# Regularity estimates for parabolic integro-differential equations and applications

Luis Silvestre

**Abstract.** We review some regularity results for integro-differential equations, focusing on Hölder estimates for equations with rough kernels and their applications. We show that if we take advantage of the integral form of the equation, we can obtain simpler proofs than for second order equations. For the equations considered here, the Harnack inequality may not hold.

**Mathematics Subject Classification (2010).** Primary 45E99; Secondary 35K99.

**Keywords.** Nonlocal equations. Regularity.

## 1. Introduction

We consider parabolic integro-differential equations of the following general form

$$u_t + b(t,x) \cdot \nabla u + \int_{\mathbb{R}^d} (u(x+y) - u(x) - y \cdot \nabla u(x) \ \chi_{B_1}(y)) K(t,x,y) \, \mathrm{d}y = f. \quad (1)$$

The equations corresponds to generators of discontinuous Levy processes with drift, but without a diffusion part. The purpose of this article is to discuss the regularization effect of the integral part of the equation under general conditions on $b$ and $K$. We present a Hölder regularity result for the solution $u$ to the equation (1) for $f$ bounded and under some conditions on $b$ and $K$ which are discussed in section 3. It is important that there is no smoothness assumption on the kernel values $K(t,x,y)$ with respect to the variables $x$ and $t$.

Integro-differential equations are a natural, fractional order, generalization of classical parabolic equations. Indeed, second order parabolic equations arise as asymptotic limits of integral equations. In this respect, we can classify regularity estimates for integral equations in two types. A regularity estimate that is uniform in the order of the equation (provided only that it is bounded away from zero) can be passed to the limit to local equations and is thus a generalization of regularity estimates for second order parabolic equations (for example the results in [9], [10], [26] and [34]). The term *robust* is sometimes used for this type of estimates. In other types of regularity results, the estimates depend on the order of the equation and blow up as it approaches two. These estimates are the ones which take advantage of the non local integral structure of the equation. The loss of the estimates in the classical limit is certainly an undesirable quality. However, it is interesting to understand how the non local structure of the equation can be used to our advantage. In fact, this second type of estimates usually have simpler proofs even than their classical local counterparts.

In this article, we discuss several results in the area. We also show a proof of Hölder estimates for solutions to a rather general form of the equation (1). The proof we show gives an estimate which is **not** robust. Our methods here depend strongly on the integral structure of the equation. In exchange, we can analyze a large class of kernels $K$ and still have a rather short and clean proof. The Hölder estimates are based on a version of the weak Harnack inequality (Theorem 5.1) which has interest in itself and is not true for second order parabolic equations in non divergence form.

The Hölder estimates are closely related with the Harnack inequality. For the equations we consider in this paper we prove the Hölder estimates but we show that the Harnack inequality does not hold.

Note that these estimates are a fractional order version of the classical theory for second order parabolic equations due to Krylov and Safonov [32]. In comparison, the proofs for second order equations are much more complicated than then proofs of the non-robust estimates for integral equations. Also, the Harnack inequality certainly holds in the second order case.

We review related results and a brief history of the subject in section 2. In the last section we give a quick summary of some of the main applications of the regularity estimates.

## 2. A review on regularity results

**2.1. Classical results for second order equations.** Integro-differential equations are a natural extension of second order equations of elliptic and parabolic type. There are two types of regularity results for second order equations, those which apply to equations in divergence form and those for equations in non divergence form.

The first result for elliptic and parabolic equations with rough coefficients and without any smallness condition is the classical result of De Giorgi, Nash and Moser which was obtained in the late 1950's. This result provides a Hölder estimate for equations of the form

$$u_t - \partial_i a_{ij}(x,t)\partial_j u \partial_j u = 0.$$

where the coefficients $a_{ij}$ are assumed to satisfy the point-wise bounds

$$\lambda I \leq \{a_{ij}\} \leq \Lambda I. \tag{2}$$

No assumption is made in terms of the regularity of $a_{ij}$ with respect to either $x$ or $t$. The Harnack inequality also holds for this type of equations.

This classical result plays a crucial role in the regularity theory of solutions to equations in divergence form. It was the key to solve Hilbert's 19th problem. It is essential for all its applications that no regularity assumption on the coefficients is necessary.

The corresponding result for equations in non divergence form was obtained in 1979 by Krylov and Safonov [33], [32]. It applies to equations of the form

$$u_t - a_{ij}(x,t)\partial_{ij}u = 0,$$

with identical assumptions on the coefficients $a_{ij}$. Before the result of Krylov and Safonov, the only regularity results available for equations in non divergence form applied to either continuous coefficients $a_{ij}$ or coefficients with small oscillation (i.e. $|a_{ij} - \delta_{ij}| < \varepsilon$). Note that any of these extra assumptions would allow us to approximate the equation locally with an equation with constant coefficients. The result of Krylov and Safonov is more delicate because it deals with a different scale invariant class of equations.

The Hölder estimate and Harnack inequality by Krylov and Safonov have multiple applications. It is a central result in the study of regularity of solutions to fully non linear elliptic equations. These are equations of the form $F(D^2u) = 0$ where $F$ is an arbitrary nonlinear function which satisfies $\lambda I \leq \partial F/\partial X_{ij} \leq \Lambda I$. The canonical examples of equations of this form come from the study of stochastic games and are the Bellman equation

$$u_t - \sup_r a^r_{ij}\partial_{ij}u = 0,$$

or the Isaacs equation

$$u_t - \inf_s \sup_r a^{rs}_{ij}\partial_{ij}u = 0.$$

Note that even tough the matrices $a^r_{ij}$ or $a^{rs}_{ij}$ in these equations may be independent of $x$ and $t$, a different one is chosen at every point $(x,t)$ and we have no a priori estimate on the optimal choice $a_{ij}(x,t)$ other than a quantitative point estimate like (2). That's why a result like the Hölder estimates by Krylov and Safonov guarantees some initial regularity for solutions to this kind of problems. More importantly, the derivative of the solution also satisfies an equation with (a priori) rough coefficients, which we can use to deduce that the solutions are $C^{1,\alpha}$ both in space and time. This is the best regularity currently known for the Isaacs equation and in fact it is known to be optimal in high dimensions [37], [38]. The solutions to the Isaacs equation in 2D are always $C^{2,\alpha}$ in space. It is still an outstanding open problem whether singular solutions exists in dimensions 3 or 4. For the Bellman equation, the solutions are always $C^{2,\alpha}$ in space, and therefore classical, due to the celebrated theorem proved independently by Evans [22] and Krylov [31].

**2.2. Integro-differential equations.** Just like in the second order case, the first regularity results that appear for integro-differential equations were for equations with a variational structure. They are a fractional order version of the classical results by De Giorgi Nash and Moser for second order equations in divergence form. See [30], [1], [5], [26], [23] and [8] among others. In this article we will concentrate in results of non variational form. There is a good survey on Hölder estimates for divergence form integro-differential equations in [28].

The first Hölder estimate, together with a Harnack inequality, was obtained by Bass and Levin [4] for elliptic integro-differential equations of the form

$$\int_{\mathbb{R}^d} (u(x+y) - u(x) - y \cdot \nabla u(x) \chi_{B_1}(y))\, K(x,y)\, \mathrm{d}y = 0,$$

assuming that $K(x,y) = K(x,-y)$ and

$$\frac{\lambda}{|y|^{d+\alpha}} \leq K(x,y) \leq \frac{\Lambda}{|y|^{d+\alpha}}, \tag{3}$$

for positive constants $\lambda$ and $\Lambda$ uniform in $x$ and $y$.

Note that there is no regularity assumption of $K$ with respect to $x$. The assumption (3) is a uniform ellipticity condition of order $\alpha \in (0,2)$ comparable to (2) for the second order case.

The result in [4] is obtained using probabilistic techniques and the estimates are not robust since the constants in the Hölder estimate are not uniformly bounded as $\alpha \to 2$.

Just as in the case of Krylov-Safonov theorem for second order equations, this result applies to a scale invariant class of equations. The assumption (2) which gives us a pointwise bound for the coefficients $a_{ij}$ in the second order case can be reproduced in the integro-differential case by the assumption (3), but there are also many other alternatives. For each point $x$, the possible kernels are non negative functions (in terms of $y$) from the full space $\mathbb{R}^d$ into $\mathbb{R}$. This is naturally a much richer class than the coefficients of a second order equation, which is just a symmetric matrix in $\mathbb{R}^{d \times d}$.

The original result of Bass and Levin was extended to more general classes of equations in [45], [3], [2] and [41]. The last one was the first one to use purely analytic methods instead of probabilistic techniques.

The first robust Hölder estimate and Harnack inequality appear in [9] for symmetric kernels satisfying the same condition (3). In this paper, the Hölder estimates were used to derive a $C^{1,\alpha}$ estimate for the non local Isaacs equation from stochastic games driven by Levy processes,

$$\inf_r \sup_s \int_{\mathbb{R}^d} (u(x+y) - u(x) - y \cdot \nabla u(x) \chi_{B_1}(y))\, K^{rs}(y)\, \mathrm{d}y = 0.$$

This result was extended to some Isaacs equations with variable coefficients in [11]. Moreover, the integro-differential Bellman equation

$$\inf_r \int_{\mathbb{R}^d} (u(x+y) - u(x) - y \cdot \nabla u(x) \chi_{B_1}(y))\, K^{r}(y)\, \mathrm{d}y = 0,$$

has classical solutions in the class $C^{\alpha+\varepsilon}$ if all kernels $K$ satisfy (3) plus some smoothness condition with respect to $y$ [10].

There is a number of variations of the results mentioned above, including, among others, generalizations to parabolic equations [35], to non symmetric kernels [13] and to kernels with logarithmic singularities [27].

As mentioned above, the condition (3) is a version of uniform ellipticity of fractional order. However, given that the possible choices of kernel measures are so rich, there are several possible alternative definitions of uniform ellipticity. It is not clear at the moment what the optimal condition should be, and there is probably never going to be one.

In [6] and [39], the lower bound condition in (3) was relaxed. The authors observed that it is enough to let it hold only in a cone of directions, which can change from point to point, but must have a uniform width. This is a remarkable improvement which shows that the previous condition (3) was far from optimal. It turns out that for this kind of integro-differential equations the Hölder estimates hold but the Harnack inequality does not. This is also quite remarkable, since the two properties are closely related to each other, and the Hölder estimates are often proved as a consequence of the Harnack inequality. Note that actually the Harnack inequality is claimed to be true in [6], but there are some issues in the proof and a counterexample was given in [39] using ideas from [7].

In this article we prove a version of the Hölder estimate for parabolic integro-differential equations. We keep the lower bound of (3) but we relax the upper bound. For this type of equations we show that the Harnack inequality does not hold either. We give a rather simple proof at the expense of making the estimate not robust (that is, the constants blow up as $\alpha \to 2$).

In a work in progress with Russell Schwab, we are working on a robust estimate for parabolic equations whose kernels satisfy the same upper bound condition as in this paper, but whose lower bound only holds in a cone of directions.

Note that a robust estimate like the one in [9] in particular implies the Krylov Safonov theorem about the regularity of uniformly elliptic second order equations. We can recognize the main ideas of the proof of Krylov and Safonov in the proof of the corresponding estimates for integral equations in [9]. In particular, there is some replacement for the Alexandrov-Bakelman-Pucci estimate. However, a perfect nonlocal analog for the ABP estimate is unknown. More precisely, the following is an open problem. Assume

$$\int_{\mathbb{R}^d} (u(x+y) - u(x) - y \cdot \nabla u(x) \chi_{B_1}(y))\, K(x,y) \,\mathrm{d}y = -\chi_A \qquad \text{in } B_1,$$

$$u(x) \leq 0 \qquad \text{for all } x \notin B_1$$

Is there an estimate of the form

$$\max_{B_1} u \leq \mu(|A|),$$

for any function $\mu$ so that $\mu(m) \to 0$ as $m \to 0$?

Here $\chi_A$ denotes the characteristic function of the set $A$.

In spite of its apparent simplicity, the estimate above has only been proved for a very special class of kernels [24].

## 3. A class of kernels

In this section, we describe a special set of assumptions that we make on the kernels for the result that we prove in this article.

The integro-differential equations that we study have an associated order $\alpha \in (0,2)$. The assumptions on the kernel $K$ depend on this value. Typically, we will look at kernels $K$ such that $K(t,x,y) \approx |y|^{-d-\alpha}$, but our assumptions are more general than that. We assume the following two inequalities.

$$K(t,x,y) \geq \frac{\lambda}{|y|^{d+\alpha}}, \tag{4}$$

$$\int_{\partial B_r} K(t,x,y) \, \mathrm{d}S(y) \leq \Lambda r^{-1-\alpha} \text{ for all } r > 0. \tag{5}$$

The estimates (4) and (5) also make sense for some singular jump measures instead of the absolutely continuous measure $K(t,x,y)\,\mathrm{d}y$. We stick to the absolutely continuous form only for the sake of clarity.

Note that (5) is more general than the usual assumption $K(t,x,y) \leq \Lambda |y|^{-d-\alpha}$. In particular (5) allows us to consider kernels containing singular measures. An extra factor $(2-\alpha)$ would be needed in the right hand side of both (4) and (5) in order to pass to the limit as $\alpha \to 2$ and recover uniformly parabolic equations in non divergence form. Since the estimates in this article are not uniform in $\alpha$, it makes no difference to have this factor or not for the purposes of our results.

We do not assume that $K$ is symmetric, (i.e. $K(t,x,y) \neq K(t,x,-y)$. Note that the purpose of the gradient term $y \cdot \nabla u(x) \chi_{B_1}(x)$ in the integrand in (1) is for the integral to be well defined around the origin. When $K$ is symmetric, this term can be safely ignored by computing the integral in the principal value sense. For non symmetric kernels $K$, this term is necessary for the integral to make sense if $\alpha \geq 1$. The choice of the radius of the ball $B_1$ is arbitrary. If we replace $B_1$ with $B_R$ for any other value of $R$, the difference of the integral operators would be absorbed by the gradient term. This ambiguity in the structure of the integral operator is inconvenient for the proofs in this article because it depends on scale. In our proofs we often rescale solutions of the equation and we need our assumptions to be invariant by this scaling. In order to avoid this ambiguity, we modify the structure of the equation depending on whether $\alpha < 1$ or $\alpha > 1$. We use the following notation

$$\delta_y u(x) := \begin{cases} u(x+y) - u(x) & \text{if } \alpha < (0,1), \\ u(x+y) - u(x) - y \cdot \nabla u(x) \chi_{B_1}(y) & \text{if } \alpha = 1, \\ u(x+y) - u(x) - y \cdot \nabla u(x) & \text{if } \alpha \in (1,2), \end{cases} \tag{6}$$

In any case, we study an equation of the form

$$u_t + b(t,x) \cdot \nabla u + \int_{\mathbb{R}^d} \delta_y u(x) K(t,x,y) \, \mathrm{d}y = f \tag{7}$$

The case $\alpha = 1$ is special. In this case we need to assume an additional symmetry assumption in the kernel $K$. Assuming $K(t,x,y) = K(t,x,-y)$ would be enough, but we make a mildly more general assumption

$$\int_{\partial B_r} yK(t,x,y)\,\mathrm{d}S(y) = 0 \quad \text{for every } r > 0. \tag{8}$$

Some symmetry assumption like (8) is required to make the proof using the current methods. See [14] for a more precise description of assumptions that work in the elliptic case with $\alpha = 1$.

These alternative structures for the non symmetric integral operators depending on the order $\alpha$ were stated in the work of Hongjie Dong and Doyoon Kim for elliptic nonlocal equations [19] and [20].

The drift term does not contribute to the regularization of the solution. We need to be able to control it with the integral part. We will assume that $b(t,x)$ is bounded if $\alpha \geq 1$ and that $b(t,\cdot)$ is uniformly bounded in $C^{1-\alpha}$ if $\alpha < 1$. For the right hand side $f$, we always assume it is a bounded function.

Note that the class of equations of the form (1) is no different from (7). Indeed, if $\alpha < 1$ then $yK(y)$ is integrable at the origin. Thus, the integral of $y\cdot\nabla u(x)\chi_{B_1}(y)$ is of the form $\tilde{b}\cdot\nabla u(x)$ and therefore it is absorbed into the first order term of (7). In the case $\alpha > 1$, $yK(y)$ is integrable at infinity. Thus, the integral of $y\cdot\nabla u(x)\chi_{\mathbb{R}^d\setminus B_1}(y)$ can also be absorbed in the first order term.

## 4. Extremal operators and viscosity solutions

Note that the equation (1) does not have a variational form and therefore we cannot define a weak solution in the sense of distributions. In order to make sense of the equation in the viscosity sense, we define the extremal Pucci-like operators which represent the maximum and minimum possible value of the integral term in (7).

**Definition 4.1.** Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be bounded in $\mathbb{R}^d$ and $C^{1,1}$ at $x$. We define $M_\alpha^\pm\varphi$ in the following way.

$$M_\alpha^+\varphi(x) = \sup\left\{\int_{\mathbb{R}^d}\delta_y\varphi(x)K(y)\,\mathrm{d}y : \text{for all } K(y) \text{ satisfying (5) and (4),}\right.$$
$$\left.\text{and also (8) if } \alpha = 1\right\}$$

The extremal operator $M_\alpha^-$ is defined similarly exchanging the sup with an inf.

We will omit the subscript $\alpha$ whenever this value is clear to avoid notation clutter.

The operators $M_\alpha^+$ and $M_\alpha^-$ have an explicit formula which is given in the next lemma.

**Lemma 4.2.** *Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be bounded in $\mathbb{R}^d$ and $C^{1,1}$ at $x$. The operators $M_\alpha^+$ and $M_\alpha^-$ have the following formula*

$$M_\alpha^+ \varphi(x) = \int_{\mathbb{R}^d} \delta_y \varphi(x) \frac{\lambda}{|y|^{d+\alpha}} \,\mathrm{d}y + \omega_d(\Lambda - \lambda) \int_0^\infty \left( \sup_{y \in \partial B_r} \delta_y \varphi(x)^+ \right) r^{-1-\alpha} \,\mathrm{d}r,$$

$$M_\alpha^- \varphi(x) = \int_{\mathbb{R}^d} \delta_y \varphi(x) \frac{\lambda}{|y|^{d+\alpha}} \,\mathrm{d}y + \omega_d(\lambda - \Lambda) \int_0^\infty \left( \sup_{y \in \partial B_r} \delta_y \varphi(x)^- \right) r^{-1-\alpha} \,\mathrm{d}r,$$

*where $\omega_d$ represents the surface area of $\partial B_1$ in $\mathbb{R}^d$.*

The proof of the lemma above is a straight forward interpretation of Definition 4.1

The following is a rather simple proposition stating essentially that $M_\alpha^\pm \varphi$ is well defined for $\varphi \in C^2$ and has some basic stability properties.

**Proposition 4.3.** *Assume $\varphi$ is a continuous bounded function in $\mathbb{R}^d$ so that $\varphi \in C^2(B_2)$. Then $M_\alpha^+ \varphi$ and $M_\alpha^- \varphi$ are continuous in $B_1$. Moreover, if $\varphi_k$ is a uniformly bounded sequence of such functions such that $\varphi_k \to \varphi$ locally uniformly in $\mathbb{R}^d$ and $\varphi_k \to \varphi$ in $C^2(B_2)$, then $M_\alpha^\pm \varphi_k \to M_\alpha^\pm \varphi$ uniformly in $B_1$.*

*Proof.* Let $K(y)$ be one kernel satisfying (4) and (5). Let $Lu$ be the linear (translation invariant) operator

$$Lu(x) := \int_{\mathbb{R}^d} \delta_y u(x) K(y) \,\mathrm{d}y. \tag{9}$$

If $x \in B_1$, then we can estimate $\delta_y u(x)$ with respect to the norms of $u$. If $x \in B_1$ and $x + y \in B_{3/2}$, we have

$$|\delta_y u(x)| \leq \begin{cases} |y| |\nabla u|_{L^\infty(B_2)} & \text{if } \alpha > 1, \\ |y|^2 |D^2 u|_{L^\infty(B_2)} \chi_{|y|<1} & \text{if } \alpha = 1 \text{ and } |y| < 1, \\ 2|u|_{L^\infty}(B_{3/2}) \chi_{|y|<1} & \text{if } \alpha = 1 \text{ and } |y| \geq 1, \\ |y|^2 |D^2 u|_{L^\infty(B_2)} & \text{if } \alpha > 1. \end{cases}$$

Also, if $x \in B_1$ and $x + y \notin B_{3/2}$,

$$|\delta_y u(x)| \leq \begin{cases} |u(x+y)| + |u|_{L^\infty(B_1)} & \text{if } \alpha > 1, \\ |u(x+y)| + |u|_{L^\infty(B_1)} + |y| |\nabla u|_{L^\infty(B_1)} \chi_{|y|<1} & \text{if } \alpha = 1 \text{ and } |y| < 1, \\ |u(x+y)| + |u|_{L^\infty(B_1)} & \text{if } \alpha = 1 \text{ and } |y| \geq 1, \\ |u(x+y)| + |u|_{L^\infty(B_1)} + |y| |\nabla u|_{L^\infty(B_1)} & \text{if } \alpha > 1. \end{cases}$$

All these inequalities, combined with the assumption (5) for $K(y)$ tell us that the expression in (9) is integrable and

$$|Lu(x)| \leq C \left( \|u\|_{C^2(B_2)} + \||x|^\beta u(x)\|_{L^\infty(\mathbb{R}^d)} \right). \tag{10}$$

Here $\beta$ is any non negative number so that $\beta < \alpha$. The constant $C$ depends on $\lambda$, $\Lambda$, $d$, $\alpha$ and $\beta$, but not on the particular choice of the kernel $K(y)$. In particular it also holds for the suppremum and infimum value of $Lu(x)$ for all admissible choices of $K$, and that is $M_\alpha^+ u(x)$ and $M_\alpha^- u(x)$.

From (10), we see that if $\varphi_n$ is a uniformly bounded sequence so that $\varphi_n \to \varphi$ in $C^2(B_2)$ and locally uniformly in $\mathbb{R}^d$, then $L\varphi_n \to L\varphi$ uniformly in $B_1$ and at a rate independent of $K$. In particular $M_\alpha^\pm \varphi_n$ converges to $M_\alpha^\pm \varphi$ uniformly in $B_1$, which proves the second part of the proposition.

Assume now that $u \in C^3(B_2) \cap C^1(\mathbb{R}^d)$, since $\partial_i[Lu] = L[\partial_i u]$, we deduce from (10) that in this case $Lu$ is Lipschitz continuous in $B_1$.

If $\varphi$ is any bounded continuous function in $\mathbb{R}^d$ which is $C^2$ in $B_2$, we can approximate it with a bounded sequence $\varphi_n \in C^3(B_2) \cap C^1(\mathbb{R}^d)$, which converges to $\varphi$ in $C^2(B_2)$ and locally uniformly in $\mathbb{R}^d$. Therefore, $L\varphi_n \to L\varphi$ uniformly in $B_1$, and therefore $L\varphi$ is continuous in $B_1$. □

Note that the fact that $u$ solves an equation of the form (7), for some kernel $K$ satisfying our assumptions, is equivalent to the fact that $u$ satisfies the following two inequalities.

$$u_t + b \cdot \nabla u - M_\alpha^+ u \leq 0, \tag{11}$$

$$u_t + b \cdot \nabla u - M_\alpha^- u \geq 0. \tag{12}$$

Even though we defined $M_\alpha^+$ and $M_\alpha^-$ using kernels which do not depend on $t$ and $x$, ultimately the choice of kernel in $M_\alpha^\pm u(t,x)$ is different at every point. Thus, the equations (11) and (12) are equivalent to (7) without any regularity assumption of $K(t,x,y)$ with respect to $t$ and $x$.

The advantage of the inequalities (11) and (12) with respect to the equation (7) is that the former can be defined in the viscosity sense.

**Definition 4.4.** Assume $b$ and $f$ are continuous in some domain $\Omega \subset \mathbb{R} \times \mathbb{R}^d$. We say that (11) (resp. (12)) holds in $\Omega$ if every time there is a function $\varphi$ : $(t_0 - \varepsilon, t_0] \times B_\varepsilon(x_0)$ such that $\varphi$ is left differentiable in time and point-wise second differentiable in space at $(t_0, x_0)$ and

$$\varphi(t_0, x_0) = u(t_0, x_0),$$
$$\varphi \geq u \text{ in } (t_0 - \varepsilon, t_0] \times B_\varepsilon(x_0) \quad \text{(resp. } \leq\text{)},$$

Then, if we construct the function

$$v(t,x) = \begin{cases} \varphi(t,x) & \text{if } (t,x) \in (t-\varepsilon, t_0] \times B_\varepsilon(x_0), \\ u(t,x) & \text{otherwise,} \end{cases}$$

we get

$$v_t(t_0,x_0) + b(t_0,x_0) \cdot \nabla v(t_0,x_0) - M_\alpha^+ v(t_0,x_0) \leq f(t_0,x_0).$$

(resp. $M_\alpha^-$ and $\geq$)

Note that $M^+$ and $M^-$ satisfy the relation $M^+\varphi = -M^-(-\varphi)$. Moreover, $u$ satisfies $u_t + b \cdot \nabla u - M^- u \geq f$ in the viscosity sense, if and only if $-u$ satisfies $(-u)_t + b \cdot \nabla(-u) - M^+(-u) \leq -f$.

One of the keys in the study of regularity properties of second order equations in non divergence form lies in the difficulty to obtain estimates in integral form. This is achieved estimating the measure of some contact sets or through the Alexandrov-Bakelman-Pucci inequality. The following lemma is a simple property of viscosity solutions of the integral equations we consider in this paper. At the same time, it is a crucial ingredient in our regularity theory since it provides a simple integral quantity associated with every point which can be touched by one side with a smooth function.

**Lemma 4.5.** *Assume* (4) *holds in the viscosity sense. Let $\varphi$ be a test function as in the Definition 4.4 such that $u(t_0, x_0 + y) \geq \varphi(t_0, x_0 + y)$ for all $y \in \mathbb{R}^d$. Then, we have*

$$\begin{aligned} &\varphi_t(t_0,x_0) + b(t_0,x_0)\cdot\nabla\varphi(t_0,x_0) - M^-\varphi(t_0,x_0) \\ &\qquad - \int_{\mathbb{R}^d} (u(t_0,x_0+y) - \varphi(t_0,x_0+y)) \frac{\lambda}{|y|^{d+\alpha}} \, \mathrm{d}y \geq f(t_0,x_0). \end{aligned}$$

*Proof.* Let $v$ be the function constructed out of $u$, $\varphi$ and some $\varepsilon > 0$ in Definition 4.4. Note that if $|y| > \varepsilon$ we have $\delta_y v(t_0,x_0) = (u(t_0,x_0+y) - \varphi(t_0,x_0+y)) + \delta_y\varphi(t_0,x_0)$. Moreover, $v_t(t_0,x_0) = \varphi_t(t_0,x_0)$ and $\nabla v(t_0,x_0) = \nabla\varphi(t_0,x_0)$. Therefore

$$\begin{aligned} f(t_0,x_0) &\leq v_t(t_0,x_0) + b(t_0,x_0)\cdot\nabla v(t_0,x_0) - M^+_\alpha v(t_0,x_0), \\ &= \varphi_t(t_0,x_0) + b(t_0,x_0)\cdot\nabla\varphi(t_0,x_0) - M^-\varphi(t_0,x_0) \\ &\qquad - \int_{\mathbb{R}^d\setminus B_\varepsilon} (u(t_0,x_0+y) - \varphi(t_0,x_0+y)) \frac{\lambda}{|y|^{d+\alpha}} \, \mathrm{d}y \end{aligned}$$

We finish the proof taking $\varepsilon \to 0$ and applying the monotone convergence theorem. □

**Remark 4.6.** Applying the previous result to $-u$, we can obtain a corresponding result for viscosity sub-solutions of (11).

## 5. The weak Harnack inequality

We use the following notation for parabolic cylinders

$$\begin{aligned} Q_r &:= (-r^\alpha, 0] \times B_r, \\ Q_r(t_0,x_0) &:= (t_0,x_0) + Q_r = (t_0 - r^\alpha, t_0] \times B_r(x_0). \end{aligned}$$

The following result is the weak Harnack inequality and is the key ingredient in the proof of the Hölder estimates. Its proof is inspired by a similar (but somewhat weaker) result for a particular case of (1) which appeared in [43]. Some of the ideas can be traced back to [41] for the elliptic case.

**Theorem 5.1.** *Let $u$ be a function which satisfies the following inequality in the viscosity sense.*

$$u_t + \Lambda|\nabla u| - M^- u \geq -a \text{ in } Q_1,$$

*Assume that $u \geq 0$ in $Q_1$, then*

$$\min_{[-1/2,0]\times B_{1/2}} u \geq c \left( \int_{(-1,-1/2]\times\mathbb{R}^d} \frac{u(t,x)}{(1+|x|)^{d+\alpha}} \,\mathrm{d}x\,\mathrm{d}t \right) - a,$$

*where $c$ is a positive constant depending on $\lambda$, $\Lambda$, $\alpha$ and the dimension $d$.*

*Proof.* Multiplying $u$ by a scalar, we can assume that

$$\int_{(-1,-1/2]\times\mathbb{R}^d} \frac{u(t,x)}{(1+|x|)^{d+\alpha}} \,\mathrm{d}x\,\mathrm{d}t = 1. \tag{13}$$

We should then prove that there is a small constant $c$ so that

$$\min_{[-1/2,0]\times B_{1/2}} u \geq c - a,$$

for any value of $a$. Of course this inequality is non trivial when $a < c$ only.

Let $\theta$ be the constant, depending on dimension and $\alpha$ only, so that

$$\int_{(-1,-1/2]\times\mathbb{R}^d} \frac{\theta}{(1+|x|)^{d+\alpha}} \,\mathrm{d}x\,\mathrm{d}t = 1/2.$$

From (13), we deduce that

$$\int_{(-1,-1/2]\times\mathbb{R}^d} \frac{(u(t,x)-\theta)^+}{(1+|x|)^{d+\alpha}} \,\mathrm{d}x\,\mathrm{d}t \geq 1/2.$$

Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be a smooth, non negative function supported in $B_{3/4}$ so that $\varphi \equiv 1$ in $B_{1/2}$. We will construct a bound from below of the form $m(t)\varphi(x) - (1+t)a$. The function $m$ is the solution of the following ODE, for some positive constants $c_0$ and $C_1$.

$$m(-1) = 0, \tag{14}$$

$$m'(t) = c_0 \left( \int_{\mathbb{R}^d} \frac{(u(t,x)-\theta)^+}{(1+|x|)^{d+\alpha}} \,\mathrm{d}x \right) - C_1 m(t). \tag{15}$$

The ODE above has the explicit solution

$$m(t) = c_0 \int_{-1}^{t} \int_{\mathbb{R}^d} \frac{(u(t,x)-\theta)^+}{(1+|x|)^{d+\alpha}} e^{C_1(s-t)} \,\mathrm{d}x\,\mathrm{d}s \geq \frac{c_0}{e^{C_1}} \iint_{[-1,-1/2]\times\mathbb{R}^d} \frac{(u(t,x)-\theta)^+}{(1+|x|)^{d+\alpha}} \,\mathrm{d}x\,\mathrm{d}t.$$

Therefore, if we proved that $u(t,x) \geq m(t)\varphi(x) - (1+t)a$, we would finish the proof with $c = c_0 e^{-C_1}/2$. Let us assume the contrary and let $\varepsilon > 0$ be an arbitrarily small number. Since $m(-1) = 0$ and $\varphi$ is supported in $B_{3/4}$, there exists a first crossing point $(t_0, x_0)$ so that

$$\begin{aligned} u(t_0,x_0) &= m(t_0)\varphi(x_0) - (1+t_0)a - \varepsilon, \\ u(t,x) &\geq m(t)\varphi(x) - (1+t)a - \varepsilon \text{ for every } t < t_0 \text{ and } x \in \mathbb{R}^d. \end{aligned}$$

We observe that we can use $m(t)\varphi(x) - (1+t)a - \varepsilon$ as a test function for Definition 4.4 and from Lemma 4.5 we have

$$\begin{aligned} m'(t_0)\varphi(x_0) - a + \Lambda m(t_0)|\nabla\varphi(x_0)| - m(t_0)M^-\varphi(t_0,x_0) \\ - \int_{\mathbb{R}^d} (u(t_0,x_0+y) - \varphi(t_0,x_0+y)) \frac{\lambda}{|y|^{d+\alpha}} \,\mathrm{d}y \geq -a. \end{aligned}$$

The $a$ terms cancel out and we get

$$\begin{aligned} m'(t_0)\varphi(x_0) + \Lambda m(t_0)|\nabla\varphi(x_0)| - m(t_0)M^-\varphi(t_0,x_0) \\ - \int_{\mathbb{R}^d} (u(t_0,x_0+y) - \varphi(t_0,x_0+y)) \frac{\lambda}{|y|^{d+\alpha}} \,\mathrm{d}y \geq 0. \end{aligned} \tag{16}$$

The last integral term is crucial to obtain the contradiction.

The first condition that we need in the choice of $c_0$ is that it must be chosen small enough so that $m(t) < \theta$ for all $t \in [-1, 0]$. This is guaranteed simply by the condition $c_0 \leq \theta$. Under this condition, we obtain the following

$$\begin{aligned} \int_{\mathbb{R}^d} (u(t_0,x_0+y) - \varphi(t_0,x_0+y)) \frac{\lambda}{|y|^{d+\alpha}} \,\mathrm{d}y &\geq \int_{\mathbb{R}^d} (u(t,y)-\theta)^+ \left( \min_{x_0 \in B_{1/2}} \frac{\lambda}{|y-x_0|^{d+\alpha}} \right) \mathrm{d}y \\ &\geq c_0 \left( \int_{\mathbb{R}^d} \frac{(u(t,x)-\theta)^+}{(1+|x|)^{d+\alpha}} \,\mathrm{d}x \right) \end{aligned}$$

The last inequality holds provided that $c_0 \leq \lambda(5/4)^{d+\alpha}$.

Substituting back into (16), we have

$$m'(t_0)\varphi(x_0) + \Lambda m(t_0)|\nabla\varphi(x_0)| - m(t_0)M^-\varphi(x_0) - c_0 \left( \int_{\mathbb{R}^d} \frac{(u(t,x)-\theta)^+}{(1+|x|)^{d+\alpha}} \,\mathrm{d}x \right) \geq 0.$$

Recalling the ODE (15) and that $\varphi \leq 1$, it follows that

$$-C_1 m(t_0)\varphi(x_0) + \Lambda m(t_0)|\nabla\varphi(x_0)| - m(t_0)M^-\varphi(x_0) \geq 0.$$

Since $u(t_0,x_0) = m(t_0)\varphi(x_0) - a(t+1) - \varepsilon$, then certainly $m(t_0) > 0$ and we can factor it out from the previous inequality.

$$-C_1\varphi(x_0) + \Lambda|\nabla\varphi(x_0)| - M^-\varphi(x_0) \geq 0.$$

We are left with choosing $C_1$ large enough in order to contradict this last inequality. Note that we can certainly do so in the set $\{x : \varphi(x) > \rho\}$ for any fixed $\rho > 0$. However, we must address the fact that $\varphi(x_0)$ might be arbitrarily small.

The key to solve this extra difficulty is to observe that $M^-\varphi(x) > 0$ wherever $\varphi(x) = 0$. Indeed $M^- f(x) > 0$ if $f$ achieves its global minimum at $x$. Moreover, also $\nabla\varphi = 0$ wherever $\varphi = 0$. Since $\varphi$ is $C^2$, then $M^-\varphi$ is continuous. Let $-\delta := \min\{M^-\varphi(x) : \varphi(x) = 0\}$. Then there is a $\rho > 0$ so that if $\varphi(x) < \rho$ then $|\nabla\varphi(x)| < \delta/(2\Lambda)$ and $M^-\varphi(x) < -\delta/2$. Therefore, every time $\varphi(x) < \rho$ we have

$$-C_1\varphi(x_0) + \Lambda|\nabla\varphi(x_0)| - M^-\varphi(x_0) \le -C_1\varphi(x_0) < 0.$$

For the points $x$ where $\varphi(x) \ge \rho$, we choose $C_1$ large enough so as to obtain a contradiction. This finishes the proof. □

Note that the lower bound provided in Theorem 5.1 involves a weighted integral of $u$ in the full space $\mathbb{R}^d$. This is obviously something that would not be expected for a local equation. The following is an immediate corollary of Theorem 5.1 in which we simply replace the integral in the right hand side by an integral in a subdomain. In this way, we obtain the weaker version of the result which is more similar to the weak Harnack inequality for local equations.

**Corollary 5.2.** *Let $u$ be a function which satisfies the following inequality in the viscosity sense.*

$$u_t + \Lambda|\nabla u| - M^-u \ge -a \text{ in } Q_1,$$

*Assume that $u \ge 0$ in $Q_1$, then*

$$\min_{[-1/2,0]\times B_{1/2}} u \ge c\left(\int_{(-1,-1/2]\times B_{1/2}} u(t,x)\,\mathrm{d}x\,\mathrm{d}t\right) - a,$$

*where $c$ is a positive constant depending on $\lambda$, $\Lambda$, $\alpha$ and the dimension $d$.*

Even the result of Corollary 5.2 is not true for second order equations. Instead, the integral on the right hand side must be replaced with the $L^\varepsilon$ norm of $u$ in $[-1,-1/2] \times B_{1/2}$ (See theorem 4.15 in [25]). It is relatively simple to construct stationary examples of the form $u(t,x) = |x|^{-p}$ for large $p$, to check that indeed a small power $\varepsilon$ is required in the second order case if $\Lambda/\lambda$ is large.

# 6. Hölder estimates

We first state the Hölder estimates in the case $\alpha \ge 1$. This case is relatively easier than $\alpha < 1$ because the diffusion is of higher order than the drift. The proof is a rather standard iterative improvement of oscillation. If the drift term vanishes, the same proof works for all $\alpha \in (0,2)$.

**Theorem 6.1.** *Assume $\alpha \ge 1$. Let $u : [-1,0] \times \mathbb{R}^d \to \mathbb{R}$ be a continuous bounded function which satisfies the following two inequalities in the viscosity sense*

$$u_t + \Lambda|\nabla u| - M^-u \ge -A \text{ in } Q_1,$$
$$u_t - \Lambda|\nabla u| - M^+u \le A \text{ in } Q_1.$$

*Then, $u \in C^\gamma(Q_{1/2})$ and there is an estimate*

$$\|u\|_{C^\gamma(Q_{1/2})} \leq C\left(\|u\|_{L^\infty([-1,0]\times\mathbb{R}^d} + A\right).$$

*Here the constant $C$ depends on $\lambda$, $\Lambda$, $\alpha$, and the dimension $d$.*

*Proof.* Replacing $u$ with $u/(\|u\|_{L^\infty} + A/\varepsilon_0)$, we can assume $\|u\|_{L^\infty} \leq 1/2$ and $A \leq \varepsilon_0$. We must now find a universal upper bound for the Hölder norm of $u$ in $Q_{1/2}$. We will prove that there for all $r > 0$,

$$\operatorname*{osc}_{Q_r} u \leq r^\gamma. \tag{17}$$

This shows that $u$ is $C^\gamma$ in space and $C^{\gamma/\alpha}$ in time at the point $(0,0)$. The regularity is extended to the cylinder $Q_{1/2}$ by a standard scaling and translation argument. Therefore, we only need to show (17).

Note that since we have $\|u\|_{L^\infty} \leq 1/2$, then (17) holds for all $r \geq 1$. We will show it holds for all $r > 0$ by induction in $r$. The inductive step we need to prove is that if (17) holds for all $r > r_0$, with $r_0 \leq 1$, then it also holds for all $r > r_0/2$.

So, assume (17) holds for all $r > r_0$. Let us consider the rescaled function

$$u_{r_0}(t,x) := (2R)^{-\gamma} u\left((2R)^\alpha t, 2Rx\right).$$

This is the scaling for which the values of $u_{r_0}$ in $Q_{1/2}$ correspond to the values of $u$ in $Q_{r_0}$. The function $u_{r_0}$ satisfies the equations

$$u_t + {r_0}^{\alpha-1}\Lambda|\nabla u| - M^- u \geq -{r_0}^{\alpha}\varepsilon_0 \text{ in } Q_{1/r_0}, \tag{18}$$
$$u_t - {r_0}^{\alpha-1}\Lambda|\nabla u| - M^+ u \leq {r_0}^{\alpha}\varepsilon_0 \text{ in } Q_{1/r_0}. \tag{19}$$

Since $r_0 < 1$ and $\alpha \geq 1$, both powers of $r_0$ in the formula above are less or equal to one. In particular

$$u_t + \Lambda|\nabla u| - M^- u \geq -\varepsilon_0 \text{ in } Q_{1/r_0},$$
$$u_t - \Lambda|\nabla u| - M^+ u \leq \varepsilon_0 \text{ in } Q_{1/r_0}.$$

Moreover, from the inductive hypothesis (17) for $r \geq r_0$, we have that

$$\operatorname*{osc}_{Q_r} u_{r_0} \leq r^\gamma \quad \text{for all } r \geq 1/2. \tag{20}$$

Let $m := \min_{Q_1} u_{r_0}$ and $M = \max_{Q_1} u_{r_0}$. From (20) we know that $M - m \leq 1$. Therefore, for every point $(t,x)$ in $Q_1$ we have at least one of the inequalities $u_{r_0}(t,x) \leq m + 1/2$ or $u(t,x) \geq M - 1/2$. Thus, one of the following two statements will hold

$$\left|\{u_{r_0} < m + 1/2\} \cap [-1,-1/2] \times B_{1/2}\right| > \frac{1}{4}|B_{1/2}| \quad \text{or} \tag{21}$$

$$\left|\{u_{r_0} > M - 1/2\} \cap [-1,-1/2] \times B_{1/2}\right| > \frac{1}{4}|B_{1/2}| \tag{22}$$

Without loss of generality, we assume (22) (otherwise, we will proceed with the rest of the proof with $-u_{r_0}$ instead of $u_{r_0}$).

Note that from (20), we deduce that $u_{r_0}(t,x) > M-1$ in $Q_1$ and $u_{r_0}(t,x) > M-|x|^\gamma$ for all $x \notin B_1$ and $t \in [-1,0]$.

Let $v(t,x)$ be the non negative function

$$v(t,x) = (u_{r_0}(t,x) - M + 2^\gamma)^+ .$$

Note that $v \geq u_{r_0}$ and for $t \in [-1,0]$ we have $v(t,x) - u(t,x) \leq (|x|^\gamma - 2^\gamma)^+$. In particular, $u(t,x) = v(t,x)$ if $t \in [-1,0]$ and $x \in B_2$.

Let $U(x) = (|x|^\gamma - 2^\gamma)^+$, so that $0 \leq v(t,x) - u(t,x) \leq U(x)$. The function $v$ satisfies the following equation (in the viscosity sense)

$$v_t + \Lambda |\nabla v| - M^- v \geq -\varepsilon_0 - M^- U \quad \text{in } Q_1.$$

Taking $\gamma$ small, we can make $M^- U$ arbitrarily small in $B_1$. Therefore, for small enough $\gamma$,

$$v_t + \Lambda |\nabla v| - M^- v \geq -2\varepsilon_0 \quad \text{in } Q_1.$$

We now apply Corollary 5.2 to $v$ and obtain a lower bound in $Q_{1/2}$,

$$\min_{Q_{1/2}} v \geq c \int_{[-1,-1/2]\times B_{1/2}} v \, \mathrm{d}x \, \mathrm{d}t - 2\varepsilon_0.$$

We now apply (22) and obtain

$$\min_{Q_{1/2}} v \geq c(2^\gamma - 1/2) - 2\varepsilon_0 > \delta.$$

This lower bound $\delta$ does not depend on $\gamma$ or $\varepsilon_0$ provided that $\varepsilon_0$ is sufficiently small.

Bringing this information back into $u_{r_0}$, this means that $u_{r_0} \geq M - 2^\gamma + \delta$ in $Q_{1/2}$. Here we also choose $\gamma$ sufficiently small so that $2^\gamma - \delta < 1 - \delta/2$. Therefore we have that $\operatorname{osc}_{Q_{1/2}} u_{r_0} \leq 1 - \delta/2$. This also means that

$$\operatorname*{osc}_{Q_{r_0/2}} u \leq (1-\delta/2) {r_0}^\gamma.$$

In particular $\operatorname{osc}_{Q_r} u \leq (1-\delta/2){r_0}^\gamma$ for all $r < r_0$. Choosing $\gamma$ sufficiently small one last time so that $2^{-\gamma} > 1 - \delta/2$, we proved that $\operatorname{osc}_{Q_r} u \leq r^\gamma$ for all $r \in [r_0/2, r_0]$.

This finishes the proof of (17) by induction in $r$. □

We now state and prove the corresponding theorem for $\alpha < 1$. In this case the interaction between the diffusion and drift is more subtle and we must make a change of variables partially following the flow in order to obtain the necessary cancellation to prove the theorem. This idea originated in [44].

**Theorem 6.2.** *Assume $\alpha < 1$. Let $u : [-1,0] \times \mathbb{R}^d \to \mathbb{R}$ be a continuous bounded function which satisfies the following two inequalities in the viscosity sense*

$$u_t + b(t,x) \cdot \nabla u - M^- u \geq -A \text{ in } Q_1,$$
$$u_t + b(t,x) \cdot \nabla u - M^+ u \leq A \text{ in } Q_1.$$

*Assume that $b$ is a continuous vector field and $\|b(t,\cdot)\|_{C^{1-\alpha}(B_1)}$ is bounded uniformly in $t$. Then, $u \in C^\gamma(Q_{1/2})$ and there is an estimate*

$$\|u\|_{C^\gamma(Q_{1/2})} \leq C\left(\|u\|_{L^\infty([-1,0]\times\mathbb{R}^d} + A\right).$$

*Here the constant $C$ depends on $\lambda$, $\Lambda$, $\alpha$, and the dimension $d$.*

*Proof.* The general strategy of the proof is similar to the proof of Theorem 6.1. The problem in the iterative argument is that in the rescaled equations (18) and (19), the factor ${r_0}^{\alpha-1}$ is large for small values of $r_0$. It is crucial that this factor remains bounded for the induction argument to succeed.

The solution is to change the shape of the parabolic cylinders we use, so that they follow the flow. Let us defined the following modified parabolic cylinders.

$$\tilde{Q}_r := \{(t,x) : |x - X(t)| < r \wedge t \in (-r^\alpha, 0]\}.$$

Here $X(t)$ is one solution to the backward ODE

$$\begin{aligned}X(0) &= 0,\\ X'(t) &= b(t, X(t)), \quad \text{for } t < 0.\end{aligned}$$

The corresponding scaled function $\tilde{u}_{r_0}$ is now

$$\tilde{u}_{r_0} = {r_0}^{-\gamma} u({r_0}^\alpha t, r_0(X(t) + x)).$$

This function solves the equations

$$\begin{aligned}u_t + {r_0}^{\alpha-1}\left(b(t,x) - b(t,X(t))\right)|\nabla u| - M^- u &\geq -{r_0}^\alpha \varepsilon_0 \text{ in } Q_{1/r_0},\\ u_t - {r_0}^{\alpha-1}\left(b(t,x) - b(t,X(t))\right)|\nabla u| - M^+ u &\leq {r_0}^\alpha \varepsilon_0 \text{ in } Q_{1/r_0}.\end{aligned}$$

The Hölder continuity assumption on $b$ assures that ${r_0}^{\alpha-1}|b(t,x) - b(t,X(t))| \leq \Lambda$ for some constant $\Lambda > 0$. This allows us to continue with the rest of the proof as in Theorem 6.1. □

# 7. Failure of the Harnack inequality

The Harnack inequality is a property of non negative solutions to some elliptic and parabolic equations. For parabolic equations, it says that there is some universal constant $C$ so that if $u$ is a solution of the equation in $Q_1$ which is non negative in $[-1,0]\times\mathbb{R}^d$, then

$$\sup_{[-3/4,-1/2]\times B_{1/2}} u \leq C \inf_{[-1/4,0]\times B_{1/2}} u.$$

The Harnack inequality holds for some integral equations, for example see [4], [9] and [1]. Interestingly enough, in the situation of [1], the Harnack inequality holds, but the solution of the equation is not necessarily continuous.

It turns out, however, that the Harnack inequality fails for the type of equations we consider in this paper. In this section, we construct a counterexample.

Let $e_1$ be a the unit vector $(1, 0, \dots)$ in $\mathbb{R}^d$. Consider the following integral operator

$$Lu(x) = \int_{\mathbb{R}^d} \frac{\delta_y u(x)}{|y|^{d+\alpha}} \, \mathrm{d}y + \int_{\mathbb{R}} \frac{\delta_{y_1 e_1} u(x)}{|y_1|^{1+\alpha}} \, \mathrm{d}y.$$

This operator is in fact the same as

$$Lu = -c_1(-\Delta)^{\alpha/2} u - c_2(-\partial_{x_1 x_1})^{\alpha/2} u.$$

for some positive constants $c_1$ and $c_2$. We look at the solution to the problem

$$\begin{aligned} u_t - Lu = f(x) \qquad & \text{in } (0, \infty) \times \mathbb{R}^d, \\ u(-1, x) = 0 \qquad & \text{for } x \in \mathbb{R}^d. \end{aligned} \tag{23}$$

where $f = \chi_{Q_\varepsilon}$ is the indicator function of the set

$$K_\varepsilon = \{x \in \mathbb{R}^d : |x_1 - 4| < \varepsilon \text{ and } |x'| < \varepsilon\}.$$

We used the notation $x = (x_1, x')$. The operator $L$ we are considering here is an integral operator with respect to a singular measure (singular along the axis $y' = 0$). That is the only reason why it does not have the form (1). Indeed, in (1) we implicitly assumed that for every $(t, x)$ the integral equation has a absolutely continuous measure with density $K(t, x, \cdot)$. That is a choice for convenience of notation only. Indeed, the non negative solution $u$ to the equation (23) satisfies the two inequalities in the viscosity sense

$$\begin{aligned} u_t - M^- u \geq 0 \qquad & \text{in } Q_1, \\ u_t - M^+ u \leq 0 \qquad & \text{in } Q_1. \end{aligned}$$

So, it is a valid candidate for a Harnack inequality. However, we will prove that

$$\lim_{\varepsilon \to 0} \frac{u(0, (0, x'))}{u(-1/2, 0)} = 0, \tag{24}$$

provided that $x' \neq 0$. This contradicts the parabolic Harnack inequality.

The intuition behind this counterexample is similar to the constructions given in [7] and [39] (for very different types of kernels). The equation we consider is the generator of a Levy process with frequent purely horizontal jumps. For $\varepsilon \ll 1$, a process starting at the origin would have a much higher chance to exit the domain at a point with $|x_1| < \varepsilon$ than a process that starts outside of a band $|x_1| > \delta$.

In order to verify (24), we will compute the solution to (23) almost explicitly. In order to write a formula for $u$ we will use the heat kernel associated with $L$ and Duhamel formula.

The heat kernel associated with this equation is explicit in Fourier side:

$$\hat{H}(t, \xi) = c \exp(-t|\xi|^\alpha - t|\xi_1|^\alpha).$$

In real variables, $H(t,x)$ is not explicit. But we know it must be the convolution of the heat kernel of $(-\Delta)^{\alpha/2}$ and the heat kernel of $(-\partial_{x_1x_1})$. That is

$$H(t,x) = \Phi_d(t,x) * \Phi_1(t,x),$$

where

$$\Phi_d(t,x) = t^{-d/\alpha}\Phi_d(1, x/t^{1/\alpha}) =: t^{-d/\alpha}\varphi_d\left(\frac{x}{t^{1/\alpha}}\right).$$

The exact formula for the $d$-dimensional fractional hear kernel $\varphi_d$ is not known. However, we know that

$$\varphi_d(x) \approx (1+|x|)^{-d-\alpha}.$$

Here we use the symbol $\approx$ to say that the ratio between the left hand side and right hand side is bounded below and above by positive constants.

The heat kernel $\Phi_1(t,x)$ is simply the one dimensional heat kernel of the fractional Laplacian in the variable $x_1$. Thus, it is a singular measure supported on the line $x' = 0$ with a density of the form

$$t^{-1-\alpha}\varphi_1\left(\frac{x_1}{t^{1/\alpha}}\right).$$

As before, we have $\varphi_1(x) \approx (1+|x|)^{-1-\alpha}$.

Using Duhamel formula, and following routine arithmetic manipulations, we arrive at a formula for $u(T,x)$.

$$u(T,x) = \int_0^{T+1}\int_{K_\varepsilon}\int_{\mathbb{R}} t^{-\frac{d+1}{\alpha}}\varphi_d\left(\frac{x-z-y_1e_1}{t^{1/\alpha}}\right)\varphi_1\left(\frac{y_1}{t^{1/\alpha}}\right)\,\mathrm{d}y_1\,\mathrm{d}z\,\mathrm{d}t.$$

We will first estimate $u(-1/2,0)$ in terms of $\varepsilon$. That is, we set $T=-1/2$ and $x=0$.

Assume $\varepsilon \ll 1$. We obtain a lower bound by restricting the domain of integration to a smaller domain.

$$u(-1/2,0) \geq \int_{K_\varepsilon}\int_0^{|z'|^\alpha}\int_{-4-\varepsilon}^{-4+\varepsilon} t^{-\frac{d+1}{\alpha}}\varphi_d\left(\frac{-z-y_1e_1}{t^{1/\alpha}}\right)\varphi_1\left(\frac{y_1}{t^{1/\alpha}}\right)\,\mathrm{d}y_1\,\mathrm{d}t\,\mathrm{d}z,$$

In this whole domain of integration we have

$$\varphi_d\left(\frac{-z-y_1e_1}{t^{1/\alpha}}\right) \geq c\left(1+\frac{\varepsilon}{t^{1/\alpha}}\right)^{-d-\alpha} \geq c\varepsilon^{-d-\alpha}t^{d/\alpha+1},$$

$$\varphi_1\left(\frac{y_1}{t^{1/\alpha}}\right) \geq c\left(1+\frac{4}{t^{1/\alpha}}\right)^{-1-\alpha} \geq ct^{1+1/\alpha}$$

Therefore

$$u(-1/2,0) \geq c\varepsilon^{-d-\alpha+1}\int_{K_\varepsilon}\int_0^{|z'|^\alpha} t^2\,\mathrm{d}t\,\mathrm{d}z,$$

$$= c\varepsilon^{2\alpha+1}$$

Now we estimate $u(0,(0,x'))$ from above in terms of $\varepsilon$. That is, we set $T = -1/2$ and $x = (0,x')$ for some non zero $x' \in \mathbb{R}^{d-1}$ with $|x'| < 1/2$. Using that $\varphi_d(x) \leq C|d|^{-d-\alpha}$ and $\varphi_1(x) \leq C|x|^{-1-\alpha}$, we get

$$u(0,(0,x')) \leq C \int_0^{T+1} \int_{K_\varepsilon} \int_{\mathbb{R}} t^{1-\frac{1}{\alpha}} \left(|x'-z'| + |x_1 - z_1 - y_1|\right)^{-d-\alpha} \varphi_1\left(\frac{y_1}{t^{1/\alpha}}\right) \,\mathrm{d}y_1 \,\mathrm{d}z \,\mathrm{d}t.$$

Since $x' \neq 0$, then $|x'-z'| > |x'|/2$ provided that $\varepsilon < |x'|/2$. That is, $|x'-z'|$ is of order one as $\varepsilon \to 0$. Therefore

$$\begin{aligned} u(0,(0,x')) &\leq C \int_0^{T+1} \int_{K_\varepsilon} \int_{\mathbb{R}} t^{1-\frac{1}{\alpha}} \varphi_1\left(\frac{y_1}{t^{1/\alpha}}\right) \,\mathrm{d}y_1 \,\mathrm{d}z \,\mathrm{d}t, \\ &= C \int_0^{T+1} \int_{K_\varepsilon} t \,\mathrm{d}z \,\mathrm{d}t = C\varepsilon^d. \end{aligned}$$

So, if $d > 1+2\alpha$, we obtain that $u(0,(0,x')) \ll u(-1/2,0)$ as $\varepsilon \to 0$. Therefore the Harnack inequality does not hold.

**Remark 7.1.** It is not clear whether the condition $d > 1+2\alpha$ is a limitation of this construction or the standard Harnack inequality actually holds for $d \leq 1+2\alpha$. We leave it as an open question.

## 8. Applications

In this last section we give a brief summary of applications to the estimates of Theorems 6.1 and (6.2).

**8.1. Fully nonlinear parabolic equations.** One of the canonical applications of the Hölder estimates presented in this paper is the $C^{1,\gamma}$ regularity for solutions to the parabolic Isaacs equation.

$$u_t - \inf_a \sup_b \int_{\mathbb{R}^d} \delta_y u(x) K_{ab}(y) \,\mathrm{d}y = 0.$$

The Isaacs equation models the value function for the optimal strategy in a zero-sum stochastic game. For the purpose of this article, we consider games driven by Levy processes without diffusion. We point out that discontinuous Levy processes have a number of applications in finance [46] and physics [36].

It is easy to check that if all kernels $K_{ab}$ satisfy the assumptions (5) and (4), then the incremental quotients $v_h(x) = (u(x+h) - u(x))/|h|$ satisfy the assumptions of Theorem (6.1) or Theorem (6.2). This quickly leads to a $C^{1,\gamma}$ regularity result for $u$ at least if the equation holds in the whole space $(0,\infty)\times\mathbb{R}^d$. If the equation holds in a bounded domain, there are some extra difficulties. The method is explained in [40]. The result there applies to a less general class of kernels but it is robust, since it is based on the Hölder estimates from [34].

**8.2. Active scalar equations.** There exist several active scalar equations of the form

$$\theta_t + B(\theta) \cdot \theta + (-\Delta)^{\alpha/s}\theta = 0, \tag{25}$$

that have attracted attention in recent years. Here $B(\theta)$ is a vector field which depends on the solution $\theta$ of the equation. This dependence makes the equation non linear. Some examples of $B$ which are of interest are the following.

- **Conservation laws with fractional diffusion.** $B(\theta) = F_i'(\theta)$. See for example [21].
- **Surface quasi-geostrophic equation.** $B(\theta) = R^{\perp}\theta$, where $R$ stands for the Riesz tranform. See for example [29], [12] and [16] among many others.
- **Modified surface quasi-geostrophic equation.** $B(\theta) = R^{\perp}(-\Delta)^{1-\alpha}\theta$, where $R$ stands for the Riesz tranform. See for example [15].
- **Incompressible flow in porous media.** $B(\theta) = (0, -\theta) - \nabla p$, so that $\operatorname{div} B = 0$. See for example [17].

The solution to any of these equations is a priori bounded in $L^\infty$ from the maximum principle. The key step in order to prove that they posses a classical global solution is to be able to obtain a regularity estimate for the solution which goes beyond $L^\infty$. Once a Hölder estimate is established, it can be bootstrapped into higher regularity using the result from [42] in any of the models above.

Theorem 6.1 gives us a Hölder estimate for conservation laws with critical fractional diffusion $\alpha = 1$. It also gives us a Hölder estimate for the modified surface quasi-geostrophic equation if $\alpha \in (0, 1)$. Therefore, the classical well-posedness of both models follows.

The study of either the surface quasi-geostrophic model or the fluid in porous media with critical diffusion $\alpha = 1$ does not follow immediately from Theorem 6.1. This is because Theorem 6.1 requires the vector field to be bounded and in these cases $B(\theta)$ is a priori only controlled in $L^\infty((0,\infty), BMO)$. A version of Theorem 6.1 for vector fields $b \in L^\infty(BMO)$ was given in [12] provided that $\operatorname{div} b = 0$, but the result are of very different nature. Indeed, the result in [12] is based on the variational structure of the equation and uses De Giorgi's technique.

**8.3. The space homogeneous Boltzmann equation.** The Boltzmann equation models the evolution of dilute gasses. In the space homogeneous case, the equation takes the form

$$f_t = Q(f, f),$$

where

$$Q(g, f) = \int_{\mathbb{R}^d} \int_{S_1} \left(g(v_*')f(v') - g(v_*)f(v)\right) \, \mathrm{d}\sigma \, \mathrm{d}v_*$$

and we write

$$v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2}\sigma, \qquad v_*' = \frac{v + v_*}{2} - \frac{|v - v_*|}{2}\sigma.$$

For every fixed $g$, the operator $f \mapsto Q(g,f)$ is an integro-differential operator which has the form

$$Q(g,f) = R_g f(v) + \int_{\mathbb{R}^d} \delta_{v'} f(v) K_g(v, v'-v) \,\mathrm{d}v'.$$

The function $R_g$ and the kernel $K_g$ can be computed in terms of $g$, although the formula is quite involved.

Under some conditions, the kernel $K_f$ satisfies the assumptions (5) and (4) and consequently Theorems 6.1 and 6.2 may be used to prove a local Hölder continuity result for the Boltzmann equation.

In order to apply the result of Theorems 6.1 and 6.2 we would need to consider a collision kernel without Grad's angular cutoff condition. Moreover, we should know a priori that $f$ is bounded below in order to guarantee that $K_f$ satisfies (4). This last assumption in particular is quite undesirable. In a work in progress of Russell Schwab and the author, we are developing a more general Hölder estimate which, for some collision kernels, would only depend on observable quantities associated with $f$ (mass, energy and entropy).

Note that the regularity of the solutions to the inhomogeneous Boltzmann equation is rather well understood by completely different methods [18].

# References


[1] Martin T. Barlow, Richard F. Bass, Zhen-Qing Chen, and Moritz Kassmann. Non-local Dirichlet forms and symmetric jump processes. *Trans. Amer. Math. Soc.*, 361(4):1963–1999, 2009.

[2] Richard F Bass and Moritz Kassmann. Harnack inequalities for non-local operators of variable order. *Transactions of the American Mathematical Society*, pages 837–850, 2005.

[3] Richard F Bass and Moritz Kassmann. Hölder continuity of harmonic functions with respect to operators of variable order. *Communications in Partial Differential Equations*, 30(8):1249–1259, 2005.

[4] Richard F Bass and David A Levin. Harnack inequalities for jump processes. *Potential Analysis*, 17(4):375–388, 2002.

[5] Richard F. Bass and David A. Levin. Transition probabilities for symmetric jump processes. *Trans. Amer. Math. Soc.*, 354(7):2933–2953 (electronic), 2002.

[6] C. Bjorland, L. Caffarelli, and A. Figalli. Non-local gradient dependent operators. *Adv. Math.*, 230(4-6):1859–1894, 2012.

[7] Krzysztof Bogdan and Paweł Sztonyk. Harnack's inequality for stable Lévy processes. *Potential Anal.*, 22(2):133–150, 2005.

[8] Luis Caffarelli, Chi Hin Chan, and Alexis Vasseur. Regularity theory for parabolic nonlinear integral operators. *J. Amer. Math. Soc.*, 24(3):849–869, 2011.

[9] Luis Caffarelli and Luis Silvestre. Regularity theory for fully nonlinear integro-differential equations. *Communications on Pure and Applied Mathematics*, 62(5):597–638, 2009.

[10] Luis Caffarelli and Luis Silvestre. The evans-krylov theorem for nonlocal fully nonlinear equations. *Annals of mathematics*, 174(2):1163–1187, 2011.

[11] Luis Caffarelli and Luis Silvestre. Regularity results for nonlocal equations by approximation. *Archive for rational mechanics and analysis*, 200(1):59–88, 2011.

[12] Luis Caffarelli and Alexis Vasseur. Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation. *Ann. of Math*, 171(3):1903–1930, 2010.

[13] Héctor Chang Lara and Gonzalo Dávila. Regularity for solutions of nonlocal, nonsymmetric equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 29(6):833–859, 2012.

[14] Hector A. Chang Lara. Regularity for fully non linear equations with non local drift. *arXiv preprint arXiv:1210.4242*, 2013.

[15] Peter Constantin, Gautam Iyer, and Jiahong Wu. Global regularity for a modified critical dissipative quasi-geostrophic equation. *arXiv preprint arXiv:0803.1318*, 2008.

[16] Peter Constantin and Vlad Vicol. Nonlinear maximum principles for dissipative linear nonlocal operators and applications. *Geometric And Functional Analysis*, 22(5):1289–1321, 2012.

[17] Diego Córdoba, Francisco Gancedo, and Rafael Orive. Analytical behavior of two-dimensional incompressible flow in porous media. *Journal of mathematical physics*, 48(6):065206, 2007.

[18] Laurent Desvillettes and Bernt Wennberg. Smoothness of the solution of the spatially homogeneous boltzmann equation without cutoff. *Comm. Partial Differential Equations*, 29(1-2):133–155, 2004.

[19] Hongjie Dong and Doyoon Kim. Schauder estimates for a class of non-local elliptic equations. *arXiv preprint arXiv:1103.5069*, 2011.

[20] Hongjie Dong and Doyoon Kim. On $L_p$-estimates for a class of non-local elliptic equations. *Journal of Functional Analysis*, 262(3):1166–1199, 2012.

[21] Jérôme Droniou and Cyril Imbert. Fractal first-order partial differential equations. *Archive for Rational Mechanics and Analysis*, 182(2):299–331, 2006.

[22] Lawrence C. Evans. Classical solutions of fully nonlinear, convex, second-order elliptic equations. *Comm. Pure Appl. Math.*, 35(3):333–363, 1982.

[23] Matthieu Felsinger and Moritz Kassmann. Local Regularity for Parabolic Nonlocal Operators. *Comm. Partial Differential Equations*, 38(9):1539–1573, 2013.

[24] Nestor Guillen and Russell W Schwab. Aleksandrov–bakelman–pucci type estimates for integro-differential equations. *Archive for Rational Mechanics and Analysis*, 206(1):111–157, 2012.

[25] Cyril Imbert and Luis Silvestre. An introduction to fully nonlinear parabolic equations. In *An Introduction to the Kähler-Ricci Flow*, pages 7–88. Springer, 2013.

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

[27] Moritz Kassmann and Ante Mimica. Intrinsic scaling properties for nonlocal operators. *arXiv preprint arXiv:1310.5371*, 2013.

[28] Moritz Kassmann and Russell Schwab. Regularity results for nonlocal parabolic equations. *arXiv preprint arXiv:1305.5418*, 2013.

[29] A. Kiselev, F. Nazarov, and A. Volberg. Global well-posedness for the critical 2D dissipative quasi-geostrophic equation. *Invent. Math.*, 167(3):445–453, 2007.

[30] Takashi Komatsu. Uniform estimates for fundamental solutions associated with non-local Dirichlet forms. *Osaka J. Math.*, 32(4):833–860, 1995.

[31] N. V. Krylov. Boundedly inhomogeneous elliptic and parabolic equations. *Izv. Akad. Nauk SSSR Ser. Mat.*, 46(3):487–523, 670, 1982.

[32] Nikolai Vladimirovich Krylov and Mikhail V Safonov. A certain property of solutions of parabolic equations with measurable coefficients. *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, 44(1):161–175, 1980.

[33] NV Krylov and MV Safonov. Estimate of the possibility of diffusion process hitting into a set of positive measure. *Doklady Akademii Nauk SSSR*, 245(1):18–20, 1979.

[34] Héctor Chang Lara and Gonzalo Dávila. Regularity for solutions of non local parabolic equations. *Calculus of Variations and Partial Differential Equations*, 49(1-2):139–172, 2014.

[35] Héctor Chang Lara and Gonzalo Dávila. Regularity for solutions of non local parabolic equations. *Calc. Var. Partial Differential Equations*, 49(1-2):139–172, 2014.

[36] Ralf Metzler and Joseph Klafter. The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. *Journal of Physics A: Mathematical and General*, 37(31):R161, 2004.

[37] Nikolai Nadirashvili and Serge Vlăduţ. Nonclassical solutions of fully nonlinear elliptic equations. *Geometric and Functional Analysis*, 17(4):1283–1296, 2007.

[38] Nikolai Nadirashvili and Serge Vlăduţ. Singular solutions of hessian elliptic equations in five dimensions. *Journal de Mathématiques Pures et Appliquées*, 100(6):769–784, 2013.

[39] Marcus Rang, Moritz Kassmann, and Russell W Schwab. Hölder regularity for integro-differential equations with nonlinear directional dependence. *arXiv preprint arXiv:1306.0082*, 2013.

[40] Joaquim Serra. Regularity for fully nonlinear nonlocal parabolic equations with rough kernels. *arXiv preprint arXiv:1401.4521*, 2014.

[41] Luis Silvestre. Holder estimates for solutions of integro-differential equations like the fractional laplace. *Indiana University mathematics journal*, 55(3):1155–1174, 2006.

[42] Luis Silvestre. On the differentiability of the solution to an equation with drift and fractional diffusion. *arXiv preprint arXiv:1012.2401*, 2010.

[43] Luis Silvestre. On the differentiability of the solution to the hamilton–jacobi equation with critical fractional diffusion. *Advances in mathematics*, 226(2):2020–2039, 2011.

[44] Luis Silvestre. Hölder estimates for advection fractional-diffusion equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 11(4):843–855, 2012.

[45] Renming Song and Zoran Vondraček. Harnack inequality for some classes of Markov processes. *Math. Z.*, 246(1-2):177–202, 2004.

[46] Peter Tankov. *Financial modelling with jump processes.* CRC Press, 2004.

Luis Silvestre
Dept. of Mathematics
University of Chicago
5734 S. University Avenue, Room 208C
Chicago, Illinois 60637
E-mail: luis@math.uchicago.edu