M_{\sigma_j}^+ v_{\alpha_j} \ge 0$ in the annulus $B_{2|x_0|} \setminus B_{|x_0|}$. We extract a subsequence σ_{j_k} that converges to some number $\tilde{\sigma} \in [\sigma_0, 2]$. We apply Lemma 4.3 and we obtain that $M_{\tilde{\sigma}}^+ l_M(x_0) \ge 0$. But this is a contradiction since $M_{\tilde{\sigma}}^+ l_M(x_0) \le -1$. $\quad\square$

References

[1] G. Barles, E. Chasseigne, and C. Imbert. Hölder continuity of solutions of second-order nonlinear elliptic integro-differential equations. *Journal of the European Mathematical Society*, 2007.

[2] L. Caffarelli and L. Silvestre. Regularity theory for fully nonlinear integro-differential equations. *Comm. Pure Appl. Math*, 62(5):597–638, 2009.

[3] L. A. Caffarelli. Interior a priori estimates for solutions of fully nonlinear equations. *Ann. of Math. (2)*, 130(1):189–213, 1989.

[4] L. A. Caffarelli and Xavier Cabré. *Fully nonlinear elliptic equations*, volume 43 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1995.

[5] Heinz Otto Cordes. Über die erste Randwertaufgabe bei quasilinearen Differentialgleichungen zweiter Ordnung in mehr als zwei Variablen. *Math. Ann.*, 131:278–312, 1956.

[6] M. Kassmann. A priori estimates for integro-differential operators with measurable kernels. *Calculus of Variations and Partial Differential Equations*, 34(1):1–21, 2009.

[7] R. Mikulyavichyus and G. Pragarauskas. Nonlinear potentials of the Cauchy-Dirichlet problem for the Bellman integro-differential equation. *Liet. Mat. Rink.*, 36(2):178–218, 1996.

[8] L. Nirenberg. On a generalization of quasi-conformal mappings and its application to elliptic partial differential equations. In *Contributions to the theory of partial differential equations*, Annals of Mathematics Studies, no. 33, pages 95–100. Princeton University Press, Princeton, N. J., 1954.

[9] L. Silvestre. Hölder estimates for solutions of integro-differential equations like the fractional laplace. *Indiana University Mathematics Journal.*, 55(3):1155–1174, 2006.