Annals of Mathematics **174** (2011), 1163–1187
http://dx.doi.org/10.4007/annals.2011.174.2.9

The Evans-Krylov theorem for nonlocal fully nonlinear equations

By Luis Caffarelli and Luis Silvestre

Abstract

We prove a regularity result for solutions of a purely integro-differential Bellman equation. This regularity is enough for the solutions to be understood in the classical sense. If we let the order of the equation approach two, we recover the theorem of Evans and Krylov about the regularity of solutions to concave uniformly elliptic partial differential equations.

1. **Introduction**

In 1982, L. Evans and N. Krylov proved independently ([6] and [7]) the following celebrated interior regularity result for elliptic partial differential equations: If u is a bounded solution to $F(D^2u)$ in B_1, where F is uniformly elliptic and concave, then $u \in C^{2,\alpha}(B_{1/2})$ for some $\alpha > 0$. In this paper we prove a nonlocal version of that theorem. We prove that solutions to concave integro-differential equations of order σ have regularity $C^{\sigma+\alpha}$ for some $\alpha > 0$. This is enough regularity to consider the solutions to be classical.

The equations we study arise in stochastic control problems with jump processes (see for example [9], [8]). In [9] a $C^{2,\alpha}$ regularity of the solutions of Bellman equations for Levy processes is obtained, but the equation is required to have a uniformly elliptic second order part which is ultimately the source of the regularity. In [1] a purely integro-differential Bellman equation is studied. They only consider the case of the maximum of two linear operators. They obtain solutions in the fractional Sobolev space $H^{\sigma/2}$ up to the boundary and H^σ in the interior of the domain. As they point out in their paper, the solutions to these equations are expected to be more regular in the interior of the domain.

In this paper we consider purely integro-differential equations and obtain an interior regularity result. Since we do not require our equations to have a second order part, our estimate comes only from the regularization effects of the integrals.

The constants in our estimates do not blow up as $\sigma \to 2$, so we can recover the usual Evans-Krylov theorem as a limit case. It is interesting to follow what the ideas of the proofs become as $\sigma \to 2$. Interestingly, the ideas we present in this paper provide a different proof of the Evans-Krylov theorem for second order elliptic equations.

We consider the equation

$$Iu(x) := \inf_{a \in \mathcal{A}} L_a u(x) = \inf_{a \in \mathcal{A}} \int_{\mathbb{R}^n} (u(x+y)+u(x-y)-2u(x)) K_a(y) \, \mathrm{d}y = 0. \tag{1.1}$$

As in [3], we will choose each linear operator L_a in some class $\mathcal{L}$. Consequently, the operator I will be elliptic with respect to $\mathcal{L}$ in the sense described in [5].

We describe below the appropriate classes of linear operators that we will use in this paper.

We say that an operator L belongs to $\mathcal{L}_0$ if its corresponding kernel K satisfies the uniform ellipticity assumption

$$(2-\sigma)\frac{\lambda}{|y|^{n+\sigma}} \leq K(y) \leq (2-\sigma)\frac{\Lambda}{|y|^{n+\sigma}}. \tag{1.2}$$

We will also assume that the kernels K in the class $\mathcal{L}_0$ are symmetric: $K(y) = K(-y)$. This assumption is somewhat implicit in the expression (1.1), since all the kernels K_a can be symmetrized without altering the equation.

The ellipticity assumption (1.2) is the essential assumption that leads to a local regularization. Our proofs, as usual, involve an improvement of oscillation of the solution to the equation (or an operator applied to it) in a decreasing sequence of balls around a point in the domain. Since the equations are nonlocal, every argument in our proofs will have to take into account the influence of the values of the solution at points outside those balls. We will often need to say that the part of the integral in (1.1) outside a neighborhood of the origin is a smooth enough function. That is why we define the following classes of smooth kernels.

We say that $L \in \mathcal{L}_1$ if, in addition to (1.2) and symmetry, the kernel K is C^1 away from the origin and satisfies

$$|\nabla K(y)| \leq \frac{C}{|y|^{n+1+\sigma}}.$$

Finally, we say $L \in \mathcal{L}_2$ if the kernel is C^2 away from the origin and satisfies

$$D^2 K(y) \leq \frac{C}{|y|^{n+2+\sigma}}. \tag{1.3}$$

We consider the corresponding maximal operators

$$\begin{aligned}
\mathrm{M}_0^+ u(x) &= \sup_{L \in \mathcal{L}_0} Lu(x) = (2-\sigma)\int_{\mathbb{R}^n} \frac{\Lambda(\delta u(x,y))^+ - \lambda(\delta u(x,y))^-}{|y|^{n+\sigma}} \, \mathrm{d}y, \\
\mathrm{M}_1^+ u(x) &= \sup_{L \in \mathcal{L}_1} Lu(x), \\
\mathrm{M}_2^+ u(x) &= \sup_{L \in \mathcal{L}_2} Lu(x).
\end{aligned}$$

Recall that we write $\delta u(x,y) = (u(x+y)+u(x-y)-2u(x))$ as in [3]. Naturally we have the inequalities $\mathrm{M}_0^+ u \geq \mathrm{M}_1^+ u \geq \mathrm{M}_2^+ u$. The minimal operators M^- are defined likewise.

We do not know a closed form for M_1^+ or M_2^+. Since $\mathcal{L}_2 \subset \mathcal{L}_1 \subset \mathcal{L}_0$, we have the relations $\mathrm{M}_2^+ u \leq \mathrm{M}_1^+ u \leq \mathrm{M}_0^+ u$ and $\mathrm{M}_2^- u \geq \mathrm{M}_1^- u \geq \mathrm{M}_0^- u$.

Our main result states that under the hypothesis that all operators L_a belong to $\mathcal{L}_2$, the solutions are classical in the sense that there is enough regularity so that all integrals are well defined and Hölder continuous.

Theorem 1.1. *Assume every L_a in* (1.1) *belongs to the class $\mathcal{L}_2$. If u is a bounded function in $\mathbb{R}^n$ such that $Iu = 0$ in B_1, then $u \in C^{\sigma+\alpha}(B_{1/2})$. Moreover*,

$$\|u\|_{C^{\sigma+\alpha}(B_{1/2})} \leq C \|u\|_{L^\infty(\mathbb{R}^n)} . \tag{1.4}$$

The interested reader may verify, in following the arguments, that global boundedness of u may be substituted by an appropriate moderated growth at infinity (see remark at the end of this paper).

For values of σ less or equal to 1, this theorem does not provide any improvement with respect to the $C^{1,\alpha}$ estimates in [5]. Thus, for the purpose of proving Theorem 1.1 we will assume σ to be strictly larger than 1 in this paper. The result becomes most interesting when σ is close to 2 and $\sigma+\alpha > 2$.

Note that the result of the theorem remains true if I is convex instead of concave (a sup of linear operators instead of an inf). Indeed, we can transform one situation in the other by considering the equation $-I(-u) = 0$.

In previous papers ([3] and [5]) we started developing the regularity theory for nonlocal equations. In [3] we obtained a nonlocal version of Krylov-Safonov theory with estimates that do not blow up as $\sigma \to 2$. This allowed us to obtain $C^{1,\alpha}$ estimates for general fully nonlinear integro-differential equations that are translation invariant. In [5], we extended those results to variable coefficient equations using perturbative methods. In this paper we use the results in our previous two papers extensively.

2. A regularization procedure

In this section we show a simple technique to approximate uniformly the solutions to the integro-differential equation (1.1) by $C^{2,\alpha}$ functions that solve an approximate equation with the same structure. This procedure works exclusively for integro-differential equations and cannot be done using only second order equations. It makes it unnecessary to use sup- or inf- convolutions and simplifies the technicalities of several proofs. Essentially the idea is that if we prove an estimate assuming the solutions u is $C^{2,\alpha}$ (but the estimate does not depend on the $C^{2,\alpha}$ norm), then we can pass to the limit using this approximation technique to extend the estimate to all viscosity solutions. In this respect, the technicalities in the integro-differential setting simplify very much compared to the second order counterpart.

LEMMA 2.1. *Let u be a continuous function in $\mathbb{R}^n$ solving* (1.1) *in B_1 with every L_a belonging to the class $\mathcal{L}_2$ (resp. $\mathcal{L}_1$ or $\mathcal{L}_0$). There is a sequence of* regularized *equations in the same class*

$$\begin{aligned} I^\varepsilon u^\varepsilon &= \inf_a L_a^\varepsilon u^\varepsilon = 0 && \text{in } B_1, \\ u^\varepsilon &= u && \text{in } \mathbb{R}^n \setminus B_1 \end{aligned}$$

so that the solutions u^ε are $C^{2,\alpha}$ in the interior of B_1 for every $\varepsilon > 0$ and $\lim_{\varepsilon \to 0} u^\varepsilon = u$ uniformly in B_1.

Proof. Let η be a smooth function such that

$$\begin{aligned} 0 \leq \eta \leq 1 &\text{ in } \mathbb{R}^n, \\ \eta = 0 &\text{ in } \mathbb{R}^n \setminus B_1, \\ \eta = 1 &\text{ in } B_{1/2} \end{aligned}$$

and let $\eta_\varepsilon(x) = \eta(x/\varepsilon)$.

Let us consider the following regularized kernels:

$$K_a^\varepsilon(y) = \eta_\varepsilon(y) \lambda \frac{2-\sigma}{|y|^{n+\sigma}} + (1 - \eta_\varepsilon(y)) K_a(y).$$

Correspondingly, we define

$$\begin{aligned} L_a^\varepsilon v &= \int_{\mathbb{R}^n} \delta v(x,y) K_a^\varepsilon(y) \,\mathrm{d}y, \\ I^\varepsilon v &= \inf_a L_a^\varepsilon v. \end{aligned}$$

Note that if $L_a \in \mathcal{L}_i$ then also $L_a^\varepsilon \in \mathcal{L}_i$ for $i = 0, 1, 2$.

Let u^ε be the solution of the following Dirichlet problem:

$$\begin{aligned} I^\varepsilon u^\varepsilon &= \inf_a L_a^\varepsilon u^\varepsilon = 0 && \text{in } B_1, \\ u^\varepsilon &= u && \text{in } \mathbb{R}^n \setminus B_1. \end{aligned}$$

The solution u^ε to this problem is $C^{2,\alpha}$ by Theorem 6.6 in [5].

It is clear that if $v \in C^2(x)$ and $|v(y) - v(x) - (y - x) \cdot \nabla v(x)| \leq M|y - x|^2$ in B_1, then

$$|I^\varepsilon v(x) - Iv(x)| \leq CM\varepsilon^{2-\sigma}$$

so $\|I^\varepsilon - I\| \to 0$ as $\varepsilon \to 0$ (recall $\sigma < 2$), where the norm $\|I^\varepsilon - I\|$ is computed in the sense of Definition 2.2 in [5]. Then, by Lemma 4.9 in [5], u^ε converges to u uniformly in $\overline{B}_1$ as $\varepsilon \to 0$. □

Remark 2.2. The concavity of I is not used in Lemma 2.1. The exact same idea works for equations of the type

$$Iu(x) := \sup_b \inf_a L_{ab} u(x) = \sup_b \inf_a \int_{\mathbb{R}^n} (u(x+y) + u(x-y) - 2u(x)) K_{ab}(y) \,\mathrm{d}y = 0.$$

3. Average of subsolutions is a subsolution

The main ingredient in the Evans-Krylov theorem is the fact that concavity of the equation makes second order incremental quotients subsolutions of the linearized equation. In order to prove that, one first observes that an average of solutions to a concave equation is a subsolution to the same equation.

In this section we prove that also in the nonlocal case the average of subsolutions to a concave equation is a subsolution of the same equation. This is a very straightforward computation if the solutions are classical. Making the proof for viscosity solutions adds a technical difficulty. We can overcome that difficulty by using the approximation technique of Section 2. As the referee pointed out, it would be possible to prove the lemmas in this section directly without using the approximation. On the other hand, we like to use the two lemmas below as the first example on how the approximation can be useful.

Proposition 3.1. *Let u and v be subsolutions of $Iu = 0$ and $Iv = 0$ in a domain Ω, u, v continuous in $\mathbb{R}^n$; then $I(u+v)/2 \geq 0$ in Ω.*

Proof. If $u, v \in C^2$ the proposition follows simply by the concavity of I. So we used the regularization procedure described in Section 2.

Let $I^\varepsilon u^\varepsilon = 0$ and $I^\varepsilon v^\varepsilon = 0$ be the approximate equations of Lemma 2.1. The functions u^ε and v^ε are C^2 so all the integro-differential operators $L_a^\varepsilon u^\varepsilon$ and $L_a^\varepsilon v^\varepsilon$ in the formula for I^ε are well defined and continuous. We can make a direct computation:

$$\begin{aligned} I^\varepsilon (u^\varepsilon + v^\varepsilon)/2 &= \inf_a \frac{L_a^\varepsilon u^\varepsilon + L_a^\varepsilon v^\varepsilon}{2} \\ &\geq \frac{\inf_a L_a^\varepsilon u^\varepsilon + \inf_a L_a^\varepsilon v^\varepsilon}{2} \\ &\geq 0 \qquad \text{in } \Omega. \end{aligned}$$

Since $u^\varepsilon \to u$ and $v^\varepsilon \to v$ uniformly in Ω and $I^\varepsilon \to I$, then $I(u+v)/2 \geq 0$ in Ω by Lemma 4.9 in [5]. ☐

The same idea shows that any average of solutions is a subsolution. In particular we have the following proposition.

Proposition 3.2. *Let u be a solution of $Iu = 0$ in B_1 and η be a mollifier; i.e.,*

(1) $\eta \geq 0$;
(2) $\int \eta = 1$;
(3) $\operatorname{supp} \eta \subset B_\delta$.

We have $I(\eta * u) \geq 0$ in $B_{1-\delta}$.

4. **The linear theory of integro-differential equations**

In this section we present some regularity theorems for linear integro-differential equations with constant coefficients. Naturally in this simple case, we can easily obtain more powerful results than for the nonlinear case. The results we present in this section are just the ones that we will need in the rest of this paper.

THEOREM 4.1. *Let L be an integro-differential operator in the class $\mathcal{L}_1$ with $\sigma \geq \sigma_0 > 1$. Suppose that u is an integrable function in the weighted space $L^1(\mathbb{R}^n, \frac{1}{1+|y|^{n+\sigma}})$ that solves the equation $Lu = 0$ in B_1, then $u \in C^{2,\alpha}(B_{1/2})$, and we have the estimates*

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

The value of the constant C and α depends on n, λ, Λ and σ_0 but not on σ.

Proof. We will prove the *a priori* estimate. The regularity estimate for a weak or viscosity solution follows by mollifying the solution or using the regularization procedure of the previous section.

First we apply Theorem 2.8 in [5] to obtain that $u \in C^{1,\alpha}(B_{3/4})$ and obtain the estimate

$$||u||_{C^1(B_{3/4})} \leq C \|u\|_{L^1(\mathbb{R}^n, \frac{1}{1+|y|^{n+\sigma}})} .$$

The idea is to apply the same $C^{1,\alpha}$ estimate to every directional derivative u_e. Since we do not have an L^∞ estimate of u_e outside of $B_{3/4}$, we have to use our usual *integration by parts* trick. We know that

$$\int_{\mathbb{R}^n} u_e(y) K(x+y) \, \mathrm{d}y = 0.$$

Let η be a smooth cutoff function such that

$$\begin{aligned} &0 \leq \eta \leq 1 && \text{in } \mathbb{R}^n, \\ &\eta = 0 && \text{outside } B_{3/4}, \\ &\eta = 1 && \text{in } B_{5/8}. \end{aligned}$$

We compute

$$\begin{aligned} &\left| \int_{\mathbb{R}^n} u_e(y) \eta(y) K(x+y) \, \mathrm{d}y \right| \\ &\qquad = \left| \int_{\mathbb{R}^n} u_e(y) (\eta(y) - 1) K(x+y) \, \mathrm{d}y \right| \\ &\qquad = \left| \int_{\mathbb{R}^n} u(y) (\eta_e(y) K(x+y) + (\eta(y)-1) K_e(x+y)) \, \mathrm{d}y \right| \\ &\qquad \leq C \|u\|_{L^1(\mathbb{R}^n, \frac{1}{1+|y|^{n+\sigma}})} . \end{aligned}$$

Thus we can apply Theorem 6.1 in [5] to the function ηu_e to conclude that $u_e \in C^{1,\alpha}(B_{1/2}))$ for every direction e and thus $u \in C^{2,\alpha}$. (Note that we are using $\sigma > 1$ here.) $\square$

In order to have better interior regularity estimates than $C^{2,\alpha}$, we would need to impose more regularity to the kernel K in L than C^1 away from the origin.

The next theorem says that in L^2 all linear operators have a comparable norm.

THEOREM 4.2. *Let L_0 and L_1 be two linear integro-differential operators in the class $\mathcal{L}_0$. Suppose that $L_0 u \in L^2(\mathbb{R}^n)$ then $L_1 u \in L^2(\mathbb{R}^n)$.*

Proof. Since we are dealing with L^2 norms and translation invariant linear operators, we will use the Fourier transform to prove this theorem.

Given a function u and $y \in \mathbb{R}^n$, we have

$$\widehat{\delta u(x,y)} = (u(.+y)+u(.-y)-2u)\widehat{\ } = (e^{iy\cdot\xi}+e^{-iy\cdot\xi}-2)\hat{u}(\xi) = 2(\cos(y\cdot\xi)-1)\hat{u}(\xi).$$

We use this identity to compute the symbol $s(\xi)$ of an operator $-L$ as a pseudo differential operator

$$\begin{aligned} -\widehat{Lu}(\xi) &= -\Big(\int_{\mathbb{R}^n} \delta u(x,y) K(y)\,\mathrm{d}y\Big)\widehat{\ }(\xi) \\ &= \Big(\int_{\mathbb{R}^n} 2(1-\cos(y\cdot\xi))K(y)\,\mathrm{d}y\Big)\,\hat{u}(\xi) =: s(\xi)\hat{u}(\xi). \end{aligned}$$

Note that for every ξ function $(1-\cos(y\cdot\xi))$ is C^2 and bounded, so the integral in the right-hand side is well defined. Let us estimate it from above and below.

For any $R > 0$,

$$\begin{aligned} s(\xi) &= \int_{\mathbb{R}^n} 2(1-\cos(y\cdot\xi))K(y)\,\mathrm{d}y \\ &\le \int_{B_R} 2|y\cdot\xi|^2(2-\sigma)\frac{\Lambda}{|y|^{n+\sigma}}\,\mathrm{d}y + \int_{\mathbb{R}^n\setminus B_R} 2(2-\sigma)\frac{\Lambda}{|y|^{n+\sigma}}\,\mathrm{d}y \\ &\le C|\xi|^2 R^{2-\sigma} + C\frac{(2-\sigma)}{\sigma}R^{-\sigma} \end{aligned}$$

so we obtain $s(\xi) \le C|\xi|^\sigma$ by choosing $R = |\xi|^{-1}$.

On the other hand, note that $(1-\cos(y\cdot\xi))$ is nonnegative and so is $K(y)$. Thus the integrand is nonnegative and we have

$$\begin{aligned} s(\xi) = \int_{\mathbb{R}^n} 2(1-\cos(y\cdot\xi))K(y)\,\mathrm{d}y &\ge \int_{B_{|\xi|^{-1}/2}} \frac{1}{4}|y\cdot\xi|^2(2-\sigma)\frac{\lambda}{|y|^{n+\sigma}}\,\mathrm{d}y \\ &\ge c|\xi|^{-\sigma}. \end{aligned}$$

So the symbol $s(\xi)$ is comparable to $|\xi|^{-\sigma}$ for any operator L in $\mathcal{L}_0$. Thus by classical Fourier analysis, we have that $L_1 L_0^{-1}$ has a bounded symbol and maps L^2 functions into L^2. ☐

The following theorem is a direct combination of Theorems 4.1 and 4.2.

THEOREM 4.3. *Let L be an integro-differential operator in the class $\mathcal{L}_1$ with $\sigma \geq \sigma_0 > 1$. Suppose u is a function in the weighted space $L^1(\mathbb{R}^n, \frac{1}{1+|y|^{n+\sigma}})$ that solves the equation $Lu = f$ in B_1 for some $f \in L^2$. Let L_1 be an operator in $\mathcal{L}_0$; then $L_1 u \in L^2(B_{1/2})$ and*

$$\|L_1 u\|_{L^2(B_{1/2})} \leq C\left(\|u\|_{L^1(\mathbb{R}^n, \frac{1}{1+|y|^{n+\sigma}})} + \|f\|_{L^2(B_1)}\right)$$

for some constant C depending on n, λ, Λ and σ_0.

Proof. Consider the function v that solves

$$Lv = f\chi_{B_1} \text{ in } \mathbb{R}^n.$$

From Theorem 4.2 we get that $L_1 v \in L^2(\mathbb{R}^n)$. Theorem 4.2 can also be applied to $L_2 = (-\triangle)^{\sigma/2}$; thus v is in the homogeneous fractional Sobolev space $\dot{H}^{\sigma/2}$. By Sobolev embedding, $v \in L^p(\mathbb{R}^n)$ where $p = 2n/(n-2\sigma)$. In particular, $v \in L^1(\mathbb{R}^n, \frac{1}{1+|y|^{n+\sigma}})$.

Now we apply Theorem 4.1 to $u - v \in L^1(\mathbb{R}^n, \frac{1}{1+|y|^{n+\sigma}})$ to conclude the proof. ☐

Remark 4.4. The ellipticity constants λ and Λ of L_0 and L_1 do not need to coincide. Indeed if each L_i is elliptic with constants λ_i and Λ_i, then they would both be elliptic with respect to the constants $\min(\lambda_0, \lambda_1)$ and $\max(\Lambda_0, \Lambda_1)$.

5. Subsolutions in L^1 are bounded above

The theorem below is a weak version of the mean value theorem. Its proof uses the same ideas as the proof of the Harnack inequality in [3]. We include it here for completeness.

THEOREM 5.1. *Let u be a function such that u is continuous in $\overline{B}_1$ and assume that*

$$\int_{\mathbb{R}^n} \frac{|u(y)|}{1+|y|^{n+\sigma}}\, \mathrm{d}y \leq C_0,$$

$$\mathrm{M}_0^+ u \geq -C_0 \quad \text{in } B_1;$$

then

$$u(x) \leq CC_0 \text{ in } B_{1/2}$$

for every $x \in B_{1/2}$, where C is a universal constant.

Proof. Dividing u by C_0, we can assume without loss of generality that $C_0 = 1$.

Let us consider the minimum value of t such that

$$u(x) \leq h_t(x) := t(1 - |x|)^{-n} \text{ for every } x \in B_1.$$

There must be an $x_0 \in B_1$ such that $u(x_0) = h_t(x_0)$; otherwise we could make t smaller. Let $d = (1 - |x_0|)$ be the distance from x_0 to ∂B_1.

For $r = d/2$, we want to estimate the portion of the ball $B_r(x_0)$ covered by $\{u < u(x_0)/2\}$ and by $\{u > u(x_0)/2\}$. We will show that t cannot be too large. In this way we obtain the result of the theorem, since the upper bound $t < C$ implies that $u(x) < C(1 - |x|)^{-n}$.

Let us first consider $A := \{u > u(x_0)/2\}$. By assumption, we have $u \in L^1(B_1)$; thus

$$\begin{aligned} |A \cap B_1| &\leq C \left| \frac{2}{u(x_0)} \right| \\ &\leq C t^{-1} d^n. \end{aligned}$$

Whereas $|B_r| = Cd^n$, so if t is large, A can cover only a small portion of $B_r(x_0)$ at most

$$|\{u > u(x_0)/2\} \cap B_r(x_0)| \leq C t^{-1} |B_r| . \tag{5.1}$$

In order to get a contradiction, we will show that $|\{u < u(x_0)/2\} \cap B_r(x_0)| \leq (1 - \alpha)|B_r|$ for a positive constant α independent of t.

We estimate $|\{u < u(x_0)/2\} \cap B_{\theta r}(x_0)|$ for $\theta > 0$ small. For every $x \in B_{\theta r}(x_0)$, we have $u(x) \leq h_t(x) \leq t(d - \theta d/2)^{-n} \leq u(x_0)(1 - \theta/2)^{-n}$, with $(1 - \theta/2)^{-n}$ close to one.

Let us consider

$$v(x) = (1 - \theta/2)^{-n} u(x_0) - u(x)$$

so that $v \geq 0$ in $B_{\theta r}$ and also $\mathrm{M}_0^- v \leq 1$ since $\mathrm{M}_0^+ u \geq -1$. We would want to apply Theorem 10.4 in [3] (the L^ε estimate) to v. The only problem is that v is not positive in the whole domain but only on $B_{\theta r}$. In order to apply such a theorem, we have to consider $w = v^+$ instead and estimate the change in the right-hand side due to the truncation error.

We want to find an upper bound for $\mathrm{M}_0^- w = \mathrm{M}_0^- v^+$ instead of $\mathrm{M}_0^- v$. We know that

$$\mathrm{M}_0^- v(x) = (2 - \sigma) \int_{\mathbb{R}^n} \frac{\lambda \delta v(x, y)^+ - \Lambda \delta v(x, y)^-}{|y|^{n+\sigma}} \, \mathrm{d}y \leq 1.$$

Therefore, if $x \in B_{\theta r/2}(x_0)$,

$$
\begin{aligned}
(5.2) \qquad \mathrm{M}_0^- w &= (2-\sigma) \int_{\mathbb{R}^n} \frac{\lambda \delta w(x,y)^+ - \Lambda \delta w(x,y)^-}{|y|^{n+\sigma}} \,\mathrm{d}y \\
&\leq 1 + 2(2-\sigma) \int_{\mathbb{R}^n \cap \{v(x+y)<0\}} -\Lambda \frac{v(x+y)}{|y|^{n+\sigma}} \,\mathrm{d}y \\
&\leq 1 + 2(2-\sigma) \int_{\mathbb{R}^n \setminus B_{\theta r/2}} \Lambda \frac{(u(x+y) - (1-\theta/2)^{-n} u(x_0))^+}{|y|^{n+\sigma}} \,\mathrm{d}y \\
&\leq 1 + C(2-\sigma)(\theta r)^{-n-\sigma} \int_{\mathbb{R}^n} \Lambda \frac{|u(y)|}{1+|y|^{n+\sigma}} \,\mathrm{d}y \leq C(\theta r)^{-n-\sigma}.
\end{aligned}
$$

Now we can apply Theorem 10.4 from [3] to w in $B_{\theta r/2}(x_0)$. Recall $w(x_0) = ((1-\theta/2)^{-n} - 1)u(x_0)$, and we have

$$
\begin{aligned}
&\left| \left\{ u < \frac{u(x_0)}{2} \right\} \cap B_{\frac{\theta r}{4}(x_0)} \right| \\
&\quad = |\{w > u(x_0)((1-\theta/2)^{-n} - 1/2)\} \cap B_{\theta r/4}(x_0)| \\
&\quad \leq C(\theta r)^n \Big(((1-\theta/2)^{-n} - 1) u(x_0) + C(\theta r)^{-n-\sigma} (r\theta)^\sigma \Big)^\varepsilon \Big(u(x_0)((1-\theta/2)^{-n} - \frac{1}{2}) \Big)^{-\varepsilon} \\
&\quad \leq C(\theta r)^n \big(((1-\theta/2)^{-n} - 1)^\varepsilon + \theta^{-n\varepsilon} t^{-\varepsilon} \big).
\end{aligned}
$$

Now let us choose $\theta > 0$ so that the first term is small:

$$
C(\theta r)^n ((1-\theta/2)^{-n} - 1)^\varepsilon \leq \frac{1}{4} \left| B_{\theta r/2} \right|.
$$

Notice that the choice of θ is independent of t. For this fixed value of θ, we observe that if t is large enough, we will also have

$$
C(\theta r)^n \theta^{-n\varepsilon} t^{-\varepsilon} \leq \frac{1}{4} \left| B_{\theta r/2} \right|
$$

and therefore

$$
|\{u < u(x_0)/2\} \cap B_{\theta/r4}(x_0)| \leq \frac{1}{2} \left| B_{\theta r/4}(x_0) \right|
$$

which implies that for t large,

$$
|\{u > u(x_0)/2\} \cap B_{\theta r/4}(x_0)| \geq c\,|B_r|\,.
$$

But this contradicts (5.1). Therefore t cannot be large. Rescaling back, we obtain

$$
u(x) \leq CC_0
$$

for any x in $B_{1/2}$. $\square$

6. **Each $L_a u$ is bounded**

The idea of this section is to show that averages of second order incremental quotients are subsolutions (of the maximal operator M^+). In particular, each $v_a := L_a u$ is a subsolution to

$$\begin{aligned} v_a &\geq 0 \text{ in } B_1, \\ \mathrm{M}^+ v_a &\geq -C||u||_{L^\infty} \text{ in } B_1. \end{aligned}$$

Then we would estimate the integral of v_a in $B_{1/2}$ and use Theorem 5.1 to prove that $L_a u$ is bounded.

Lemma 6.1. *Assume $u \in C^2$ and $Iu = 0$ in B_1. Let K be a symmetric kernel satisfying $K(y) \leq (2-\sigma)\Lambda|y|^{-n-\sigma}$ (but not necessarily the bound below). Then for every bump function b such that*

$$\begin{aligned} 0 \leq b(x) &\leq 1 \textit{ in } \mathbb{R}^n, \\ b(x) &= b(-x) \textit{ in } \mathbb{R}^n, \\ b(x) &= 0 \textit{ in } \mathbb{R}^n \setminus B_{1/2}, \end{aligned}$$

we have

$$\mathrm{M}_2^+ \left(\int_{\mathbb{R}^n} \delta u(x,y) K(y) b(y) \, \mathrm{d}y \right) \geq 0 \qquad \textit{in } B_{1/2}.$$

Proof. Let ϕ_k be the L^1 function $\phi_k(y) = \chi_{\mathbb{R}^n \setminus B_{1/k}}(y) K(y) b(y)$. Since $u \in C^2$ and K is symmetric, we can approximate the value of the integral uniformly by

$$\begin{aligned} \int_{\mathbb{R}^n} \delta u(x,y) K(y) b(y) \, \mathrm{d}y &= \lim_{k\to\infty} \int_{\mathbb{R}^n} \delta u(x,y) \phi_k(y) \, \mathrm{d}y \\ &= 2 \left(\lim_{k\to\infty} u * \phi_k - \|\phi_k\|_{L^1} u \right). \end{aligned}$$

Applying Proposition 3.1, we have

$$I\left(u * \frac{\phi_k}{\|\phi_k\|_{L^1}} \right) \geq 0.$$

On the other hand, we know that $Iu = 0$; then by Lemma 5.8 in [3] and the fact that M^+ is homogeneous,

$$\mathrm{M}_2^+ (u * \phi_k - \|\phi_k\|_{L^1} u) = \|\phi_k\|_{L^1} \mathrm{M}_2^+ \left(u * \frac{\phi_k}{\|\phi_k\|_{L^1}} - u \right) \geq 0.$$

Since we have $\mathrm{M}_2^+(u * \phi_k - \|\phi_k\|_{L^1} u) \geq 0$ for every $k > 0$, the result follows by Lemma 4.3 in [5] by taking limit as $k \to \infty$. □

Lemma 6.2. *Assume $u \in C^2$ and $Iu = 0$ in B_1. Then there is a universal constant C such that for every operator $L \in \mathcal{L}_2$,*

$$\mathrm{M}_2^+ (Lu) \geq -C||u||_{L^\infty} \qquad \textit{in } B_{1/2}.$$

Proof. As in the proof of Lemma 6.1, we let ϕ_k be the L^1 function $\phi_k(y) = \chi_{\mathbb{R}^n \setminus B_{1/k}}(y) K(y)$ and we approximate the value of the integral uniformly by

$$Lu(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} \delta u(x,y) \phi_k(y) \, \mathrm{d}y = 2 \lim_{k\to\infty} (u * \phi_k - u \left\|\phi_k\right\|_{L_1}).$$

Let b be a smooth bump function such that

$$\begin{aligned} 0 \leq b(x) &\leq 1 \text{ in } \mathbb{R}^n, \\ b(x) &= b(-x) \text{ in } \mathbb{R}^n, \\ b(x) &= 0 \text{ in } \mathbb{R}^n \setminus B_{1/2}, \\ b(x) &= 1 \text{ in } B_{1/4}. \end{aligned}$$

As in the proof of Lemma 6.1,

$$I\left(u * \frac{\phi_k(x) b(x)}{\left\|\phi_k(x) b(x)\right\|_{L^1}} \right) \geq 0.$$

Therefore $\mathrm{M}_2^+ \left(u * (\phi_k b) - \left\|\phi_k b\right\|_{L^1} u\right) \geq 0$.

On the other hand, we estimate $I(u * (\phi_k(1-b)))$ in $B_{1/2}$:

$$\begin{aligned} I(u * (\phi_k(1-b))) &= \inf_a \int_{\mathbb{R}^n} L_a(u * (\phi_k(1-b))) \, \mathrm{d}y \\ &= \inf_a \int_{\mathbb{R}^n} u * L_a(\phi_k(1-b)) \, \mathrm{d}y. \end{aligned}$$

Since by (1.2), $\phi_k(1-b) \in L^1$ and by (1.3), $D^2 \phi_k(1-b) \in L^1$ uniformly in k, then $L(\phi_k(1-b))$ is bounded in L^1 uniformly for all $L \in \mathcal{L}_0$. Therefore,

$$|I(u * (\phi_k(1-b)))| \leq C \left\|u\right\|_{L^\infty}.$$

Using Lemma 5.8 in [3] and the homogeneity of M^+,

$$\begin{aligned} &\mathrm{M}_2^+ \left(u * \phi_k - \left\|\phi_k\right\|_{L^1} u\right) \\ &\qquad \geq I\left(u * \phi_k\right) - \left\|\phi_k\right\|_{L^1} Iu = I\left(u * \phi_k\right) \\ &\qquad \geq I\left(u * (\phi_k b)\right) + I\left(u * (\phi_k(1-b))\right) \qquad \text{by the concavity of } I \\ &\qquad \geq -C \left\|u\right\|_{L^\infty}. \end{aligned}$$

We finish the proof of the lemma by taking $k \to \infty$, using Lemma 4.3 in [5]. □

LEMMA 6.3. *Let u be a solution of (1.1) in B_1. Assume $L_a \in \mathcal{L}_2$ for every a. Then for every a, $L_a u \leq C \left\|u\right\|_{L^\infty}$ in $B_{1/8}$ for some universal constant C.*

Proof. By Lemma 2.1, we can assume the function u is C^2. Indeed, for every $\varepsilon > 0$, we can approximate u with a C^2 function u^ε that satisfies the same kind of equation. If we can prove the estimate for u^ε with a universal constant C that does not depend on ε, then we would prove it for u by passing to the limit as $\varepsilon \to 0$. So we assume that $u \in C^2$, and thus all the integrals are well defined.

From Lemma 6.2, we know that for each L_a,

$$\mathrm{M}_2^+(L_a u) \geq -C \|u\|_{L^\infty} \text{ in } B_{1/2}.$$

We would want to apply Theorem 5.1 to $L_a u$. For that we still need an estimate at least in $L^1((1+|y|)^{-n-\sigma})$. We can easily obtain an estimate in $L^1(B_{1/2})$ using the fact that $L_a u \geq 0$ in B_1 because of equation (1.1).

Let b be a smooth cutoff function such that $0 \leq b \leq 1$ in $\mathbb{R}^n$, $b = 1$ in $B_{1/2}$ and $b = 0$ outside B_1. We multiply $L_a u$ by b and *integrate by parts*:

$$\int_{\mathbb{R}^n} L_a u(x)\ b(x) \,\mathrm{d}x = \int_{\mathbb{R}^n} L_a b(x)\ u(x) \,\mathrm{d}x \leq C \|u\|_{L^\infty}$$

for some universal constant C. Since $L_a u \geq 0$ in B_1, then it is in $L^1(B_{1/2})$.

We would still need some control on the values of $L_a u$ away from $B_{1/2}$ in order to apply Theorem 5.1. We do not want to assume any regularity for u outside B_1, so our only choice is to cut off again.

Let $c(x) := b(2x)$ and $w(x) = c(x)\ L_a u(x)$. We will estimate $\mathrm{M}_2^+ w(x)$ for $x \in B_{1/4}$. For that, let us consider any operator $L \in \mathcal{L}_2$ and estimate

$$\begin{aligned}
Lw(x) &= \int_{\mathbb{R}^n} \delta w(x,y) K(y) \,\mathrm{d}y \\
&= \int_{\mathbb{R}^n} \delta L_a u(x,y) K(y) \,\mathrm{d}y - \int_{\mathbb{R}^n} \delta(L_a u(1-c))(x,y) K(y) \,\mathrm{d}y \\
&\geq L(L_a u) - 2\int_{\mathbb{R}^n} L_a u(x+y)(1-c(x+y)) K(y) \,\mathrm{d}y \\
&\geq L(L_a u) - 2\int_{\mathbb{R}^n} u(x+y) L_a\left((1-c(x+.))K(.)\right)(y) \,\mathrm{d}y \\
&\geq L(L_a u) - C \|u\|_{L^\infty}.
\end{aligned}$$

For the last inequality, we used the fact that $L_a\left((1-c(x+.))K\right)$ is in L^1 uniformly for $x \in B_{1/4}$. This follows from the estimates (1.2) and (1.3) of the kernel K.

Taking supremum in a, we obtain

$$\mathrm{M}_2^+ w(x) \geq \mathrm{M}_2^+(L_a u) - C \|u\|_{L^\infty} \geq -C \|u\|_{L^\infty}.$$

Now we can apply Theorem 5.1 to w to obtain that $w \leq C \|u\|_{L^\infty}$ in $B_{1/8}$. but $w = L_a u$ in $B_{1/4}$, so we finish the proof. □

7. Extremal operators are bounded

In this section we prove Theorem 7.4. We will achieve this result by showing that $Lu(x)$ is bounded uniformly for all $L \in \mathcal{L}_0$. The ideas are very similar to Section 6 but now we apply them to operators that are not *a priori* bounded below, so we use the L^2 estimates from Section 4 instead.

LEMMA 7.1. *Assume $u \in C^2$. Then for every symmetric kernel K satisfying $K(y) \leq (2-\sigma)\Lambda |y|^{-n-\sigma}$ (but not necessarily the bound below) and every smooth bump function b such that*

$$\begin{aligned} 0 \leq b(x) \leq 1 & \text{ in } \mathbb{R}^n, \\ b(x) = b(-x) & \text{ in } \mathbb{R}^n, \\ b(x) = 0 & \text{ in } \mathbb{R}^n \setminus B_{1/2}, \\ b(x) = 1 & \text{ in } B_{1/4}, \end{aligned}$$

we have

$$\mathrm{M}_2^+ \left(b(x) \int_{B_{1/2}} \delta u(x,y) K(y) \, \mathrm{d}y \right) \geq -C \|u\|_{L^\infty} \qquad \text{in } B_{1/2}.$$

Proof. Let us call

$$L^t u(x) = \int_{B_{1/2}} \delta u(x,y) K(y) \, \mathrm{d}y.$$

By Lemma 6.1, we have $\mathrm{M}_2^+(L^t u) \geq 0$. Let L be any operator in $\mathcal{L}_2$; thus we estimate

$$\begin{aligned} L(bL^t u)(x) &= \int_{\mathbb{R}^n} \delta(L^t u)(x,y) K(y) \, \mathrm{d}y - \int_{\mathbb{R}^n} \delta((1-b)L^t u)(x,y) K(y) \, \mathrm{d}y \\ &\geq -2 \int_{\mathbb{R}^n} (1 - b(x+y)) L^t u(x+y) K(y) \, \mathrm{d}y \\ &\geq -2 \int_{\mathbb{R}^n} u(x+y) L^t((1 - b(x+\cdot))K)(y) \, \mathrm{d}y \\ &\geq -C \|u\|_{L^\infty}, \end{aligned}$$

where we used that $L^t((1-b(x+\cdot))K)$ is bounded in L^1 uniformly in x. This is due to the fact that $D^2((1-b(x+\cdot))K) \in L^1(\mathbb{R}^n)$ because of (1.3). □

LEMMA 7.2. *Let u be a solution of* (1.1) *in B_1 with all operators L_a in $\mathcal{L}_2$. There is a constant C such that for every operator L with a symmetric kernel K satisfying $K(y) \leq (2-\sigma)\Lambda |y|^{-n-\sigma}$ (but not necessarily the bound below), we have*

$$|Lu(x)| \leq C \|u\|_{L^\infty} \qquad \text{in } B_{1/2}.$$

Proof. As in the proof of Lemma 6.3, we can and will assume that $u \in C^2$. We will write the proof assuming that $\|u\|_{L^\infty(\mathbb{R}^n)} = 1$. Moreover, we will prove the estimate in $B_{1/64}$ instead of $B_{1/2}$. The general estimate follows directly by scaling and a standard covering argument.

Let L_a be one of the operators used in the infimum in (1.1). We know from Lemma 6.3 that L_a is bounded in $B_{1/2}$ by a constant C. In particular, $\|L_a u\|_{L^2(B_{1/2})} \leq C$. Note that $||u||_{L^1(\mathbb{R}^n, 1/(1+|y|^{n+\sigma}))} \leq C \|u\|_{L^\infty}$. From

Theorem 4.3, we have an L^2 estimate for every $L \in \mathcal{L}_0$:

$$\|Lu\|_{L^2(B_{1/4})} \leq C.$$

We split the integral of Lu into two domains

$$Lu(x) = \int_{B_{1/2}} \delta u(x,y) K(y) \, \mathrm{d}y + \int_{\mathbb{R}^n \setminus B_{1/2}} \delta u(x,y) K(y) \, \mathrm{d}y.$$

The second integral is clearly bounded since $K(y)$ is a function in $L^1(\mathbb{R}^n \setminus B_{1/2})$. Thus we still have an estimate in L^2 for the first term:

$$\left\| \int_{B_{1/2}} \delta u(x,y) K(y) \, \mathrm{d}y \right\|_{L^2(B_{1/4})} \leq C.$$

On the other hand, from Lemma 6.1,

$$\mathrm{M}_2^+ \left(\int_{B_{1/2}} \delta u(x,y) K(y) \, \mathrm{d}y \right) \geq 0.$$

For a bump function $c(x)$ such that

$$\begin{aligned} \operatorname{supp} c &= B_{1/4}, \\ c &\equiv 1 \text{ in } B_{1/8}, \end{aligned}$$

we define

$$w(x) := c(x) \int_{B_{1/2}} \delta u(x,y) K(y) \, \mathrm{d}y.$$

We know that $w \in L^2(\mathbb{R}^n)$ and $w = 0$ outside $B_{1/4}$. In particular w is bounded in the weighted L^1 space: $L^1((1+|y|^{n+\sigma})^{-1})$ needed for Theorem 5.1.

From Lemma 7.1, we have $\mathrm{M}_2^+ w \geq -C$ in $B_{1/16}$. Thus we can apply Theorem 5.1 to w to obtain that $w \leq C$ in $B_{1/32}$, which naturally implies $Lu \leq C$ in $B_{1/32}$.

We got the desired bound from above only. In order to get the corresponding bound from below we must use equation (1.1) again.

Recall that $L_a u$ is bounded by Lemma 6.3, and the formulas of L_a and L are given by

$$\begin{aligned} L_a u(x) &= \int \delta u(x,y) K_a(y) \, \mathrm{d}y, \\ Lu(x) &= \int \delta u(x,y) K(y) \, \mathrm{d}y, \end{aligned}$$

where K_a is bounded below by $(2-\sigma)\lambda/|y|^{n+\sigma}$ and both K and K_a are bounded above by $(2-\sigma)\Lambda/|y|^{n+\sigma}$.

Consider the kernel

$$K_d = \frac{2}{\lambda} K_a - \frac{1}{\Lambda} K$$

and the corresponding linear operator L_d. The kernel K_d satisfies the ellipticity conditions $(2-\sigma)/|y|^{n+\sigma} \leq K_d \leq (2-\sigma)(2\Lambda/\lambda - \lambda/\Lambda)/|y|^{n+\sigma}$, so L_d is in the class $\mathcal{L}_0$ with ellipticity constants 1 and $(2\Lambda/\lambda - \lambda/\Lambda)$. The same proof as above tells us that $L_d u \leq C$ in $B_{1/32}$. But then since L_a is bounded, we obtain a bound below for L in $B_{1/32}$:

$$Lu = 2\frac{\Lambda}{\lambda} L_a - \Lambda L_d \geq -C.$$

Thus, we have both bounds and we obtain $|Lu| \leq C$ in $B_{1/32}$. □

Corollary 7.3. *$\mathrm{M}_0^+ u$ and $\mathrm{M}_0^- u$ are bounded in $B_{1/2}$.*

Proof. Since $\mathrm{M}_0^+ u = \sup_{L \in \mathcal{L}_0} Lu$ and for every L in $\mathcal{L}_0$ we have $|Lu| \leq C \|u\|_{L^\infty}$ with C independent of the choice of L in $\mathcal{L}_0$, then also $|\mathrm{M}_0^+ u| \leq C \|u\|_{L^\infty}$ with the same constant C. □

Now we are going to prove that all integrals in (1.1) are absolutely convergent. This already implies that the solution is classical in some way.

Theorem 7.4. *Assume every L_a in* (1.1) *belongs to the class $\mathcal{L}_2$. If u is a bounded function in $\mathbb{R}^n$ such that $Iu = 0$ in B_1 in the viscosity sense, then we have the following estimate*:

$$\int_{\mathbb{R}^n} |\delta u(x,y)| \frac{(2-\sigma)}{|y|^{n+\sigma}} \, \mathrm{d}y \leq C \|u\|_{L^\infty(\mathbb{R}^n)} \quad \textit{in } B_{1/2}.$$

Proof. Applying Lemma 7.2 to $L = -(-\triangle)^{\sigma/2}$, we get

$$|-(-\triangle)^{\sigma/2} u(x)| = \left| c_\sigma \int_{\mathbb{R}^n} \delta u(x,y) \frac{(2-\sigma)}{|y|^{n+\sigma}} \, \mathrm{d}y \right| \leq C \|u\|_{L^\infty(\mathbb{R}^n)} \quad \text{in } B_{1/2}.$$

On the other hand, applying Corollary 7.3 with any pair $\lambda < \Lambda$, we get

$$\begin{aligned} |\mathrm{M}_0^+ u(x)| &= \left| \int_{\mathbb{R}^n} (\Lambda \delta u(x,y)^+ - \lambda \delta u(x,y)^-) \frac{(2-\sigma)}{|y|^{n+\sigma}} \, \mathrm{d}y \right| \\ &\leq C \|u\|_{L^\infty(\mathbb{R}^n)} \quad \text{in } B_{1/2}. \end{aligned}$$

Subtracting, we obtain

$$\mathrm{M}_0^+ u(x) + \lambda(-\triangle)^{\sigma/2} u(x) \leq C \|u\|_{L^\infty(\mathbb{R}^n)},$$

$$(\Lambda - \lambda) \int_{\mathbb{R}^n} \delta u(x,y)^+ \frac{(2-\sigma)}{|y|^{n+\sigma}} \, \mathrm{d}y \leq C \|u\|_{L^\infty(\mathbb{R}^n)}.$$

On the other hand, by subtracting $\Lambda(-\triangle)^{\sigma/2} u(x) - \mathrm{M}_0^+ u(x)$ we obtain the bound

$$(\Lambda - \lambda) \int_{\mathbb{R}^n} \delta u(x,y)^- \frac{(2-\sigma)}{|y|^{n+\sigma}} \, \mathrm{d}y = \Lambda(-\triangle)^{\sigma/2} u(x) - \mathrm{M}_0^+ u(x) \leq C \|u\|_{L^\infty(\mathbb{R}^n)}.$$

Combining the two estimates above, we finish the proof. □

8. **Outline of the strategy: the second order case**

Theorem 7.4 provides an estimate slightly stronger than $u \in C^\sigma$. In the case $\sigma \to 2$ it becomes an estimate of the $C^{1,1}$ norm of u. Comparing with the proof of Evans-Krylov theorem (as in [6], [7] or [2]), the underlying strategy of the proof up to this point is essentially the same but adapted to the integro-differential setting using the ideas in our previous papers [3] and [5].

The next step in the proof of our main result is to pass from this C^σ estimate to a $C^{\sigma+\alpha}$ estimate. In the second order case it corresponds to the *a priori* estimate $\|u\|_{C^{2,\alpha}(B_{1/2})} \leq C \|u\|_{C^{1,1}(B_1)}$. The presently known proofs of this *a priori* estimate seem difficult to adapt to the nonlocal setting. Thus we present a different strategy for the proof. The key tools that the proof is based on are similar, but in our approach they are organized differently and arguably more directly. We plan to publish a short note [4] focusing only on this new proof for concave second order elliptic equations.

In order to better understand our proof in the next section, we first sketch its adaptation to the second order case.

We consider a C^2 solution of a fully nonlinear equation $F(D^2u) = 0$ with F concave and uniformly elliptic. We transform this into an integral equation by pointing out that a linear equation can be written as an integral on the unit sphere S_1:

$$a_{ij}\partial_{ij}u(x) = \int_{S_1} \partial_{\sigma\sigma}u(x)\ w(\sigma)\ \mathrm{d}\sigma$$

with the weight $w(\sigma) = 1/(\det\{a_{ij}\}a^{ij}\sigma_i\sigma_j)$, where $\{a^{ij}\} = \{a_{ij}\}^{-1}$. If the coefficients a_{ij} are uniformly elliptic, then $w(\sigma)$ will be bounded away from zero.

We recall that F being concave implies that pure second derivatives are all subsolutions of the linearized operator. Therefore, for any fixed $A \subset S_1$, the following function is also a subsolution:

$$v_A = \int_A \partial_{\sigma\sigma}u(x)\ \mathrm{d}\sigma.$$

We also recall that for a (nonnecessarily concave) fully nonlinear equation

$$F(D^2u) = \sup_b \inf_a L_{ab}u = 0,$$

a solution u satisfies (just because it is an inf sup) that for any two points x and y in the domain, there exists an operator L_{ab} for which

$$L_{ab}u(x) - L_{ab}u(y) \geq 0.$$

In our approach, this means that there is a weight $w(\sigma)$, bounded below and above depending on the ellipticity constants, such that

$$\int_{S_1} (\partial_{\sigma\sigma}u(x) - \partial_{\sigma\sigma}u(y))\, w(\sigma)\ \mathrm{d}\sigma \geq 0.$$

In particular, since there is another weight which gives the same inequality exchanging x and y, we must have that the following quantities are comparable:

$$\begin{aligned} \int_{S_1} (\partial_{\sigma\sigma} u(x) - \partial_{\sigma\sigma} u(y))^+ \, \mathrm{d}\sigma &\approx \int_{S_1} (\partial_{\sigma\sigma} u(x) - \partial_{\sigma\sigma} u(y))^- \, \mathrm{d}\sigma \\ &\approx \int_{S_1} |\partial_{\sigma\sigma} u(x) - \partial_{\sigma\sigma} u(y)| \, \mathrm{d}\sigma. \end{aligned} \tag{8.1}$$

At this point we define

$$h(x,\sigma) = \partial_{\sigma\sigma} u(x) - \partial_{\sigma\sigma} u(0),$$
$$w_A(x) = \int_A h(x,\sigma) \, \mathrm{d}\sigma$$

for any set $A \subset S_1$. We will use only the properties above to show that

$$\int_{S_1} |h(x,\sigma)| \, \mathrm{d}\sigma \leq |x|^\alpha.$$

The $C^{2,\alpha}$ estimate for u follows easily from this estimate.

By (8.1), we only need to prove that $w_A(x) \leq C|x|^\alpha$ for every set A, since

$$\int_{S_1} |h(x,\sigma)| \, \mathrm{d}\sigma \approx \int_{S_1} h(x,\sigma)^+ \, \mathrm{d}\sigma = \sup_A w_A(x).$$

In fact, by renormalization we only need to prove the following lemma.

LEMMA 8.1. *Assume that for x in B_1, for any set $A \subset S_1$, $w_A(x) \leq 1$. Then there is a universal constant $\theta > 0$ such that*

$$w_A(x) \leq 1 - \theta$$

for any $A \subset S_1$ and $x \in B_{1/2}$.

Sketch of the proof. Suppose that there exists an $x \in B_{1/2}$, where $w_A > 1 - \theta$ for some set $A \subset S_1$. We will arrive to a contradiction if θ is too small.

Since $w_A \leq 1$ in B_1 and w_A is a subsolution of the linearized equation, we can apply Theorem 4.8(1) in [2] (the L^ε estimate) to $1 - w_A$ (this will correspond to Theorem 10.4 in [3] in the nonlocal case). It follows that we can make

$$\Omega = \{w_A(x) \geq 1 - t\theta\}$$

cover almost all $B_{1/4}$ if we choose t large (but independently of θ).

We will now obtain a contradiction by looking at w_{A^c} in $B_{1/4}$.

For every x in Ω, the choice of the set A is almost maximal in the sense that

$$1 - t\theta \leq w_A(x) \leq \int_{S_1} h(x,\sigma)^+ \, \mathrm{d}\sigma \leq 1.$$

On the other hand, since

$$\int_{S_1} h(x,\sigma) \, \mathrm{d}\sigma = \int_{S_1} h(x,\sigma)^+ \, \mathrm{d}\sigma - \int_{S_1} h(x,\sigma)^- \, \mathrm{d}\sigma = w_A(x) + w_{A^c}(x),$$

then also

$$0 \le w_{A^c} + \int_{S_1} h(x,\sigma)^- \, \mathrm{d}\sigma \le t\theta.$$

From (8.1), we know that the integrals of h^+ and h^- are comparable. Thus in Ω, we have

$$w_{A^c} \le t\theta - C$$

for a constant C depending on λ and Λ. If we choose θ small, that means that w_{A^c} will be strictly negative in most of $B_{1/4}$. But then applying Theorem 4.8(2) in [2] (which corresponds to Theorem 5.1 in the nonlocal case), we obtain that $w_{A^c}(0) \le -c$ for some universal constant c. This is a contradiction since clearly $w_{A^c}(0) = 0$. $\square$

In the integro-differential case, the proof will be slightly lengthier in part because we have to keep track of the truncation error we make every time we localize an integral. We cover the proof in detail in the next section.

9. **Further regularity**

This section is devoted to fill the gap between Theorem 7.4 and Theorem 1.1.

From Theorem 7.4, we know that

$$\int_{B_{1/2}} |\delta u(x,y)| \frac{2-\sigma}{|y|^{n+\sigma}} \, \mathrm{d}y \le C \, \|u\|_{L^\infty} \quad \text{in } B_{1/4}. \tag{9.1}$$

Our objective is to show that

$$\int_{B_{1/2}} |\delta u(x,y) - \delta u(0,y)| \frac{2-\sigma}{|y|^{n+\sigma}} \, \mathrm{d}y \le C|x|^\alpha \, \|u\|_{L^\infty}$$

for some constant C and $\alpha > 0$ and for every $x \in B_{1/4}$. This estimate implies the Hölder continuity of the fractional Laplacian $(-\triangle)^{\sigma/2}$ from which the $C^{\sigma+\alpha}$ regularity of u follows.

We will consider all kernels K of the form

$$K_A(y) = \frac{(2-\sigma)}{|y|^{n+\sigma}} \chi_A(y),$$

where $\chi_A(y)$ is the characteristic function of a set A, which is only require to be symmetric: $A = -A$.

Let b be a bump function as in Lemma 7.1. For each set A, we write

$$w_A(x) = b(x) \int_{B_{1/2}} (\delta u(x,y) - \delta u(0,y)) K_A(y) \, \mathrm{d}y.$$

We know that w_A is uniformly bounded from Lemma 7.2. From Lemma 7.1, we have

$$\mathrm{M}_2^+ w_A \ge -C||u||_{L^\infty} \quad \text{in } B_{1/4} \text{ uniformly in } A. \tag{9.2}$$

We define the following quantities:

$$P(x) := \sup_A w_A(x) = b(x) \int_{B_{1/2}} (\delta u(x,y) - \delta u(0,y))^+ \frac{(2-\sigma)}{|y|^{n+\sigma}} \, \mathrm{d}y,$$

$$N(x) := \sup_A -w_A(x) = b(x) \int_{B_{1/2}} (\delta u(x,y) - \delta u(0,y))^- \frac{(2-\sigma)}{|y|^{n+\sigma}} \, \mathrm{d}y.$$

Note that $P(x)$ is realized by the symmetric set $A = \{y : \delta u(x,y) > \delta u(0,y)\}$ and $N(x)$ is realized by the complement of that set.

LEMMA 9.1. *Assume* $||u||_{L^\infty} = 1$. *There is a constant* C *such that for* $x \in B_{1/4}$,

$$\frac{\lambda}{\Lambda} N(x) - C|x| \le P(x) \le \frac{\Lambda}{\lambda} N(x) + C|x|.$$

Proof. For some $x \in B_{1/4}$, let $u_x(z) := u(x+z)$. Since u solves equation (1.1) in a neighborhood of x, then both u and u_x solve (1.1) in a neighborhood of 0. Thus $\mathrm{M}_2^+(u_x - u)(0) \ge 0$ and $\mathrm{M}_2^-(u_x - u)(0) \le 0$.

For every kernel K in the family $\mathcal{L}_2$, we have

$$\begin{aligned} L(u_x - u)(0) &= \int_{\mathbb{R}^n} (\delta u(x,y) - \delta u(0,y)) K(y) \, \mathrm{d}y \\ &= \int_{B_{1/2}} (\delta u(x,y) - \delta u(0,y)) K(y) \, \mathrm{d}y \\ &\qquad + \int_{\mathbb{R}^n \setminus B_{1/2}} (\delta u(x,y) - \delta u(0,y)) K(y) \, \mathrm{d}y. \end{aligned}$$

Let us analyze the second term in the right-hand side:

$$\begin{aligned} &\int_{B_{1/2}^c} (\delta u(x,y) - \delta u(0,y)) K(y) \, \mathrm{d}y \\ &= \int_{\mathbb{R}^n} \delta u(0,y) \left(K(y-x) \chi_{B_{1/2}^c}(y-x) + K(y) \chi_{B_{1/2}^c}(y-x) \right) \, \mathrm{d}y \\ &\le \int_{\mathbb{R}^n \setminus B_{1/2+|x|}} |\delta u(0,y)| \frac{C}{|y|^{n+\sigma+1}} |x| \, \mathrm{d}y + 8 \, \|u\|_{L^\infty} \int_{B_{1/2+|x|} \setminus B_{1/2}} \frac{\Lambda(2-\sigma)}{|y|^{n+\sigma}} \, \mathrm{d}y \\ &\le C|x|. \end{aligned}$$

Therefore, for every kernel K in the family $\mathcal{L}_2$, we have

$$\int_{\mathbb{R}^n} (\delta u(x,y) - \delta u(0,y)) K(y) \, \mathrm{d}y \le \int_{B_{1/2}} (\delta u(x,y) - \delta u(0,y)) K(y) \, \mathrm{d}y + C|x|.$$

Taking the supremum, we obtain

$$0 \le \mathrm{M}_2^+(u_x - u) \le \sup_K \int_{B_{1/2}} (\delta u(x,y) - \delta u(0,y)) K(y) \, \mathrm{d}y + C|x|.$$

In particular, if we take the suppremum over all kernels K in $\mathcal{L}_0$ (a larger family), we still have

$$\sup_{\lambda\frac{(2-\sigma)}{|y|^{n+\sigma}}\leq K\leq\Lambda\frac{(2-\sigma)}{|y|^{n+\sigma}}} \int_{B_{1/2}} (\delta u(x,y)-\delta u(0,y))K(y)\,\mathrm{d}y \geq -C|x|,$$

which is the same as $\Lambda P(x)-\lambda N(x)\geq -C|x|$.

The same computation with $\mathrm{M}_2^-(u_x-u)(0)\leq 0$ provides the other inequality. □

It is important to notice the following relation:

$$\begin{aligned}\int_{B_{1/2}} |\delta u(x,y)-\delta u(0,y)|\frac{2-\sigma}{|y|^{n+\sigma}}\,\mathrm{d}y &= \sup_A w_A - \inf_A w_A\\ &= P(x)+N(x).\end{aligned}$$

The strategy for proving our regularity result will be to prove that $\sup_{x\in B_r} P(x)\leq Cr^\alpha$. It is enough to prove it for $|x|$ small enough; therefore we can consider a rescaled situation by taking $\bar{w}_A(x)=\frac{1}{C}w_A(rx)$, where C is the constant from (9.1) and r is small enough so that our estimates become

(9.3) for every set A: $|w_A|\leq 1$ in $\mathbb{R}^n$,

(9.4) for every set A: $\mathrm{M}_2^+ w_A\geq -\varepsilon_1$ in B_1,

(9.5) $$\frac{\lambda}{\Lambda}N(x)-\varepsilon_1|x|^{1-\varepsilon_1}\leq P(x)\leq\frac{\Lambda}{\lambda}N(x)+\varepsilon_1|x|^{1-\varepsilon_1}$$

for ε_1 arbitrarily small.

LEMMA 9.2. *Assume $\sigma\in(1,2)$. Let $P(x)$ be the function defined above. There is a constant C and $\alpha>0$ such that*

$$P(x)\leq C|x|^\alpha||u||_{L^\infty}.$$

Proof. We assume $||u||_{L^\infty}=1$; otherwise we divide the equation by $||u||_{L^\infty}$.

As mentioned above, after an appropriate scaling, we can assume that (9.3), (9.4) and (9.5) hold with ε_1 arbitrarily small. On the other hand, given the construction in Lemma 2.1, we can assume u is C^2 and thus w_K, P and N are continuous. We will obtain the *a priori* estimates independently of the modulus of continuity of them, so the estimate holds when passing to the limit.

We will prove that there is $r>0$ and $\theta>0$ such that

(9.6) $$\sup_{B_{r^k}}|P|\leq(1-\theta)^k=r^{\alpha k}\text{ where }\alpha=\frac{\log(1-\theta)}{\log r}.$$

This is clear for $k=0$. Let us prove it is true for all values of k by induction. So let us assume it is true up to some value k.

Since (9.6) holds up to some value k, we have that

$$|w_A(x)|\leq(1-\theta)^{-1}|x|^\alpha\text{ for }|x|>r^k.$$

Consider the following rescaled functions:

$$\begin{aligned}\widetilde{w}_A(x) &= (1-\theta)^{-k} w_A(r^k x),\\ \widetilde{P}(x) &= (1-\theta)^{-k} P(r^k x) = \sup_A \widetilde{w}_A(x),\\ \widetilde{N}(x) &= (1-\theta)^{-k} N(r^k x) = \sup_A -\widetilde{w}_A(x).\end{aligned}$$

The function $\widetilde{P}$ satisfies the relations

$$\begin{aligned}&\widetilde{P}(x) \le 1 \qquad \text{in } B_1,\\ &\widetilde{P}(x) \le (1-\theta)^{-1}|x|^\alpha \qquad \text{outside } B_1.\end{aligned}$$

Moreover, from (9.5),

$$\frac{\lambda}{\Lambda}\widetilde{N}(x) - \varepsilon_1 \le \widetilde{P}(x) \le \frac{\Lambda}{\lambda}\widetilde{N}(x) + \varepsilon_1. \tag{9.7}$$

We want to show that if θ and r are chosen small enough, we will have $\widetilde{P} \le (1-\theta)$ in B_r. The proof is by contradiction. We will arrive to a contradiction if θ and r are small enough.

Let x_0 be the point where the maximum of $\widetilde{P}$ is achieved in $\overline{B}_r$ for some $r \in (0, 1/2)$. We assume $\widetilde{P}(x_0) \ge 1-\theta$ to get a contradiction. Let A be the set such that $\widetilde{P}(x_0) = \widetilde{w}_A(x_0) \ge 1-\theta$.

Let $v_A = (1-\widetilde{w}_A)^+$. We know that $\inf_{B_r} v_A \le \theta$. Moreover,

$$\begin{aligned}\mathrm{M}_2^- v_A &\le \mathrm{M}_2^-(1-\widetilde{w}_A) + \mathrm{M}_2^+(1-\widetilde{w}_A)^-\\ &\le -\mathrm{M}_2^+\widetilde{w}_A + \mathrm{M}_2^+(1-\widetilde{w}_A)^-\\ &\le C \quad \text{in } B_{1/2}\end{aligned}$$

since $\mathrm{M}_2^+\widetilde{w}_A \ge -\varepsilon_1$ and $(1-\widetilde{w}_A)^- \le ((1-\theta)^{-1}|x|^\alpha - 1)^+$.

By Theorem 10.4 in [3], for some $p > 0$ and $r < 1/4$, we have the estimate,

$$|\{v_A > t\theta\} \cap B_{2r}(x_0)| \le Cr^n(\theta + Cr^\sigma)^p (t\theta)^{-p}.$$

Let us choose r (depending on θ to be chosen later) so that $Cr^\sigma < \theta$. Therefore, we have

$$|\{v_A > t\theta\} \cap B_r(x_0)| \le Cr^n t^{-p} = ct^{-p}|B_r|.$$

Thus, by choosing t large, we will be able to make the measure of the set $\{v_A > t\theta\} \cap B_r$ a small factor of $|B_1|$ independently of θ. Note that $v_A > t\theta$ is equivalent to $w_A < 1 - t\theta$.

Let $G = \{v_A \le t\theta\} \cap B_r$. We know that $|G| \ge (1 - ct^{-p})|B_r|$. The set G is also the set where $\widetilde{w}_A \ge 1 - t\theta$. On the other hand, since $G \subset B_1$, $\widetilde{P} \le 1$ in G, then $\widetilde{P} - \widetilde{w}_A \le t\theta$ in G. This allows us to estimate the difference between $-N(x)$ and w_{A^c} in G, where A^c is the complement of the set A.

Clearly $\widetilde{w}_A + \widetilde{w}_{A^c} = \widetilde{P} - \widetilde{N}$; then $\widetilde{N} + \widetilde{w}_{A^c} = \widetilde{P} - \widetilde{w}_A \leq t\theta$ in G. Since $\widetilde{N}(x) \geq \lambda/\Lambda \widetilde{P}(x) - \varepsilon_1$, we have that in G,

$$\begin{aligned} \widetilde{w}_{A^c}(x) &\leq -\widetilde{N}(x) + t\theta \\ &\leq -\frac{\lambda}{\Lambda}(1 - t\theta) + t\theta + \varepsilon_1 \\ &\leq -\frac{\lambda}{2\Lambda} \qquad \text{if } \theta \text{ and } \varepsilon_1 \text{ are small enough (depending on } t). \end{aligned}$$

Consequently, $|\{\widetilde{w}_{A^c} \leq -\frac{\lambda}{2\Lambda}\} \cap B_r| \geq (1 - ct^{-p})|B_r|$.

For some small $\kappa > 0$, we define $v_c(x) = (\widetilde{w}_{A^c}(\kappa r x) + \frac{\lambda}{2\Lambda})^+$. We know $\mathrm{M}^+ v_c \geq -\varepsilon_1$ in B_2; thus we can apply Theorem 5.1 to $v_c(\kappa r x)$ for some small $r > 0$ and get

$$\begin{aligned} v_c(0) &\leq C\varepsilon_1 + C \int_{\mathbb{R}^n} \frac{|v_c(y)|}{1 + |y|^{n+\sigma}} \, \mathrm{d}y \\ &\leq C\varepsilon_1 + C \int_{|y| \leq \kappa^{-1}} \frac{|v_c(y)|}{1 + |y|^{n+\sigma}} \, \mathrm{d}y + C \int_{|y| > \kappa^{-1}} \frac{|v_c(y)|}{1 + |y|^{n+\sigma}} \, \mathrm{d}y. \end{aligned}$$

Using that $|\{v_c > 0\} \cap B_{\kappa^{-1}}| < Ct^{-p}\kappa^{-n}$,

$$\leq C\varepsilon_1 + C\kappa^{-n}t^{-p} + C \int_{|y| > \kappa^{-1}} \frac{\widetilde{w}_{A^c}(\kappa r y)^+}{1 + |y|^{n+\sigma}} \, \mathrm{d}y.$$

Since $r < 1$, we can bound the third term independently of r:

$$\begin{aligned} &\leq C\varepsilon_1 + C\kappa^{-n}t^{-p} + C \int_{|y| > \kappa^{-1}} \frac{2(\kappa |y|)^{\alpha}}{1 + |y|^{n+\sigma}} \, \mathrm{d}y \\ &\leq C\varepsilon_1 + C\kappa^{-n}t^{-p} + C\kappa^{\sigma}. \end{aligned}$$

Thus we can choose κ and ε_1 so that $C\varepsilon_1 + C\kappa^{\sigma} < \lambda/(8\Lambda)$ and then t such that $C\kappa^{-n}t^{-p} < \lambda/(8\Lambda)$. Therefore, we get the following estimate:

$$v_c(0) \leq \frac{\lambda}{4\Lambda}.$$

But this means that $\widetilde{w}_{A^c}(0) \leq -\frac{\lambda}{4\Lambda}$ which is a contradiction since $\widetilde{w}_{A^c}(0) = 0$.

The contradiction comes from saying that $\widetilde{P}(x_0) \geq (1 - \theta)$ for some x_0 in B_r. Thus $\widetilde{P} < (1-\theta)$ in B_r. In the original scale, this means that $P \leq (1-\theta)^{k+1}$ in $B_{r^{k+1}}$, which finishes the inductive step and the proof. □

Using Lemma 9.2, we can finally prove Theorem 1.1.

Proof of Theorem 1.1. As it was mentioned before, the case $\sigma \leq 1$ is already covered in [3], so we prove the case $\sigma \in (1, 2)$ only.

Let us consider the fractional Laplacian of order σ,

$$-(-\triangle)^{\sigma/2} u(x) = c_\sigma (2 - \sigma) \int_{\mathbb{R}^n} \delta u(x, y) \frac{1}{|y|^{n+\sigma}} \, \mathrm{d}y,$$

where the constant c_σ remains bounded below and above for $\sigma \in (1,2)$.

Let b be a bump function as in Lemma 7.1. For $x \in B_{1/4}$ we have the identity

$$\begin{aligned}&(-\triangle)^{\sigma/2}u(0) - (-\triangle)^{\sigma/2}u(x)\\ &\qquad = c_\sigma \left(P(x) - N(x) + \int_{\mathbb{R}^n \setminus B_{1/2}} (\delta u(x,y) - \delta u(0,y)) \frac{(2-\sigma)}{|y|^{n+\sigma}} \, \mathrm{d}y \right).\end{aligned}$$

The third term is bounded by $C||u||_{L^\infty}|x|$ and the first two by $C\,\|u\|_{L^\infty(\mathbb{R}^n)}\,|x|^\alpha$. Thus, for $x \in B_{1/4}$,

$$|(-\triangle)^{\sigma/2}u(x) - (-\triangle)^{\sigma/2}u(0)| \le C\,\|u\|_{L^\infty(\mathbb{R}^n)}\,|x|^\alpha.$$

Therefore, by a standard translation of the estimate, we obtained that $(-\triangle)^{\sigma/2}u \in C^\alpha(B_{1/2})$, with the estimate

$$\left\|(-\triangle)^{\sigma/2}u\right\|_{C^\alpha(B_{1/2})} \le C\,\|u\|_{L^\infty(\mathbb{R}^n)}\,.$$

But if $(-\triangle)^{\sigma/2}u \in C^\alpha$, then $u \in C^{\sigma+\alpha}$ with the corresponding estimate from a classical result (see for example [10]). So we finish the proof. ☐

Remark 9.3. We have not used the homogeneity of I essentially in any proof in this paper. With the same arguments we can obtain the same regularity result for equations of the form

$$\inf_{a\in\mathcal{A}} (L_a u(x) + b_a) = \inf_{a\in\mathcal{A}} \left(\int_{\mathbb{R}^n} (u(x+y) + u(x-y) - 2u(x)) K_a(y) \, \mathrm{d}y + b_a \right) = 0$$

for a bounded family of real numbers b_a. This is the general form of a concave uniformly elliptic nonlocal operator of order σ.

For the estimates, we would have to include the values of b_a in the right-hand side:

$$||u||_{C^{\sigma+\alpha}(B_{1/2})} \le C(\|u\|_{L^\infty} + \sup_a b_a).$$

Remark 9.4. The assumption $u \in L^\infty(\mathbb{R}^n)$ is not sharp. It could easily be replaced in all estimates by $u \in L^1(\mathbb{R}^n, 1/(1+|y|^{n+\sigma}))$. We kept the L^∞ norm for simplicity of the exposition.

Acknowledgments. Both authors were partially supported by NSF grants. Part of this work took place while both authors were visiting the Institute for Advanced Study during the special program in geometric PDE. The authors would like to thank the referee for pointing out a number of misprints and corrections in an earlier version of this paper.

References

[1] H. Abels and M. Kassmann, An analytic approach to purely nonlocal Bellman equations arising in models of stochastic control, *J. Differential Equations* **236** (2007), 29–56. MR `2319919`. Zbl `1119.47048`. http://dx.doi.org/10.1016/j.jde.2006.12.013.

[2] L. Caffarelli and X. Cabré, *Fully Nonlinear Elliptic Equations*, *Amer. Math. Soc. Colloq. Publ.* **43**, Amer. Math. Soc., Providence, RI, 1995. MR `1351007`. Zbl `0834.35002`.

[3] L. Caffarelli and L. Silvestre, Regularity theory for fully nonlinear integro-differential equations, *Comm. Pure Appl. Math.* **62** (2009), 597–638. MR `2494809`. Zbl `1170.45006`. http://dx.doi.org/10.1002/cpa.20274.

[4] ———, On the Evans-Krylov theorem, *Proc. Amer. Math. Soc.* **138** (2010), 263–265. MR `2550191`. Zbl `1180.35229`.

[5] ———, Regularity results for nonlocal equations by approximation, *Archive of Rational Mechanics and Analysis* **200** (2011), 59–88. http://dx.doi.org/10.1007/s00205-010-0336-4.

[6] L. C. Evans, Classical solutions of fully nonlinear, convex, second-order elliptic equations, *Comm. Pure Appl. Math.* **35** (1982), 333–363. MR `0649348`. Zbl `0469.35022`. http://dx.doi.org/10.1002/cpa.3160350303.

[7] N. V. Krylov, Boundedly inhomogeneous elliptic and parabolic equations, *Izv. Akad. Nauk SSSR Ser. Mat.* **46** (1982), 487–523, 670. MR `0661144`. Zbl `0529.35026`.

[8] J.-L. Menaldi and M. Robin, Ergodic control of reflected diffusions with jumps, *Appl. Math. Optim.* **35** (1997), 117–137. MR `1424785`. Zbl `0876.93095`. http://dx.doi.org/10.1007/s002459900040.

[9] R. Mikulyavichyus and G. Pragarauskas, Classical solutions of boundary value problems for some nonlinear integro-differential equations, *Liet. Mat. Rink.* **34** (1994), 347–361. MR `1355243`. http://dx.doi.org/10.1007/BF02336537.

[10] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, *Princeton Math. Ser.* **30**, Princeton Univ. Press, Princeton, N.J., 1970. MR `0290095`. Zbl `0207.13501`.

(Received: May 8, 2009)
(Revised: September 26, 2010)

The University of Texas at Austin, Austin, TX
E-mail: caffarel@math.utexas.edu

The University of Chicago, Chicago, IL
E-mail: luis@math.uchicago.edu