

## $C^{\sigma,\alpha}$ ESTIMATES FOR CONCAVE, NON-LOCAL PARABOLIC EQUATIONS WITH CRITICAL DRIFT

HÉCTOR CHANG LARA AND GONZALO DÁVILA

Communicated by Stig-Olof Londen

**ABSTRACT.** Given a concave integro-differential operator  $I$ , we study regularity for solutions of fully nonlinear, nonlocal, parabolic equations of the form  $u_t - Iu = 0$ . The kernels are assumed to be smooth but non necessarily symmetric, which accounts for a critical non-local drift. We prove a  $C^{\sigma+\alpha}$  estimate in the spatial variable and  $C^{1,\alpha}$  estimates in time assuming time regularity for the boundary data. The estimates are uniform in the order of the operator  $I$ , hence allowing us to extend the classical Evans-Krylov result for concave parabolic equations.

**1. Introduction.** In this work, we are interested in studying regularity of solutions of

$$(1.1) \quad u_t - \inf_{L \in \mathcal{L}} Lu = 0 \text{ in } B_1 \times (-1, 0],$$

where  $L$  is a linear integro differential operator of the form,

$$Lu(x) := (2 - \sigma) \int_{\mathbb{R}^n} \delta u(x; y) \frac{K(y)}{|y|^{n+\sigma}} dy + b \cdot Du(x),$$

and  $\delta u$  is given by

$$\delta u(x; y) := u(x + y) - u(x) - Du(x) \cdot y \chi_{B_1}(y).$$

Here, the class  $\mathcal{L}$  corresponds to family of linear operators with kernels  $K/|y|^{n+\sigma}$  comparable to the kernel of the fractional Laplacian of order  $\sigma \in [1, 2)$ , that is,  $\lambda/|y|^{n+\sigma} \leq K(y)/|y|^{n+\sigma} \leq \Lambda/|y|^{n+\sigma}$ . In Section 2,

---

2010 AMS *Mathematics subject classification.* Primary 35B45, 35B65, 35D40, 35K55, 35R09.

*Keywords and phrases.* Non-local parabolic equations, fully non-linear concave operators, Evans-Krylov estimate.

The second author was partially supported by Fondecyt grant No. 11150880. Received by the editors on November 4, 2014, and in revised form on December 3, 2015.

we give a more precise definition of this class, but at this point we stress the fact that we are not assuming that the kernels are symmetric.

As was discussed in a previous paper [10], the odd part of the kernel brings a non-local drift term after rescaling the equation. This is the main reason why we included the first order term  $b \cdot Du$  above, since it allows us to account for the standard drift. We point out that the non-local drift term has the same order as the non-local diffusion. Therefore, it cannot be absorbed by rescaling the equation which creates additional difficulties. This is one of the main differences with respect to the second order case, where the lower order drift can be absorbed in the estimates of the purely second order equation at sufficiently small scales.

These types of equations appear naturally when studying stochastic control problems (see economic applications [11], ergodic control problems [16, 20]), in which the random part is given by a purely jump process and most of the time is non necessarily symmetric. The particular concave case can be seen as a one-player stochastic game, where at each step he can choose a strategy to minimize the expected value of some fixed function evaluated at the first exit point of a given domain.

In the local case ( $\sigma = 2$ ) this problem was first studied independently by Evans and Krylov (see [12, 15]. Recently, Caffarelli and Silvestre provided a new proof of the classic results in [6]). They obtain  $C^{2,\alpha}$  a priori estimates, and therefore, the existence of classical solutions by the continuity method.

A first study in the elliptic nonlocal case was done by Abels and Kassmann [1], in which the non-linearity was the maximum of two linear operators. They obtain  $H^\sigma$  interior estimates and  $H^{\sigma/2}$  up to the boundary using a variational approach. Also using a variational approach Felsinger and Kassmann proved interior regularity and a weak Harnack inequality in the parabolic nonlocal linear case, for a general kernel and general right hand side (see [13]).

Caffarelli and Silvestre proved in [5] that solutions of the elliptic problem  $Iu = 0$ , where  $I$  is a concave operator with smooth kernels, are  $C^{\sigma+\alpha}$ . By *concave operator*, we mean that, for smooth functions  $u$  and  $v$ ,  $s \in [0, 1]$ , we have  $I(su + (1-s)v) \geq sI(u) + (1-s)I(v)$ . The arguments in their work rely on the theory of viscosity solutions developed in [3, 4]. The regularity obtained is enough to evaluate the

operator in the classical sense. Moreover, the estimates obtained are independent of the order of the equation and extend the theory to the classical case.

A recent improvement of the aforementioned work, done by Serra [18], allows to remove the smoothness condition for symmetric kernels in order to prove  $C^{\sigma+\alpha}$  estimates. It proceeds by a compactness argument that blows up the solution, reducing the problem to a Liouville type of result. We need to stress that this work does not account for critical non-local drifts, as pointed out by the author, since the term does persist after the blow up.

Regularity for parabolic nonlocal equations has been studied by the authors in [8, 9, 10] in which Hölder estimates are proven for general equations like (1.1) with a non zero right hand side. Recent advances include the work of Serra [17] for  $C^{1,\alpha}$  estimates with rough kernels and the work of Jin and Xiong [14] for higher order, optimal Schauder estimates.

We extend the ideas of [5] to the parabolic nonlocal case to prove the desired  $C^{\sigma+\alpha}$  interior regularity in the spatial variable. The order  $\sigma$  is assumed at least one in order for the drift to be at most comparable with the diffusion. On the other hand, for  $\sigma \in (0, 1]$ , the  $C^{1,\alpha}$  estimates established in [10] already give classical solutions. In the upcoming results, we require the boundary data to have first derivatives in time. For general boundary data, one cannot expect  $C^{1,\alpha}$  regularity in time, even for the fractional heat equation; an example is discussed in [8].

We now state the main result.

**Theorem 1.1.** *Let  $\sigma \in [1, 2)$ ,  $\mathcal{L} \subseteq \mathcal{L}_2^\sigma(\lambda, \Lambda, \beta)$  (sufficiently smooth kernels to be defined), and suppose  $u$  satisfies in the viscosity sense,*

$$u_t - \inf_{L \in \mathcal{L}} Lu = 0 \text{ in } B_1 \times (-1, 0].$$

*Then, there is some  $\alpha \in (0, 1)$  and  $C > 0$ , depending only on  $n, \lambda, \Lambda$  and  $\beta$  such that  $u \in C^{1,\alpha}(B_{1/2} \times (-1/2, 0])$  and, for all fixed  $t$ ,  $u(\cdot, t) \in C^{\sigma+\alpha}(B_{1/2})$ . More precisely, for  $t \in [-1/2, 0]$ , we have*

$$\|u(\cdot, t)\|_{C^{\sigma+\alpha}(B_{1/2})} \leq C(\|u\|_{L^1((-1,0] \rightarrow L^1(\omega_\sigma))} + [u\chi_{B_1^c}]_{C^{0,1}((-1,0] \rightarrow L^1(\omega_\sigma))}).$$

The paper is divided as follows. In Section 2, we introduce the

family of operators we are considering, the notion of viscosity solution and recall some properties. We also state some previous results that we need for the rest of this work. We use the concavity of the non-linearity in Section 3 to determine an equation for the average of a given solution; in particular, we get an equation for the fractional Laplacian. In Section 4, we use the previous equation to obtain  $C^\sigma$  estimates on the fractional Laplacian of the solution. Finally, in Section 5, we prove a diminish of oscillation lemma for the fractional Laplacian, which implies our main theorem.

**2. Preliminaries and viscosity solutions.** We begin this section with some preliminary notation.

The cylinder of radius  $r$ , height  $\tau$  and center  $(x, t)$  in  $\mathbb{R}^n \times \mathbb{R}$  is denoted by  $C_{r,\tau}(x, t) := B_r(x) \times (t - \tau, t]$ . Whenever we omit the center we are assuming that they are centered at the origin in space and time.

Given the scaling properties of linear operators with non symmetric kernels discussed in [7, 10], it is reasonable to enlarge the family of linear operators to include (classical) drift terms. With this in mind, let us introduce the following notation where the time variable has been omitted

$$(2.1) \quad \begin{aligned} L_{K,b}^\sigma u(x) &:= (2 - \sigma) \int_{\mathbb{R}^n} \delta u(x; y) \frac{K(y)}{|y|^{n+\sigma}} dy + b \cdot Du(x), \\ \delta u(x; y) &:= u(x + y) - u(x) - Du(x) \cdot y \chi_{B_1}(y). \end{aligned}$$

Consider kernels bounded from above and away from zero  $0 < \lambda \leq K \leq \Lambda < \infty$ . Since the drift comes not only from the term  $b \cdot D$  but also from the odd part of the kernel (after rescaling), we will assume that they are controlled in the following way:

$$(2.2) \quad \sup_{r \in (0,1)} \left| b + (2 - \sigma) \int_{B_1 \setminus B_r} \frac{yK(y)}{|y|^{n+\sigma}} dy \right| \leq \beta.$$

For a thorough discussion on the scaling properties of these types of operators we refer to [10].

We denote by  $\mathcal{L}_0^\sigma(\lambda, \Lambda, \beta)$  the family of all linear operators above and suppress some of its parameters whenever it is clear from the context; usually we just write  $\mathcal{L}_0$ . A sufficient regularity/integrability

condition on  $u$  to evaluate  $L_{K,b}^\sigma u(x)$  is  $u \in C^{1,1}(x) \cap L^1(\omega_\sigma)$  where  $\omega_\sigma(y) = \min(1, |y|^{-(n+\sigma)})$ .

An approach in order to prove higher regularity estimates is to assume smoothness of the kernels. This corresponds to the initial approach taken in [3, 4, 5] in order to use an integration by parts techniques to control rough oscillations of the boundary data of the solution. This work uses the same technique; therefore, let us define the class  $\mathcal{L}_1^\sigma(\lambda, \Lambda, \beta) \subseteq \mathcal{L}_0$  by

$$|DK(y)| \leq \Lambda|y|^{-1},$$

where  $K$  is a kernel defining a linear operator of the class. Similarly, we define  $\mathcal{L}_2^\sigma(\lambda, \Lambda, \beta) \subseteq \mathcal{L}_1$

$$|D^2K(y)| \leq \Lambda|y|^{-2}.$$

Recall that the smoothness hypotheses of the previous works have been lifted in [17, 18] for symmetric kernels. Their techniques apply also if drift terms or lower order terms are included because of scaling considerations. However, an odd kernel renews the non-local drift (see [10]) and keeps it comparable to the diffusion even when the scale goes to zero, so their technique does not apply directly.

Given  $\mathcal{L} \subseteq \mathcal{L}_0$  a non linearity  $I$  is given by a function  $I : \Omega \times (t_1, t_2] \times \mathbb{R}^\mathcal{L} \rightarrow \mathbb{R}$ , such that

$$Iu(x, t) := I(x, t, (Lu(x, t))_{L \in \mathcal{L}}).$$

$I$  is considered to be elliptic if it is increasing in  $\mathbb{R}^\mathcal{L}$ .

The nonlinearity in our main theorem is constructed from  $\mathcal{L} \subseteq \mathcal{L}_2$  such that

$$Iu = \mathcal{M}_\mathcal{L}^- u := \inf_{L \in \mathcal{L}} Lu.$$

It satisfies the following uniform ellipticity relation with the extremal operators

$$\mathcal{M}_\mathcal{L}^-(u - v) \leq Iu - Iv \leq \mathcal{M}_\mathcal{L}^+(u - v).$$

where  $\mathcal{M}_\mathcal{L}^+ := \sup_{L \in \mathcal{L}} L$ .

Given a family  $\mathcal{L}$ , the operators

$$\mathcal{M}_\mathcal{L}^+ := \sup_{L \in \mathcal{L}} L, \quad \mathcal{M}_\mathcal{L}^- := \inf_{L \in \mathcal{L}} L,$$

are usually referred to as the *maximal* and *minimal operators*. This definition is closely related with the classical second order maximal and minimal Pucci operators (see [2]).

**2.1. Viscosity solutions.** We recall some definitions pertaining to viscosity solutions  $u$  for the equation  $u_t - Iu = f$ . A test function  $\varphi$  needs to be sufficiently smooth/integrable near the contact point where the equation is tested. Moreover, qualitative properties, such as the continuity of  $Iu$ , require the tail of  $u$  to be at least continuous in time in the following integrable sense.

**Definition 2.1.** The space  $C((t_1, t_2] \mapsto L^1(\omega_\sigma))$  consists of all measurable functions  $u : \mathbb{R}^n \times (t_1, t_2] \rightarrow \mathbb{R}$  such that, for every  $t \in (t_1, t_2]$ ,

- (1)  $\|u(\cdot, t)^- \|_{L^1(\omega_\sigma)} < \infty$ .
- (2)  $\lim_{\tau \rightarrow 0^-} \|u(\cdot, t) - u(\cdot, t - \tau)\|_{L^1(\omega_\sigma)} = 0$ .

Here  $u(\cdot, t) = u(\cdot, t)^+ - u(\cdot, t)^-$ , with

$$u(\cdot, t)^+ = \max\{u(\cdot, t), 0\}, \quad u(\cdot, t)^- = \max\{-u(\cdot, t), 0\},$$

that is,  $u$  has been decomposed in its positive and negative part for each  $t$ .

Next we define our suitable test functions. We point out that, in order to evaluate the operator  $I$  in the classical sense, we only need the function to be  $C^{1,1}$  punctually (see [3]). As mentioned before, we require the tail of  $u$  to be continuous in an integrable sense in order to have continuity of  $I$ .

**Definition 2.2** (Test functions). A test function is a pair  $(\varphi, C_{r,\tau}(x, t))$ , such that  $\varphi \in C_x^{1,1}C_t^1(C_{r,\tau}(x, t)) \cap C((t - \tau, t] \mapsto L^1(\omega_\sigma))$ .

Whenever the cylinder in the Definition 2.2 becomes irrelevant we will refer to the test function  $(\varphi, C_{r,\tau}(x, t))$  just by  $\varphi$ .

**Definition 2.3.** Given a function  $u$  and a test function  $\varphi$ , we say that  $\varphi$  touches  $u$  from below at  $(x, t)$  if there is  $\tau > 0$  such that,

- (1)  $\varphi(x, t) = u(x, t)$ ,
- (2)  $\varphi(y, s) \leq u(y, s)$  for  $(y, s) \in \mathbb{R}^n \times (t - \tau, t]$ .

A similar definition for *contact from above* can be done by reversing the previous inequality.

**Definition 2.4** (Viscosity (super) solutions). Given an elliptic operator  $I$  and a function  $f$ , a function  $u \in C(\Omega \times (t_1, t_2]) \cap C((t_1, t_2] \mapsto L^1(\omega_\sigma))$  is said to be a viscosity super solution to  $u_t - Iu \geq f$  in  $\Omega \times (t_1, t_2]$  if, for every lower semi-continuous test function  $(\varphi, C_{r,\tau}(x, t))$  touching  $u$  from below at  $(x, t) \in \Omega \times (t_1, t_2]$ , we have that  $\varphi_{t-}(x, t) - I\varphi(x, t) \geq f(x, t)$ .

Recall that  $\varphi_{t-}$  denotes the left time derivative of  $\varphi$ , natural for time evolution problems.

The definition of  $u$  being a viscosity *sub* solution to  $u_t - Iu \leq f$  in  $\Omega \times (t_1, t_2]$  is done similarly to the definition of super solution replacing contact from below by contact from above and reversing the last inequality. A viscosity solution to  $u_t - Iu = f$  in  $\Omega \times (t_1, t_2]$  is a function which is a super and a sub solution simultaneously.

**2.2. Previous results.** Several qualitative results for viscosity solutions of our parabolic equation such as the stability, comparison principle and the existence of (viscosity) solutions have been established in [8, 9, 10]. We recall at this point some quantitative estimates for the solutions which will be used in this work.

**Theorem 2.5** (Stability). *Let  $I_k$  be a sequence of uniformly elliptic operators with respect to  $\mathcal{L}_0$ . Let  $\{u_k\}_{k \geq 1} \subseteq LSC(\bar{\Omega} \times (-T, 0]) \cap C(-T, 0; L^1(\omega))$  be such that*

- (i)  $(u_k)_t - I_k u_k \geq f_k$  in the viscosity sense in  $\Omega \times (-T, 0]$ ,
- (ii)  $u_k \rightarrow u$  in the  $\Gamma$  sense in  $\Omega \times (-T, 0]$  and in  $C(-T, 0; L^1(\omega))$ ,
- (iii)  $f_k \rightarrow f$  locally uniformly in  $\Omega \times (-T, 0]$ ,
- (iv)  $I_k \rightarrow I$  weakly in  $\Omega \times (-T, 0]$  (with respect to  $\omega$ ),
- (v)  $|u_k(x, t)| \leq C$  for every  $(x, t) \in \Omega \times (-T, 0]$ .

*Then, also  $u_t - Iu \geq f$  in the viscosity sense in  $\Omega \times (-T, 0]$ .*

**Theorem 2.6** (Point estimate). *Let  $\sigma \in [1, 2)$ . Suppose  $u \geq 0$  satisfies*

$$u_t - \mathcal{M}_{\mathcal{L}_0^-} u \geq -f(t) \quad \text{in } C_{2r, 2r^\sigma}(0, r^\sigma).$$

Then, for every  $s \geq 0$ ,

$$\frac{|\{u > s\} \cap C_{r,r^\sigma}|}{|C_{r,r^\sigma}|} \leq C \left( \inf_{C_{r,r^\sigma}(0,r^\sigma)} u + r^\sigma \int_{-r^\sigma}^{r^\sigma} f^+(s) ds \right)^\varepsilon s^{-\varepsilon},$$

for some constants  $\varepsilon$  and  $C$  depending only on  $n, \lambda, \Lambda$  and  $\beta$ .

The oscillation lemma provided in [10] controls the pointwise size of a non negative sub solution in terms of an integral norm.

**Lemma 2.7** (Oscillation lemma). *Let  $\mathcal{L} \subseteq \mathcal{L}_0, I : \Omega \times (t_1, t_2] \times \mathbb{R}^\mathcal{L} \rightarrow \mathbb{R}$  be uniformly elliptic and such that  $I0 = 0$ . Let  $u$  satisfy,*

$$u_t - Iu \leq f \quad \text{in } \Omega \times (t_1, t_2].$$

Then, for every  $\Omega' \times (t'_1, t_2] \subset\subset \Omega \times (t_1, t_2]$ ,

$$\sup_{\Omega' \times (t'_1, t_2]} u^+ \leq C \left( \|u^+\|_{L^1((t_1, t_2] \rightarrow L^1(\omega_\sigma))} + \|f^+\|_{L^1((t_1, t_2] \rightarrow L^\infty(\Omega))} \right),$$

for some universal  $C > 0$ , independent of  $\sigma \in [1, 2)$ , depending on the domains.

**Theorem 2.8** (Hölder regularity). *Let  $u$  satisfy*

$$\begin{aligned} u_t - \mathcal{M}_{\mathcal{L}_0}^+ u &\leq f(t) && \text{in } C_{1,1}, \\ u_t - \mathcal{M}_{\mathcal{L}_0}^- u &\geq -f(t) && \text{in } C_{1,1}, \end{aligned}$$

Then there is some  $\alpha \in (0, 1)$  and  $C > 0$ , depending only on  $n, \lambda, \Lambda$  and  $\beta$ , such that, for every  $(y, s)$  and  $(x, t) \in C_{1/2, 1/2}$

$$\frac{|u(y, s) - u(x, t)|}{(|x - y| + |t - s|^{1/\sigma})^\alpha} \leq C \left( \|u\|_{L^1((-1, 0] \rightarrow L^1(\omega_\sigma))} + \|f\|_{L^1(0, 1)} \right).$$

We say that  $I$  is translation invariant in space or time if the function  $I$  does not depend on the variable  $x$  or  $t$ . *Translation invariant*, without making reference to the space or time variable, means that it is translation invariant with respect to both.

**Theorem 2.9** (Regularity for translation invariant operators). *Let  $\mathcal{L} \subseteq \mathcal{L}_1, I : \mathbb{R}^\mathcal{L} \rightarrow \mathbb{R}$  be uniformly elliptic, translation invariant and*

such that  $I0 = 0$ . Let  $u$  satisfy

$$u_t - Iu = f(t) \text{ in } C_1.$$

Then there is some  $\alpha \in (0, 1)$  and  $C > 0$ , depending only on  $n, \lambda, \Lambda$  and  $\beta$ , such that for every  $(y, s), (x, t) \in C_{1/2,1/2}$ ,

$$|Du(x, t)| + \frac{|Du(x, t) - Du(y, s)|}{(|x - y| + |t - s|^{1/\sigma})^\alpha} \leq C (\|u\|_{L^1((-1,0] \rightarrow L^1(\omega_\sigma))} + \|f\|_{L^1(0,1)}).$$

The previous theorem does not give more regularity in time even if  $I$  is translation invariant in time and  $f \equiv 0$ . In [8], the authors gave an example of a function, which is not better than Lipschitz in its time variable, solving the fractional heat equation. However, better regularity in time can be obtained via the oscillation lemma if the Dirichlet data has a smoothness condition controlled by

$$[u]_{C^{0,1}((t_1,t_2] \rightarrow L^1(\omega_\sigma))} := \sup_{(t-\tau,t] \subseteq (t_1,t_2]} \frac{\|u(t) - u(t-\tau)\|_{L^1(\omega_\sigma)}}{\tau}.$$

**Theorem 2.10** (Further regularity in time). *Let  $\mathcal{L} \subseteq \mathcal{L}_0, I : \mathbb{R}^{\mathcal{L}} \rightarrow \mathbb{R}$  be uniformly elliptic and translation invariant such that  $I0 = 0$ . Let  $u$  satisfy*

$$u_t - Iu = 0 \text{ in } C_{1,1}.$$

Then there is some  $\alpha \in (0, 1)$  and  $C > 0$ , depending only on  $n, \lambda, \Lambda$  and  $\beta$ , such that, for every  $(x, t), (y, s) \in C_{1/2,1/2}$  we have

$$|u_t(x, t)| + \frac{|u_t(x, t) - u_t(y, s)|}{(|x - y| + |t - s|^{1/\sigma})^\alpha} \leq C [u]_{C^{0,1}((-1,0] \rightarrow L^1(\omega_\sigma))}.$$

**3. Equations for  $Lu$  by concavity and translation invariance.**

We fix for this section  $\sigma \in [1, 2), \mathcal{L} \subseteq \mathcal{L}_2$ , and  $u$  such that,

$$u_t - \mathcal{M}_{\mathcal{L}}^- u = 0 \text{ in } C_{8,3},$$

$$\|u\|_{L^\infty((-3,0] \rightarrow L^1(\omega_\sigma))} + [u]_{C^{0,1}((-3,0] \rightarrow L^1(\omega_\sigma))} \leq 1.$$

We can assume that  $u$  is a classical solution with smooth boundary and initial data. Otherwise, we approximate  $u$  by a sequence of classical solutions with smooth boundary and initial data and recover the

estimates of this section in the limit by the regularization procedure described in [9]. When approximating the solution we need to make sure that the a priori estimates are independent of (fractional) derivatives of  $u$ .

Note that many results in this and the following sections can be obtained by controlling  $\|u\|_{L^1((-5,0] \rightarrow L^1(\omega_\sigma))}$ , instead of the  $L^\infty$  norm. However, when the bound for  $\|u\|_{L^1((-5,0] \rightarrow L^1(\omega_\sigma))}$  is coupled with the bound for  $[u]_{C^{0,1}((-5,0] \rightarrow L^1(\omega_\sigma))}$ , one can check that this implies a bound for the  $L^\infty$  norm.

It is convenient for this section to introduce the following notation. Given  $K(y) \geq 0$ , let

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

We denote the convolution by

$$v * w(x) := \int_{\mathbb{R}^n} w(x - y)v(y) dy.$$

In particular, given that  $K \geq 0$  goes to zero near the origin with at least a quadratic rate, then we can decompose a linear operator as:

$$L_{K,b}^\sigma u = \left( K^\sigma * - \|K^\sigma\|_1 - \left( \int_{B_1} y K^\sigma(y) dy - b \right) \cdot D \right) u.$$

**Property 3.1.** Let  $\alpha \in \mathbb{R}$ ,  $b \in \mathbb{R}^n$  and  $\eta \geq 0 \in L^1(\mathbb{R}^n)$ . Then the following holds for any regular function  $v$ :

- (1) **Homogeneity.**  $\mathcal{M}_\mathcal{L}^\pm(\alpha v) = \alpha \mathcal{M}_\mathcal{L}^\pm v.$
- (2) **Translation.**  $\mathcal{M}_\mathcal{L}^-(b \cdot Dv) \leq b \cdot D\mathcal{M}_\mathcal{L}^\pm v \leq \mathcal{M}_\mathcal{L}^+(b \cdot Dv).$
- (3) **Concavity.**  $\eta * \mathcal{M}_\mathcal{L}^- v \leq \mathcal{M}_\mathcal{L}^\pm(\eta * v) \leq \eta * \mathcal{M}_\mathcal{L}^+ v.$

**Corollary 3.2.** Let  $K \geq 0$ ,  $b \in \mathbb{R}^n$  and  $\varphi \in C_0^\infty(B_2 \mapsto [0, 1])$  be such that  $\varphi = 1$  in  $B_1$ . Then

$$(L_{K,b}^\sigma)_t - \mathcal{M}_\mathcal{L}^+(L_{K,b}^\sigma) \leq ((1 - \varphi)K^\sigma)_t - \mathcal{M}_\mathcal{L}^-(((1 - \varphi)K^\sigma)_t),$$

in  $C_{6,3}$ . In particular, if  $\text{supp } K \subseteq B_1$ , then

$$(L_{K,b}^\sigma)_t - \mathcal{M}_\mathcal{L}^+(L_{K,b}^\sigma) \leq 0 \quad \text{in } C_{6,3}.$$

*Proof.* Let, for  $\varepsilon \in (0, 1)$ ,  $K_\varepsilon := \chi_{B_\varepsilon} K$ . We decompose the operator  $L_{K_\varepsilon, b}^\sigma$  as a sum of a local and nonlocal operator, where the nonlocal operator is given by  $(1 - \varphi)K^\sigma*$ ,

$$\begin{aligned} L_{K_\varepsilon, b}^\sigma &= L + NL, \\ &:= (L_{K_\varepsilon, b}^\sigma - K_\varepsilon^\sigma(1 - \varphi)*) + K_\varepsilon^\sigma(1 - \varphi)*, \\ &= \left( \varphi K_\varepsilon^\sigma * - \|K_\varepsilon^\sigma\|_1 - \left( \int_{B_1} y K_\varepsilon^\sigma(y) dy - b \right) \cdot D \right) + K_\varepsilon^\sigma(1 - \varphi)*. \end{aligned}$$

Then,

$$\begin{aligned} (L_{K_\varepsilon, b}^\sigma u)_t - \mathcal{M}_{\mathcal{L}}^+(L_{K_\varepsilon, b}^\sigma u) &\leq ((Lu)_t - \mathcal{M}_{\mathcal{L}}^+(Lu)) \\ &\quad + ((NLu)_t - \mathcal{M}_{\mathcal{L}}^-(NLu)) \\ &\leq L(u_t - \mathcal{M}_{\mathcal{L}}^-u) + ((NLu)_t - \mathcal{M}_{\mathcal{L}}^-(NLu)). \end{aligned}$$

In  $C_{6,3}$ , the first term is zero as the local operator  $L$  does not take into account the values of  $(u_t - \mathcal{M}_{\mathcal{L}}^-u)$  outside of  $B_8$ . The result now follows by the stability Theorem 2.5 and letting  $\varepsilon \searrow 0$  (see [10] for more details).  $\square$

**Property 3.3** (Integration by parts). Let  $K \geq 0$ ,  $b \in \mathbb{R}^n$ ,  $(\bar{K}(y), \bar{b}) := (K(-y), -b)$  and, for  $L = L_{K, b}^\sigma$ ,  $\bar{L} = L_{\bar{K}, \bar{b}}^\sigma$ . Then the following holds for any pair of regular/integrable functions  $v$  and  $w$ :

$$\int_{\mathbb{R}^n} vLw = \int_{\mathbb{R}^n} w\bar{L}v.$$

In particular,

$$L(v * w) = v * (Lw) = (\bar{L}v) * w.$$

**Corollary 3.4.** For  $L_{K, b}^\sigma \in \mathcal{L}_2$ , it holds that

$$(L_{K, b}^\sigma u)_t - \mathcal{M}_{\mathcal{L}}^+(L_{K, b}^\sigma u) \leq C \quad \text{in } C_{6,3},$$

for some universal constant  $C > 0$ .

*Proof.* Corollary 3.2 tells us that it suffices to estimate  $([(1 - \varphi)K^\sigma] * u)_t - \mathcal{M}_{\mathcal{L}}^-([(1 - \varphi)K^\sigma] * u)$  in  $C_{6,3}$ ,

$$\begin{aligned} ((1 - \varphi)K^\sigma] * u)_t &= [(1 - \varphi)K^\sigma] * u_t \\ &\leq C[u]_{C^{0,1}((-\infty, 0] \rightarrow L^1(\omega_\sigma))} = C, \end{aligned}$$

and so we get

$$\begin{aligned} \mathcal{M}_{\mathcal{L}}^-([(1 - \varphi)K^\sigma] * u) &\geq \inf_{L \in \mathcal{L}_2} L([(1 - \varphi)K^\sigma] * u) \\ &= \inf_{L \in \mathcal{L}_2} (\bar{L}[(1 - \varphi)K^\sigma] * u) \\ &\geq -C. \end{aligned}$$

In the last inequality, we used that  $|DK(y)| \leq \Lambda|y|^{-1}$ ,  $|D^2K(y)| \leq \Lambda|y|^{-2}$  and  $\|u\|_{L^\infty((-3,0] \rightarrow L^1(\omega_\sigma))} \leq 1$ . □

From now on, we denote, for  $r_1 > r_2 > 0$ ,  $\psi_{r_1,r_2} \in C_0^\infty(B_{r_1} \rightarrow [0, 1])$  such that  $\psi_{r_1,r_2} = 1$  in  $B_{r_2}$ .

**Corollary 3.5.** *Let  $6 \geq r_1 > r_2 > 0$ ,  $K \geq 0$ ,  $b \in \mathbb{R}^n$ , such that either  $L_{K,b}^\sigma \in \mathcal{L}_2$  or  $|b| \leq \beta'$ ,  $\text{supp } K \subseteq B_1$  and  $K(y) \in [0, \Lambda']$ . Then,*

$$(\psi_{r_1,r_2} L_{K,b}^\sigma)_t - \mathcal{M}_{\mathcal{L}}^+(\psi_{r_1,r_2} L_{K,b}^\sigma) \leq C \quad \text{in } C_{r_2,3},$$

for some universal constant  $C > 0$  depending also on  $r_1, r_2, \beta'$  and  $\Lambda'$ .

*Proof.* We use either Corollary 3.2 or 3.4 to get that  $\psi_{r_1,r_2} L_{K,b}^\sigma$  satisfies the following inequality in  $C_{r_2,3}$ :

$$\begin{aligned} &(\psi_{r_1,r_2} L_{K,b}^\sigma)_t - \mathcal{M}_{\mathcal{L}}^+(\psi_{r_1,r_2} L_{K,b}^\sigma) \\ &\leq C + \sup_{L \in \mathcal{L}_2} L((1 - \psi_{r_1,r_2})L_{K,b}^\sigma) \\ &= C + \sup_{L \in \mathcal{L}_2} K_L * ((1 - \psi_{r_1,r_2})L_{K,b}^\sigma), \end{aligned}$$

where  $K_L$  is the kernel associated to  $L \in \mathcal{L}_2$ . Notice that there are cancellations provided by the fact that  $(1 - \psi_{r_1,r_2})$  and its gradient are zero in  $B_{r_2}$ .

Now we take a closer look at  $[K_L * ((1 - \psi_{r_1,r_2})L_{K,b}^\sigma)](x, t)$  for  $(x, t) \in C_{r_2,3}$ ,

$$\begin{aligned} &[K_L * ((1 - \psi_{r_1,r_2})L_{K,b}^\sigma)](x, t) \\ &= [(K_L(1 - \psi_{r_1,r_2}(x + \cdot))) * L_{K,b}^\sigma](x, t) \\ &= [L_{K,\bar{b}}^\sigma(K_L(1 - \psi_{r_1,r_2}(x + \cdot))) * u](x, t) \\ &\leq C. \end{aligned}$$

In the last inequality we used that  $|DK_L(y)| \leq \Lambda|y|^{-1}$ ,  $|D^2K_L(y)| \leq \Lambda|y|^{-2}$  and  $\|u\|_{L^\infty((-3,0] \rightarrow L^1(\omega_\sigma))} \leq 1$ . □

**4. Estimate for  $\Delta^{\sigma/2}u$ .** In this section, we keep the same assumptions as before:  $\sigma \in [1, 2)$ ,  $\mathcal{L} \subseteq \mathcal{L}_2$  and  $u$  is such that

$$u_t - \mathcal{M}_{\mathcal{L}}^- u = 0 \text{ in } C_{8,3},$$

$$\|u\|_{L^\infty((-3,0] \rightarrow L^1(\omega_\sigma))} + [u]_{C^{0,1}((-3,0] \rightarrow L^1(\omega_\sigma))} \leq 1.$$

**Lemma 4.1.** For  $K(y) \in [0, \Lambda]$ ,  $b \in B_\beta$ ,

$$\|L_{K,b}^\sigma u\|_{L^\infty(C_{1,1})} \leq C,$$

for some universal constant  $C$ .

*Proof.* We prove it in several steps. Here is the summary of the strategy:

- (1) For  $L \in \mathcal{L}$ , we bound  $Lu$  from below by using the equation for  $u$  and the control we have for  $u_t$  inside the domain.
- (2) For  $L \in \mathcal{L}$ , we integrate by parts to control  $\|Lu(t)\|_{L^1(\omega_\sigma)}$  and then apply Lemma 2.7 to bound  $Lu$  from above.
- (3) For general  $K$  and  $b$ , we use  $L^2$  theory to control  $\|L_{K,b}^\sigma u(t)\|_{L^1(\omega_\sigma)}$  and then apply Lemma 2.7 to bound  $L_{K,b}^\sigma u$  from above.
- (4) For general  $K$  and  $b$ , we apply the previous step to

$$(K'', b'') = \Lambda(K', b') - \lambda(K, b),$$

with  $L_{K',b'}^\sigma \in \mathcal{L}$  to bound  $L_{K,b}^\sigma u$  from below.

*Step 1.*  $L \in \mathcal{L}$ , then  $Lu \geq -C$  in  $C_{8,3}$ .

It follows from the equation for  $u$  and the regularity in time that

$$Lu \geq \mathcal{M}_{\mathcal{L}}^- u = u_t \geq -[u]_{C^{0,1}((-3,0] \rightarrow L^1(\omega_\sigma))}.$$

*Step 2.*  $L \in \mathcal{L}$ , then  $Lu \leq C$  in  $C_{4,2}$ .

For this step, we recall the definition of the cut-off function  $\psi_{r_1,r_2}$ : for  $r_1 > r_2 > 0$ ,  $\psi_{r_1,r_2} \in C_0^\infty(B_{r_1} \rightarrow [0, 1])$  such that  $\psi_{r_1,r_2} = 1$  in  $B_{r_2}$ .

We apply the oscillation lemma to  $\psi_{6,5}Lu$ . By Corollary 3.5,  $\psi_{6,5}Lu$  satisfies

$$(\psi_{6,5}Lu)_t - \mathcal{M}_{\mathcal{L}}^+(\psi_{6,5}Lu) \leq C \text{ in } C_{5,3}.$$

We now estimate  $\|(\psi_{6,5}Lu)^+\|_{L^1((-3,0] \rightarrow L^1(\omega_\sigma))}$ . As  $\psi_{6,5}Lu$  is bounded from below and compactly supported, all we need is to control the following integral

$$\int_{-3}^0 \int_{\mathbb{R}^n} \psi_{6,5} Lu = \int_{-3}^0 \int_{\mathbb{R}^n} (\bar{L}\psi_{6,5}) u \leq C.$$

By Lemma 2.7, we have that  $\psi_{6,5}Lu$  is bounded from above in  $C_{4,2}$ , where it coincides with  $Lu$ .

*Step 3.* Given  $K(y) \in [0, \Lambda']$  and  $b \in B_{\beta'}$ , then  $L_{K,b}^\sigma u \leq C$  in  $C_{1,1}$ .

Given  $L \in \mathcal{L}$ , the previous steps imply that  $Lu$  is bounded in  $C_{4,2}$ . From Fourier analysis techniques we then get that (see [6, Theorem 4.3]),

$$\begin{aligned} \|L_{K,b}^\sigma u(t)\|_{L^2(B_2)} &\leq C \|Lu(t)\|_{L^2(B_3)} \leq C \\ \implies \|L_{K\chi_{B_1},b}^\sigma u(t)\|_{L^1(B_2)} &\leq C + \|L_{K\chi_{B_1^c},0}^\sigma u(t)\|_{L^1(B_2)} \\ \implies \|\psi_{3,2}L_{K\chi_{B_1},b}^\sigma u\|_{L^1((-2,0] \rightarrow L^1(\omega_\sigma))} &\leq C. \end{aligned}$$

By Corollary 3.5,  $\psi_{3,2}L_{K\chi_{B_1},b}^\sigma u$  satisfies

$$\begin{aligned} (\psi_{3,2}L_{K\chi_{B_1},b}^\sigma u)_t - \mathcal{M}_{\mathcal{L}}^+(\psi_{3,2}L_{K\chi_{B_1},b}^\sigma u) &\leq C \quad \text{in } C_{2,2} \\ \int_{-4}^0 f(t) dt &\leq C. \end{aligned}$$

By Lemma 2.7,  $\psi_{3,2}L_{K\chi_{B_1},b}^\sigma u$  gets bounded from above in  $C_{1,1}$ . By the hypotheses, we also obtain the bound for  $\psi_{3,2}L_{K,b}^\sigma u$  in  $C_{1,1}$  where it coincides with  $L_{K,b}^\sigma u$ ,

$$\psi_{3,2}L_{K,b}^\sigma u \leq C + \psi_{3,2}L_{K\chi_{B_1^c},b}^\sigma u \leq C + \|u\|_{L^\infty((-1,0] \rightarrow L^1(\omega_\sigma))}.$$

*Step 4.* Given  $K(y) \in [0, \Lambda]$  and  $b \in B_\beta$ , then  $L_{K,b}^\sigma u \geq -C$  in  $C_{1,1}$ .

Consider  $L_{K',b'}^\sigma \in \mathcal{L}$ , and

$$L_{K'',b''}^\sigma := \Lambda L_{K',b'}^\sigma - \lambda L_{K,b}^\sigma,$$

such that  $|b''| \leq (\Lambda + \lambda)\beta$ , and  $K''(y) \in [0, \Lambda^2]$ . Given the result from the second step, it suffices to show that  $L_{K'',b''}^\sigma u \geq -C$  in  $C_{1,1}$ . This is just a consequence of applying the bound of the third step to  $L_{K'',b''}^\sigma u$ .  $\square$

**Corollary 4.2.** *There is a universal constant  $C > 0$  such that*

$$(2 - \sigma) \int_{\mathbb{R}^n} \frac{|\delta u(x, t; y)|}{|y|^{n+\sigma}} dy \leq C \quad \text{in } C_{1,1}.$$

In particular, by Morrey estimates, we have that  $u \in C_x^\alpha(C_{1,1})$  for every  $\alpha \in [1, \sigma)$ , see [21].

*Proof.* Using  $K(y) := \Lambda(2 - \sigma)|y|^{-(n+\sigma)}$  in the previous lemma, we get

$$(2 - \sigma) \int_{\mathbb{R}^n} \frac{\delta u(x, t; y)}{|y|^{n+\sigma}} dy \geq -C \quad \text{in } C_{1,1}.$$

Fixing  $(x, t) \in C_{1,1}$  and using  $K(y) := \Lambda(2 - \sigma) \operatorname{sign}(\delta u(x, t; y))|y|^{-(n+\sigma)}$  in the previous lemma, we get

$$(2 - \sigma) \int_{\mathbb{R}^n} \frac{\delta^+ u(x, t; y)}{|y|^{n+\sigma}} dy \leq C.$$

Adding them up, the corollary is concluded. □

**5. Further regularity.** Regularity  $C^{2,\alpha}$  can be reduced to Hölder regularity of the Laplacian. The same holds with respect to  $C^{\sigma+\alpha}$  regularity and the fractional Laplacian  $(-\Delta)^{\sigma/2}$ . Now, the fractional Laplacian,  $(-\Delta)^{\sigma/2}u$ , can be thought as a difference of an average of  $u$  with itself, which relates with the concavity of  $\mathcal{M}_{\mathcal{L}}^-$  in a proper way. We will exploit these two facts in this section to prove our  $C^{\sigma+\alpha}$  regularity result.

We keep the previous hypotheses for this section,  $\mathcal{L} \subseteq \mathcal{L}_2$  and  $u$  satisfy

$$u_t - \mathcal{M}_{\mathcal{L}}^- u = 0 \text{ in } C_{8,3},$$

$$\|u\|_{L^\infty((-3,0] \rightarrow L^1(\omega_\sigma))} + [u]_{C^{0,1}((-3,0] \rightarrow L^1(\omega_\sigma))} \leq 1.$$

In particular we know by now that, for  $K(y) \in [0, \Lambda]$ ,  $b \in B_\beta$ ,

$$\|L_{K,b}^\sigma u\|_{L^\infty(C_{1,1})} \leq C.$$

Given  $A \subseteq B_1$ , let

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

Fix  $\varphi \in C_0^\infty(B_1 \rightarrow [0, 1])$  such that  $\varphi = 1$  in  $B_{1/2}$ , and define

$$w_A(x, t) := \varphi(x) \int_{\mathbb{R}^n} (\delta u(x, t; y) - \delta u(0, t; y)) K_A^\sigma(y) dy.$$

By the properties deduced in the previous sections we have that  $w_A$  is globally bounded and satisfies

$$(w_A)_t - \mathcal{M}_{\mathcal{L}}^+ w_A \leq C,$$

in the cylinder  $C_{1,1}$ .

Let us also consider the extremal functions

$$P(x, t) := \sup_{A \subseteq B_1} w_A = (2 - \sigma) \varphi(x) \int_{B_{1/2}} \frac{(\delta u(x, t; y) - \delta u(0, t; y))^+}{|y|^{n+\alpha}} dy,$$

$$N(x, t) := \sup_{A \subseteq B_1} (-w_A) = (2 - \sigma) \varphi(x) \int_{B_{1/2}} \frac{(\delta u(x, t; y) - \delta u(0, t; y))^-}{|y|^{n+\alpha}} dy.$$

Our goal is to prove diminishing of the oscillation lemma for  $P + N$ . This implies that  $(-\Delta)^{\sigma/2} u$  is Hölder continuous, and therefore, the desired  $C^{\sigma+\alpha}$  regularity. We start by proving that  $P$  and  $N$  are comparable modulus a controlled error.

**Lemma 5.1.** *There exist universal constants  $C > 0$  and  $\alpha \in (0, 1)$  such that, for  $(x, t) \in C_{1/8, 1/2}$ , we have*

$$\frac{\lambda}{\Lambda} N - C|x|^\alpha \leq P \leq \frac{\Lambda}{\lambda} N + C|x|^\alpha.$$

*Proof.* For  $x \in B_{1/8}$ , let  $u_x(y, t) = u(x + y, t)$ . Since  $u$  solves  $u_t - \mathcal{M}_{\mathcal{L}}^- u = 0$  in  $C_{1,1}$ , then the difference  $(u_x - u)$  satisfies in  $C_{7/8, 1}$ ,

$$(u_x - u)_t - \mathcal{M}_{\mathcal{L}}^+(u_x - u) \leq 0,$$

$$(u_x - u)_t - \mathcal{M}_{\mathcal{L}}^-(u_x - u) \geq 0.$$

To recover  $P$  and  $N$  from the previous relations we consider for  $L = L_{K,0}^\sigma \in \mathcal{L}_2$ ,

$$L(u_x - u)(0) = \int_{\mathbb{R}^n} (\delta u(x, t; y) - \delta u(0, t; y)) K^\sigma(y) dy,$$

$$\begin{aligned} \lambda P(x, t) - \lambda N(x, t) &\leq \int_{B_1} (\delta u(x, t; y) - \delta u(0, t; y)) K^\sigma(y) dy \\ &\leq \Lambda P(x, t) - \lambda N(x, t). \end{aligned}$$

Changing variables, we can rewrite

$$\begin{aligned} &\int_{B_1^c} (\delta u(x, t; y) - \delta u(0, t; y)) K^\sigma(y) dy \\ &= \int_{\mathbb{R}^n} u(y, t) (K^\sigma(y - x) \chi_{B_1^c}(y - x) - K^\sigma(y) \chi_{B_1^c}(y)) dy \\ &\quad + (u(x, t) - u(0, t)) \int_{B_1^c} K^\sigma(y) dy. \end{aligned}$$

By Theorem 2.9, the last term is of order  $|x|$ . The first term can be estimated using the smoothness hypothesis of  $K$ ,

$$\begin{aligned} &\int_{\mathbb{R}^n} |K^\sigma(y - x) \chi_{B_1^c}(y - x) - K^\sigma(y) \chi_{B_1^c}(y)| dy \\ &\leq \int_{B_{1/2}^c} |K^\sigma(y - x) - K^\sigma(y)| dz \leq C|x|. \end{aligned}$$

On the other hand, we have the estimate  $\|(u_x - u)_t\|_\infty \leq C|x|^\alpha$  from Theorem 2.10. Therefore,

$$[(u_x - u)_t - L(u_x - u)](0, t) \geq -C|x|^\alpha - \Lambda P(x) + \lambda N(x).$$

Taking the infimum over  $L \in \mathcal{L}_2$  and using the equation for  $(u_x - u)$ , we get

$$0 \geq -C|x|^\alpha - \Lambda P(x, t) + \lambda N(x, t).$$

A similar computation with  $(u_x - u)_t - \mathcal{M}_{\mathcal{L}}^-(u_x - u) \geq 0$  provides the other inequality. □

The next result is a diminishing of the oscillation lemma. As we have learned from [3, 4, 5, 8, 9] it is important to strengthen the hypothesis of being just bounded and allow some growth at infinity. This allows iteration of the lemma by noting that the tails grow in a controlled way.

By rescaling, we can further assume that, for  $\varepsilon_1 > 0$  sufficiently small (to be fixed) and for every set  $K \subseteq \mathbb{R}^n$ , we have

$$(5.1) \quad |w_K| \leq 1/2 \quad \text{in } C_{1,1},$$

$$(5.2) \quad |w_K| \leq |x|^{1/2} \quad \text{in } B_1^c \times [-1, 0],$$

$$(5.3) \quad (w_K)_t - \mathcal{M}_{\mathcal{L}}^+ w_K \leq \varepsilon_1 \quad \text{in } C_{1,1}.$$

Additionally, by the previous lemma, we can assume that in  $C_{1/2,1}$

$$(5.4) \quad \frac{\lambda}{\Lambda} N(x, y) - \varepsilon_1 |x|^\alpha \leq P(x, t) \leq \frac{\Lambda}{\lambda} N(x, t) + \varepsilon_1 |x|^\alpha.$$

**Lemma 5.2.** *Assume (5.1), (5.2), (5.3) and (5.4). There are constants  $\kappa, \theta > 0$ , sufficiently small such that in  $C_{\kappa, \kappa^\sigma}$*

$$P \leq \frac{1}{2} - \theta.$$

**Remark 5.3.** We should ask ourselves how small  $\kappa$  and  $\theta$  should be in order to be able to iterate the lemma. We need the rescaled function  $\tilde{w}_K$ , given by

$$\tilde{w}_K(x, t) = \frac{w_K(\kappa x, \kappa^\sigma t)}{1 - \theta},$$

to satisfy the same hypotheses (5.1), (5.2), (5.3) and (5.4). Now, (5.1) is immediate and (5.2) holds if  $(1 - \theta) - \kappa^{1/2} \geq \theta/2 > 0$ , which is reasonable as  $\kappa, \theta$  can be chosen even smaller. Also, (5.3) holds if  $(1 - \theta) > \kappa^\sigma$  which was already contained in the previous inequality as  $\sigma > 1/2$ . Finally, (5.4) holds if  $\kappa^{\sigma-\alpha} \leq (1 - \theta)$  which is possible because  $\sigma > 1 > \alpha$ .

*Proof.* Assume by contradiction that, for some  $(x_0, t_0) \in C_{\kappa, \kappa^\sigma}$ ,  $P(x_0, t_0) > (1/2 - \theta)$ . There is then some set  $A$  such that  $w_A(x_0, t_0) > (1/2 - \theta)$ . The function  $v_A$ , given by the following truncation,

$$v_A := \left( \frac{1}{2} - w_A \right)^+,$$

satisfies an equation coming from (5.3). The truncation introduces an error that can be controlled in the interior

$$(v_A)_t - \mathcal{M}_{\mathcal{L}}^- v_A \geq -C \quad \text{in } C_{1/2,1}.$$

We use the point estimate (Theorem 2.6) to control the distribution of  $v_A$  in  $C_{\kappa,\kappa^\sigma}(0, -\kappa^\sigma)$ ,

$$(5.5) \quad \frac{|\{v_A > s\theta\} \cap C_{\kappa,\kappa^\sigma}(0, -\kappa^\sigma)|}{|C_{\kappa,\kappa^\sigma}(0, -\kappa^\sigma)|} \leq C(\theta + \kappa^\sigma)^\varepsilon (s\theta)^{-\varepsilon}.$$

Take  $\kappa^\sigma \leq \theta$ ; we can make the right hand side  $Cs^{-\varepsilon}$  sufficiently small, independently of  $\theta$ , by taking  $s$  sufficiently large. This makes the set  $G$ , defined by

$$G := \{w_A \geq (1/2 - s\theta)\} \cap C_{\kappa,\kappa^\sigma}(0, -\kappa^\sigma),$$

cover a fraction of  $C_{\kappa,\kappa^\sigma}(0, -\kappa^\sigma)$  close to one.

In  $G$ ,  $w_A$  and  $P$  are close to  $1/2$ . Thanks to (5.4),  $N$  can also be forced to be strictly positive in  $G$  by making  $\varepsilon_1 + \theta \leq 1/4$ . Therefore, we can assume  $N$  is larger than  $\lambda/(4\Lambda)$  in  $G$ . Also, in  $G$  and for  $B = B_1 \setminus A$ ,  $w_B$  has to be close to  $-N$ . This follows from the relation  $w_A + w_B = P - N$ ; therefore,

$$0 \leq N + w_B = P - w_A \leq s\theta.$$

This allows us to make  $w_B \leq -\lambda/(8\Lambda)$  in  $G$  by choosing  $\theta < \lambda/(8s\Lambda)$ .

Now we use the oscillation lemma to obtain the contradiction. Consider, for  $\eta \in (0, 1)$ , the function  $v_B$  given by

$$v_B(x, t) = \left( w_B(\kappa\eta x, (\kappa\eta)^\sigma t - \kappa^\sigma) + \frac{\lambda}{8\Lambda} \right)^+.$$

It still satisfies in  $C_{(\kappa\eta)^{-1}, (\kappa\eta)^{-\sigma}}$ ,

$$(v_B)_t - \mathcal{M}_{\mathcal{L}}^+(v_B) \leq \varepsilon_1(\kappa\eta)^\sigma \leq \varepsilon_1.$$

Also, from (5.5), we know that

$$|\{v_B > 0\} \cap C_{\eta^{-1}, \eta^{-\sigma}}| \leq C\eta^{-(n+\sigma)} s^{-\varepsilon}.$$

By the oscillation lemma,

$$\frac{\lambda}{8\Lambda} = v_B(0, 0) \leq C \left( \varepsilon_1 + \eta^{-(n+\sigma)} s^{-\varepsilon} + \sup_{t \in [-\eta^{-\sigma}, 0]} \int_{B_{\eta^{-1}}^{C_{\eta^{-1}, \eta^{-\sigma}}}} \frac{|v_B(y, t)|}{|y|^{n+\sigma}} dy \right).$$

Changing variables,

$$\int_{B_{\eta^{-1}}^c} \frac{|v_B(y, t)|}{|y|^{n+\sigma}} dy = (\kappa\eta)^\sigma \int_{B_\kappa^c} \frac{(w_B(y, (\kappa\eta)^\sigma t - \kappa^\sigma) + (\lambda/8\Lambda))^+}{|y|^{n+\sigma}} dy, \leq C\eta^\sigma,$$

where the last inequality holds by the bounds (5.1) and (5.2). Putting it back into the estimate, we obtain

$$\frac{\lambda}{8\Lambda} \leq C \left( \varepsilon_1 + \eta^{-(n+\sigma)} s^{-\varepsilon} + \eta^\sigma \right).$$

This gives us a contradiction by choosing  $\varepsilon_1, \eta^\sigma < \lambda/(100C\Lambda)$ , and then  $s^\varepsilon > (100C\Lambda)/(\lambda\eta^{n+\sigma})$ . □

We are now able to prove the parabolic nonlocal Evans-Krylov theorem.

**Theorem 5.4** (Classical solutions). *Let  $\mathcal{L} \subseteq \mathcal{L}_2$ ,  $u$  be a bounded function in  $\mathbb{R}^n \times (-1, 0]$  solving*

$$u_t - \mathcal{M}_{\mathcal{L}}^- u = 0 \quad \text{in viscosity in } C_{1,1}.$$

*Then,  $(-\Delta)^\sigma u$  is Hölder continuous with the following estimate*

$$\|(-\Delta)^\sigma u\|_{C^\alpha(C_{1/2,1/2})} \leq C(\|u\|_{L^\infty((-1,0] \rightarrow L^1(\omega_\sigma))} + [u]_{C^{0,1}((-1,0] \rightarrow L^1(\omega_\sigma))}).$$

*Proof.* The case  $\sigma \leq 1$  is contained in [10]. By the regularization procedure of [9] we can assume that  $(-\Delta)^\sigma u$  is continuous, and by translation invariance, all we need to show is that the estimate holds at the origin. As usual, we renormalize  $u$  in order to have

$$\|u\|_{L^\infty((-1,0] \rightarrow L^1(\omega_\sigma))} + [u]_{C^{0,1}((-1,0] \rightarrow L^1(\omega_\sigma))} \leq 1.$$

By the definitions of  $N$  and  $P$ , we have the following identity in  $B_{1/8} \times (-1, 0]$

$$\begin{aligned} & (-\Delta)^\sigma u(0, t) - (-\Delta)^\sigma u(x, t) \\ &= C \left( P(x, t) + N(x, t) + (2 - \sigma) \int_{B_1^c} \frac{\delta u(x, t; y) - \delta u(0, t; y)}{|y|^{n+\sigma}} dy \right). \end{aligned}$$

The third term can be bounded by  $C|x|$  as in the proof of Lemma 5.1.

Lemma 5.2 and Remark 5.3 give a geometric decay for  $P$  around the origin which implies a Hölder modulus of continuity for it. By Lemma 5.1, this is equivalent to a similar modulus of continuity for  $N$ . Then, the first two terms above can be bounded by  $C|x|^\alpha$ , for some universal  $\alpha$ , which concludes the proof.  $\square$

## REFERENCES

1. H. Abels and M. Kassmann, *An analytic approach to purely nonlocal Bellman equations arising in models of stochastic control*, J. Diff. Equat. **236** (2007), 29–56.
2. L. Caffarelli and X. Cabré, *Fully nonlinear elliptic equations*, Amer. Math. Soc. Colloq. Publ. **43**, American Mathematical Society, Providence, RI, 1995.
3. L. Caffarelli and L. Silvestre, *Regularity theory for fully nonlinear integro differential equations*, Comm. Pure Appl. Math. **62** (2009), 597–638.
4. ———, *Regularity results for nonlocal equations by approximation*, Arch. Rat. Mech. Anal. **1** (2011), 59–88.
5. ———, *The Evans-Krylov theorem for non local fully non linear equations*, Ann. Math. **174** (2011), 1163—1187.
6. ———, *On the Evans-Krylov theorem*, Proc. Amer. Math. Soc. **138** (2010), 263–265.
7. Héctor Chang-Lara, *Regularity for fully non linear equations with non symmetric kernels up to some critical cases*, Math. Anal. PDEs **2012**, arXiv:1210.4242v2[math.AP].
8. Héctor Chang-Lara and Gonzalo Dávila, *Regularity for solutions of non local parabolic equations*, Calc. Var. Partial Diff. Equat. **49** (2014), 139–172
9. ———, *Regularity for solutions of non local parabolic equations*, J. Diff. Equat. **256** (2014), 130–156.
10. ———, *Hölder estimates for non-local parabolic equations with critical drift*, J. Diff. Equat. **260** (2016), 4237–4284.
11. S.D. Deskmukh and S.R. Pliska, *Optimal consumption and exploration of non-renewable resources under uncertainty*, Econometrica **48** (1980), 177–200.
12. Lawrence C. Evans, *Classical solutions of fully nonlinear, convex, second-order elliptic equations*, Comm. Pure Appl. Math. **35** (1982), 333–363.
13. Matthieu Felsinger and Moritz Kassmann, *Local regularity for parabolic nonlocal operators*, Comm. Part. Diff. Equat. **38** (2013), 1539–1573.
14. Tianling Jin and Jingang Xiong, *Schauder estimates for nonlocal fully nonlinear equations*, Ann. Inst. Henri Poincaré Non Lin. Anal. (2015), 0294–1449.
15. N.V. Krylov, *Boundedly inhomogeneous elliptic and parabolic equations*, Izv. Akad. Nauk **46** (1982), 487–523, 670.
16. J.-L. Menaldi and M. Robin, *Ergodic control of reflected diffusions with jumps*, Appl. Math. Optim. **35** (1997), 117–137.

17. Joaquim Serra, *Regularity for fully nonlinear nonlocal parabolic equations with rough kernels*, arXiv:1401.4521v3, 2014.

18. ———,  *$C^{\sigma+\alpha}$  regularity for concave nonlocal fully nonlinear elliptic equations with rough kernels*, Calc. Var. Part. Diff. Equat. **54**, (2015), 3571–3601.

19. Luis Silvestre, *On the differentiability of the solution to the Hamilton-Jacobi equation with critical fractional diffusion*, Adv. Math. **226** (2011), 2020–2039.

20. Halil Mete Soner, *Optimal control with state-space constraint*, II, SIAM J. Contr. Optim. **24** (1986), 1110–1122.

21. Elias M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Math. **30**, Princeton University Press, Princeton, NJ, 1970.

COLUMBIA UNIVERSITY, DEPARTMENT OF MATHEMATICS, NEW YORK, NY 10027

**Email address:** [changelara@math.columbia.edu](mailto:changelara@math.columbia.edu)

DEPARTAMENTO DE MATEMÁTICA, UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA, AVENIDA ESPAÑA 1680, VALPARAÍSO, CHILE

**Email address:** [gonzalo.davila@usm.cl](mailto:gonzalo.davila@usm.cl)