POTENTIAL THEORY FOR NONLOCAL DRIFT-DIFFUSION EQUATIONS

QUOC-HUNG NGUYEN, SIMON NOWAK, YANNICK SIRE,
AND MARVIN WEIDNER

ABSTRACT. The purpose of this paper is to prove new fine regularity results for nonlocal drift-diffusion equations via pointwise potential estimates. Our analysis requires only minimal assumptions on the divergence free drift term, enabling us to include drifts of critical order belonging merely to BMO. In particular, our results allow to derive new estimates for the dissipative surface quasi-geostrophic equation.

CONTENTS

1. INTRODUCTION

The goal of this paper is to derive new regularity results for drift-diffusion equations with a critical dissipative term in $\mathbb{R}^d$. More precisely, we establish fine estimates for solutions to such equations by means of parabolic Riesz potentials that are independent of the vector field in the drift term, except for its *a priori* regularity (see [Kis10] for a survey and references therein). The framework of this article is very general and, in particular, it yields new estimates for the critical dissipative SQG equation. Such equation has been introduced in [CMT94] as a toy model for the regularity of Navier-Stokes and several seminal works in the last years have led to major breakthroughs on this topic, see [CV10, KNV07, KN09]. Moreover, let us point out that some of the fine regularity estimates obtained in this article are even new for the standard fractional heat equation.

For $d \geq 2$, we consider the following general drift-diffusion equation

$$(1.1) \qquad \partial_t u + (b, \nabla u) + \mathcal{L}_t u = \mu \text{ in } I \times \Omega \subset \mathbb{R} \times \mathbb{R}^d = \mathbb{R}^{d+1},$$

where μ belongs to the set of finite measures in $\mathbb{R}^{d+1}$, which we denote by $\mathcal{M}(\mathbb{R}^{d+1})$, while $b : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is a vector field that belongs to either $L^\infty(\mathbb{R}^{d+1})$ or $L^\infty(\mathbb{R}; \mathrm{BMO}(\mathbb{R}^d))$, and satisfies $\mathrm{div}(b(t)) = 0$ almost everywhere in time $t \in \mathbb{R}$.

The integral operator $\mathcal{L}_t$ $(t \in \mathbb{R})$ is defined by

$$(1.2) \qquad \mathcal{L}_t u(x) = \text{P.V.} \int_{\mathbb{R}^d} (u(x) - u(y)) K_t(x, y) \mathrm{d}y,$$

where the kernel $K_t : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty)$, $(t, x, y) \mapsto K_t(x, y)$ is assumed to satisfy the following conditions for a given $s \in [1/2, 1)$:

  (i) K_t is a measurable function;
  (ii) for every $t \in \mathbb{R}$, K_t is symmetric which means that

$$K_t(x, y) = K_t(y, x), \quad \text{for a.e. } (t, x, y) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d;$$

  (iii) there exists a constant $\Lambda \geq 1$ such that

$$\Lambda^{-1} \leq |x - y|^{d+2s} K_t(x, y) \leq \Lambda, \text{ for a.e. } (t, x, y) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d.$$

An example of an operator $\mathcal{L}_t$ is of course given by the Fourier multiplier $(-\Delta)^s$. In this case equation (1.1) refers to a standard class of drift-diffusion equations. In the case $d = 2$, $s = \frac{1}{2}$, when b is given by a rotation of the vectorial Riesz transform $b = \nabla^{\perp}(-\Delta)^{-\frac{1}{2}}u$, one obtains the dissipative surface quasi-geostrophic equation (SQG equation). We note at this point that in the case $s = \frac{1}{2}$, the drift term has the same order as the diffusion term given by the nonlocal operator in (1.1), so that the problem becomes *critical* instead of subcritical, leading to substantial additional technical difficulties compared to parabolic nonlocal equations without drift.

The main results of the present paper establish potential-theoretic estimates of (weak) solutions u to (1.1) in terms of parabolic Riesz potentials of the measure μ (see Theorem 1.2 and Theorem 1.4). In this context, we extend similar results previously obtained in the nonlocal elliptic setting (see [KMS15, KMS18]) to the setting of time-dependent drift-diffusion equations, including in particular the dissipative SQG equation. Moreover, thanks to the known mapping properties of the potentials, this framework provides new estimates in terms of finer scales of function spaces. This philosophy of obtaining fine regularity estimates via potentials follows a by now long tradition started in the context of local elliptic and parabolic equations, see e.g. [KM94, TW02, Min11, DM11, Cia11, KM14, KM18, BCD$^+$18, BY19, DZ21, DF22].

Let us now explain and discuss our main results in detail. We consider the cases $b \in L^{\infty}(\mathbb{R}^{d+1})$ and $b \in L^{\infty}(\mathbb{R}; \mathrm{BMO}(\mathbb{R}^d))$ separately.

Main results for bounded drifts. A crucial step in the construction of potential estimates is to establish suitably localized Hölder estimates for weak solutions to the homogeneous problem corresponding to (1.1) in parabolic cylinders. In our case, the homogeneous problem is given by

$$(1.3) \qquad \partial_t v + (b, \nabla v) + \mathcal{L}_t v = 0 \text{ in } Q_r(t_0, x_0),$$

where $Q_r(t_0, x_0) := I_r^{\ominus}(t_0) \times B_r(x_0)$ denotes the parabolic cylinder and $I_r^{\ominus}(t_0) := (t_0 - r^{2s}, t_0)$.

Weak solutions to the homogeneous problem (1.3) and the inhomogeneous problem (1.1) are defined as follows:

Definition 1 (Weak solutions). Let $\mu \in \mathcal{M}(\mathbb{R}^{d+1})$. A function $u \in L^\infty(I; L^2(\Omega)) \cap L^2(I; H^s(\Omega)) \cap L^2(I; L^1_{2s}(\mathbb{R}^d))$ is a weak subsolution to (1.1), if for all $(t_1, t_2) \subset I$ and nonnegative test functions $\varphi \in L^2((t_1, t_2); H^s(\mathbb{R}^d)) \cap C^1((t_1, t_2); L^2(\Omega))$ with $\varphi \equiv 0$ in $(t_1, t_2) \times \mathbb{R}^d \setminus \Omega$ we have

$$-\int_{t_1}^{t_2} \int_\Omega u \left[\partial_t \varphi + (b, \nabla \varphi)\right] \mathrm{d}x\mathrm{d}t + \int_\Omega u(t_2, x)\varphi(t_2, x)\mathrm{d}x - \int_\Omega u(t_1, x)\varphi(t_1, x)\mathrm{d}x$$

$$+ \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (u(t, x) - u(t, y))(\varphi(t, x) - \varphi(t, y))K_t(x, y)\mathrm{d}y\mathrm{d}x\mathrm{d}t$$

$$\leq \int_{t_1}^{t_2} \int_\Omega \varphi \mathrm{d}\mu.$$

Here, $L^1_{2s}(\mathbb{R}^d)$ denotes the nonlocal tail space

$$L^1_{2s}(\mathbb{R}^d) := \left\{ v \in L^1_{\mathrm{loc}}(\mathbb{R}^d) \ \Big| \ \int_{\mathbb{R}^d} \frac{|v(y)|}{(1 + |y|)^{d+2s}} \mathrm{d}y < \infty \right\}.$$

A function u is called a weak supersolution if the previous estimate holds true for every nonpositive test function ϕ, and it is called a weak solution if it is a weak subsolution and a weak supersolution.
u is called a weak (sub/super)solution to (1.3) if it is a weak (sub/super)solution to (1.1) with $\mu \equiv 0$.

Next, we state the localized Hölder regularity estimate for solutions to the homogeneous problem (1.3) under the assumption that the vector field b is bounded.

Theorem 1.1 (Hölder estimate). *Let $s \in [1/2, 1)$ and $b \in L^\infty(\mathbb{R}^{d+1})$ with $\|b\|_{L^\infty(\mathbb{R}^{d+1})} \leq \Lambda$. Suppose that v is a weak solution to (1.3) in a parabolic cylinder $Q_r(t_0, x_0)$ for some $r > 0$. Then, there exist $\alpha \in (0, 1)$ and $c > 0$ depending only on d, s, Λ, q such that $v \in C^\alpha_{loc}(Q_r(t_0, x_0))$ and the following estimate holds true for any $q > 1$:*

$$\|v\|_{C^\alpha(Q_{r/2}(t_0, x_0))}$$

$$\leq cr^{-\alpha} \left[\left(\fint_{Q_r(t_0, x_0)} |v| \mathrm{d}x\mathrm{d}t \right) + \left(\fint_{I^\ominus_r(t_0)} \mathrm{Tail}(v(t); x_0, r)^q \mathrm{d}t \right)^{1/q} \right],$$

where the nonlocal tail term is defined as

$$\mathrm{Tail}(v(t); x_0, r) = r^{2s} \int_{\mathbb{R}^d \setminus B_r(x_0)} |v(t, y)||x_0 - y|^{-d-2s} \mathrm{d}y.$$

The previous result is already new in contrast to previous Hölder estimates for nonlocal drift-diffusion equations (see e.g. [CV10], [NST23], [DS18]) because of two reasons. First, our estimate is truly local in nature in the sense that it does not require v to solve an equation in the full space. Second, our estimate of the Hölder norm only contains tail terms which are merely required to belong to L^q in time for some $q > 1$. Since this quantity is always finite for any weak solution, our result does not impose any artificial restriction on the behavior of solutions at infinity. Note that Theorem 1.1 does not hold with $q = 1$ due to [KW23a, Example 5.2].

In comparison to previous results on Hölder regularity for nonlocal diffusion equations, our result can be regarded as an extension of [KW23a] since it allows for subcritical and critical drift terms. For previous results on nonlocal elliptic and parabolic equations without drift see for instance [Kas09, CCV11, FK13, DCKP16, Coz17, KW22a]. Finally, let us mention the articles [Sil11, Sil12a, Sil12b], where Hölder estimates have been obtained for nonlocal drift-diffusion equations in case $\mathcal{L}_t = (-\Delta)^s$ using non-variational techniques.

We proceed by introducing parabolic Riesz potentials:

Definition 2. Let $\mu \in \mathcal{M}(\mathbb{R}^{d+1})$ be a finite Radon measure with finite total mass and consider a parabolic cylinder $Q_\rho(t_0, x_0) = (t_0 - \rho^{2s}, t_0) \times B_\rho(x_0)$. Then we denote by $\mathcal{P}_a^R[\mu]$ the parabolic Riesz potential of μ with order $a \in (0, d + 2s)$ which is defined as follows

$$\mathcal{P}_a^R[\mu](t_0, x_0) = \int_0^R \frac{|\mu|(Q_\rho(t_0, x_0))}{\rho^{d+2s-a}} \frac{d\rho}{\rho}.$$

In particular, one has

$$\mathcal{P}_{2s}^R[\mu](t_0, x_0) = \int_0^R \frac{|\mu|(Q_\rho(t_0, x_0))}{\rho^d} \frac{d\rho}{\rho}.$$

Remark 1.1. Note that while the order $2s$ of the parabolic potential does not explicitly appear in its definition, it is intrinsically encoded via the size of the parabolic cylinder $Q_\rho(t_0, x_0)$.

We are now in position to state our first main result, a zero-order potential estimate for solutions to (1.1) under the assumption that the drift b is bounded uniformly in space and time, which in particular includes solutions to the SQG equation.

Theorem 1.2 (Potential estimate)**.** *Let $s \in [1/2, 1)$, $\mu \in \mathcal{M}(\mathbb{R}^{d+1})$, assume that $b \in L^\infty(\mathbb{R}^{d+1})$ with $\|b\|_{L^\infty(\mathbb{R}^{d+1})} \leq \Lambda$, and that u is a weak solution of* (1.1) *in $I \times \Omega \subset \mathbb{R}^{d+1}$. Then for almost all $t_0 \in I$, $x_0 \in \Omega$, any $q > 1$, and any $R > 0$ such that $Q_R(t_0, x_0) \subset I \times \Omega$, we have the estimate*
(1.4)
$$|u(t_0, x_0)| \leq c \left[\left(\fint_{Q_R(t_0, x_0)} |u|^q dxdt \right)^{1/q} + \left(\fint_{I_R^\ominus(t_0)} \mathrm{Tail}(u(t); x_0, R)^q dt \right)^{1/q} \right.$$
$$\left. + \mathcal{P}_{2s}^R[\mu](t_0, x_0) \right],$$

where c depends only on d, s, Λ, q.

Remark 1.2 (Equations without drift)**.** In case $b = 0$, that is, when (1.1) is a parabolic nonlocal diffusion equation without drift, the potential estimate Theorem 1.2 remains valid for the full range $s \in (0, 1)$ without any changes in the proof (see also Remark 4.1).

Our main result Theorem 1.2 extends corresponding potential estimates in the local parabolic case (see e.g. [DM11]) and in the nonlocal elliptic case

(see e.g. [KMS15, KMS18, KNS22, DN23]) to the setting of nonlocal drift-diffusion equations. Moreover, it seems to be new already for the fractional heat equation with $\mathcal{L}_t = (-\Delta)^s$, $s \in (0,1)$, and $b \equiv 0$ (see Remark 1.2).

Remark 1.3 (SOLA)**.** While for the sake of simplicity we work with the above notion of energy-type weak solutions, in view of standard approximation arguments we expect our main results to remain valid for distributional solutions which can be approximated by weak solutions to regularized equations in a suitable fashion. Solutions of this type are often called SOLA (= solutions obtained by limiting approximations) and can usually be shown to exist under general measure data. See for instance [KMS15, KNS22] for a precise solution concept of this type in the nonlocal elliptic setting and [BDGO97, DM11, KM14] for corresponding ones in the local parabolic setting.

Our main technical tool for proving Theorem 1.2 is the following excess-decay lemma, which in particular encodes the zero-order regularity properties of solutions to (1.1) in a precise way.

Lemma 1.1 (Excess decay)**.** *Let $s \in [1/2, 1)$, $R > 0$, $t_0 \in I$, $x_0 \in \mathbb{R}^n$, $q > 1$, $\mu \in \mathcal{M}(\mathbb{R}^{d+1})$, and assume that $b \in L^\infty(\mathbb{R}^{d+1})$ with $\|b\|_{L^\infty(\mathbb{R}^{d+1})} \leq \Lambda$. In addition, let $\alpha \in (0, s(1 - 1/q))$ be given by Theorem 1.1. Let u be a weak solution to (1.1) in $I \times \Omega$ and assume that $Q_{2R}(t_0, x_0) \subset I \times \Omega$. Then for any integer $m \geq 1$, we have*

$$(1.5) \qquad \begin{aligned} E(u, t_0, x_0, 2^{-m}R) &\leq C_0 2^{-\alpha m} E(u, t_0, x_0, R) \\ &\quad + C_0 2^{((d+2s)/q)m} R^{-d} |\mu|(Q_R(t_0, x_0)), \end{aligned}$$

where $C_0 \geq 1$ depends only on d, s, Λ, α, q and for any $r > 0$,

$$\begin{aligned} E(u, t_0, x_0, r) := &\left(\fint_{I_r^\ominus(t_0)} \left[\fint_{B_r(x_0)} |u - (u)_{Q_r(t_0, x_0)}| \mathrm{d}x \right]^q \mathrm{d}t \right)^{1/q} \\ &+ r^{-2s/q} \|\operatorname{Tail}(u(\cdot) - (u)_{Q_r(t_0, x_0)}; x_0, r)\|_{L^q(I_r^\ominus(t_0))}. \end{aligned}$$

The proof of the previous parabolic potential estimate is nonlinear in nature, not because the equation (1.1) is a nonlinear PDE, but as for the techniques we use, which are reminiscent of the works [DM11, KMS18]. In particular, we believe that our approach has the potential to be extended to obtaining similar results for nonlinear generalizations of the equation (1.1).

Let us end this section by providing an alternative perspective. In this paper, we also derive upper estimates for the fundamental solution of the operator $\partial_t + (b, \nabla \cdot) + \mathcal{L}_t$ (heat kernel). Such estimates could in principle also be used to derive zero-order potential estimates for solutions of (1.1).

Theorem 1.3 (Upper heat kernel estimate)**.** *Let $s \in [1/2, 1)$ and assume that $b \in L^\infty(\mathbb{R}^{d+1})$ with $\|b\|_{L^\infty(\mathbb{R}^{d+1})} \leq \Lambda$. Let $0 \leq \eta \leq t \leq T$ and let p be the heat kernel associated to the operator $\partial_t + (b, \nabla \cdot) + \mathcal{L}_t$ defined on $\mathbb{R}^d$. Then the following holds*

$$(1.6) \qquad p(\eta, x; t, y) \leq c \frac{t - \eta}{\left(|x - y| + |t - \eta|^{\frac{1}{2s}} \right)^{d+2s}} \qquad \forall x, y \in \mathbb{R}^d$$

for some constant $c > 0$, depending only on d, s, Λ, T.

Our proof of Theorem 1.3 is based on linear techniques and follows a strategy going back to the celebrated work by Aronson (see [Aro68]), which has recently been extended to a nonlocal setting by the last-named author (see [KW23b]). We note that similar upper bounds for the heat kernel of drift-diffusion equations have been proved in [MM13b], [MM13a] by using an adaptation of the celebrated Davies' method (see [Dav87], [Dav89]), which was first applied to nonlocal problems in [CKS87].

Fine space-time regularity. In the following, we mention some consequences of our nonlocal potential estimate Theorem 1.2. Indeed, it is well-known (see e.g. [DM11, Ngu14] and the references therein) that potential estimates lead to space-time estimates for solutions, using the mapping properties of the parabolic Riesz potential in Lorentz/Marcinkiewicz spaces. More precisely, we have the following fine regularity results.

Corollary 1.1 (Lorentz regularity). *Let $s \in [1/2, 1)$, $\mu \in \mathcal{M}(\mathbb{R}^{d+1})$, assume that $b \in L^\infty(\mathbb{R}^{d+1})$, and that u is a weak solution of* (1.1) *in $I \times \Omega \subset \mathbb{R}^{d+1}$.*

- *For any $p \in \left(1, \frac{d+2s}{2s}\right)$ and any $\sigma \in (0, \infty]$, we have the implication*

(1.7)
$$\mu \in L^{p,\sigma}(I \times \Omega) \implies u \in L_{loc}^{\frac{p(d+2s)}{d+2s-2sp}, \sigma}(I \times \Omega).$$

- *We have*

(1.8)
$$\mu \in L^{\frac{d+2s}{2s}, 1}(I \times \Omega) \implies u \in L_{loc}^\infty(I \times \Omega).$$

Remark 1.4 (Equations without drift). In case $b = 0$, Corollary 1.1 remains valid for the full range $s \in (0, 1)$ by using the corresponding potential estimate that was discussed in Remark 1.2.

Main results for BMO drifts. Although we are mainly interested in the dissipative SQG system, we emphasize that our method to prove Theorem 1.2 allows to consider more general drift-diffusion equations of the form (1.1), as far as the regularity of the vector field b is concerned. We recall that the SQG system relates via a Biot-Savart law the vector field b to the solution u. Whenever such a law is not available, the problem becomes to ask under which minimal regularity assumption on b the equation (1.1) admits a local in time suitably regular solution. It is known since the seminal work by Caffarelli and Vasseur that dissipative SQG transports BMO-vector fields; which is more general than the standing assumption from our main results stated above that the drift is in L^∞ in space and time.

The aim of this section is to state results under the aforementioned more general BMO assumption. While the adaptation of our techniques to BMO-drifts is rather straightforward when $s > 1/2$, the critical case $s = 1/2$ requires us to adjust the parabolic cylinders in space to the transport that is caused by the critical drift (see Lemma 2.8). Similar arguments were already applied in [CV10], [NST23]. In fact, let $z_r(t)$ be the solution to the ODE

$$z_r'(t) = \fint_{B_1} b(rt, rx + rz_r(t)) \mathrm{d}x, \ z_r(0) = 0.$$

Then given $t_0 \in \mathbb{R}$, $x_0 \in \mathbb{R}^d$ and $R > 0$, we introduce the slanted cylinders $\tilde{Q}_R(t_0, x_0) = I_R^\ominus(t_0) \times \tilde{B}_R(x_0)$ by

$$\tilde{Q}_R(t_0, x_0) = \{(t, x + R z_R(t/R)) : (t, x) \in Q_R(t_0, x_0)\}.$$

Moreover, we define $\widetilde{\mathrm{Tail}}(v(t); x_0, R) = \mathrm{Tail}(v(t, \cdot + R z_R(t/R)); x_0, R)$, and introduce the following *slanted* parabolic potential. Let $\mu \in \mathcal{M}(\mathbb{R}^{d+1})$ be a Radon measure with finite total mass. Then we define

$$\widetilde{\mathcal{P}}_1^R[\mu](t_0, x_0) := \int_0^R \frac{|\mu|(\tilde{Q}_\rho(x, t))}{\rho^d} \frac{d\rho}{\rho}.$$

Note that in view of (2.28), the slanted potential can be controlled as follows

(1.9)
$$\widetilde{\mathcal{P}}_1^R[\mu](t_0, x_0) \leq \int_0^R \frac{|\mu|(I_\rho^\ominus(t_0) \times B_{c\rho(C_1 + C_2 |\log(\rho)|)}(x_0))}{\rho^d} \frac{d\rho}{\rho}$$
$$\leq \int_0^{f^{-1}(R)} \frac{|\mu|(I_\rho^\ominus(t_0) \times B_\rho(x_0))}{f^{-1}(\rho)^d f'(f^{-1}(\rho))} \frac{d\rho}{f^{-1}(\rho)},$$

where $f(\rho) = c\rho(C_1 + C_2 |\log(\rho)|)$. Here c depends only on d, Λ, q, while C_1 and C_2 are given as in the statement of Theorem 1.4. We believe these estimate to be helpful when investigating mapping properties for the slanted potential, which would be required in order to obtain an analog of Corollary 1.1 for BMO-drifts in case $s = 1/2$.

We can then prove the following theorem:

Theorem 1.4. *Let $s \in [1/2, 1)$ with $b \in L^\infty(\mathbb{R}; \mathrm{BMO}(\mathbb{R}^d))$. Assume furthermore that there exist constants $C_1, C_2 > 0$ such that*

(1.10)
$$\sup_{t, x_0} \fint_{B_1(x_0)} |b(t, x)| \mathrm{d}x \leq C_1, \qquad \|b\|_{L^\infty(\mathbb{R}; \mathrm{BMO}(\mathbb{R}^d))} \leq C_2.$$

Then the following holds:

- *Suppose $s \in (1/2, 1)$. Then the statements of Proposition 1.1 and Theorem 1.2 remain true under the more general assumptions on b indicated above. The constants depend in addition on C_1, C_2.*
- *If $s = 1/2$ and v is a weak solution to (1.3) in $Q_{\theta r}(t_0, x_0)$ for some $r > 0$, where $\theta = 1 + c_0(C_1 + C_2 |\log(r)|)$ for some $c_0 > 0$ depending only on d, then $v \in C_{loc}^\alpha(Q_r(t_0, x_0))$, and the following estimate holds true for any $q > 1$:*

(1.11)
$$\|v\|_{C^\alpha(\tilde{Q}_{r/2}(t_0, x_0))} \leq c r^{-\alpha} \left[\left(\fint_{\tilde{Q}_r(t_0, x_0)} |v| \mathrm{d}x \mathrm{d}t \right) \right.$$
$$\left. + \left(\fint_{I_r^\ominus(t_0)} \widetilde{\mathrm{Tail}}(v(t); x_0, r)^q \mathrm{d}t \right)^{1/q} \right]$$

for some $c > 0$, $\alpha \in (0,1)$, depending only on d, Λ, q, C_1, C_2.

Moreover, for any weak solution u to (1.1) in $Q_{\theta r}(t_0, x_0)$ we have

$$
|u(t_0, x_0)| \leq c \left(\fint_{\tilde{Q}_r(t_0,x_0)} |u|^q \mathrm{d}x\mathrm{d}t \right)^{1/q}
$$

(1.12)

$$
+ c \left(\fint_{I_r^{\ominus}(t_0)} \widetilde{\mathrm{Tail}}(u(t); x_0, R)^q \mathrm{d}t \right)^{1/q}
$$

$$
+ c \widetilde{\mathcal{P}}_1^r[\mu](t_0, x_0),
$$

where c depends only on d, Λ, q, C_1, C_2.

Outline. This article is structured as follows. In Section 2, we prove the localized Hölder regularity estimate Theorem 1.1 for bounded drifts, and for drifts in BMO (1.11). Section 3 establishes a comparison estimate between solutions to (1.1) and the homogeneous problem (1.3). The potential estimate for bounded drifts Theorem 1.2, and for drifts in BMO (1.12) are shown in Section 4. Finally, Section 5 establishes the heat kernel upper bound Theorem 1.3.

2. LOCAL HÖLDER ESTIMATES FOR HOMOGENEOUS PROBLEMS

Hölder regularity estimates for global solutions to (1.3) have already been established in [NST23]. Our goal in this section is two-fold: First to prove a localized version of such Hölder bounds; second, to accomplish this goal under low integrability assumptions on the tail term, leading to precise estimates that are suitable to subsequently derive potential estimates. We accomplish this by proving the desired Hölder estimates directly, without relying on the harmonic extension that was used in the previous works [CV10, NST23].

Indeed, for linear parabolic equations without drift, i.e. $b \equiv 0$, the Hölder regularity of weak solutions to (1.3) has been established in various previously mentioned articles by a direct approach using either a nonlocal adaptation of the De Giorgi- or the Moser iteration. The proof consists of two main parts: a local boundedness estimate and a proof of oscillation decay. In order to run the respective iteration arguments, one needs to test the weak formulation of (1.3) with suitable test-functions and derive corresponding Caccioppoli-type estimates. In our case, the main work consists in proving the energy estimates. Starting from those estimates, the derivation of the local boundedness and the oscillation decay goes by an adaptation of the techniques recently developed in [KW23a], [Lia22] (see also [BK23]).

Remark 2.1. The nonlocal $\log(u)$-lemma seems to fail for (1.3). Therefore, we need to follow a De Giorgi-type approach which does not make use of such estimate in order to deduce the oscillation decay. We follow the technique developed in [Lia22] for the fractional p-Laplacian.

2.1. **Energy estimates.** We have the following energy estimate for solutions to (1.3). Note that they do not differ from the respective estimates for solutions to (1.3) in case $b \equiv 0$. For now, we always assume that $b \in L^\infty(\mathbb{R}^{d+1})$ with $\|b\|_{L^\infty(\mathbb{R}^{d+1})} \leq \Lambda$ until specified otherwise.

Recall the notation

$$Q_r(t_0, x_0) = (t_0 - r^{2s}, t_0) \times B_r(x_0).$$

Moreover, we write

$$I_r^{\ominus}(t_0) := (t_0 - r^{2s}, t_0), \qquad I_r^{\oplus}(t_0) := (t_0, t_0 + r^{2s}).$$

In the following, we will drop t_0 from the notation and also write $B_r(x_0) = B_r$ and $Q_r(t_0, x_0) = Q_r$, whenever no confusion can arise.

Moreover, we write

$$(2.1) \qquad \mathcal{E}^{K_t}(u, v) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (u(x) - u(y))(v(x) - v(y)) K_t(x, y) \mathrm{d}y \mathrm{d}x$$

for the energy associated to $\mathcal{L}_t$, and denote

$$\mathcal{E}_{B_r}^s(u, v) = \int_{B_r} \int_{B_r} (u(x) - u(y))(v(x) - v(y)) |x - y|^{-d-2s} \mathrm{d}y \mathrm{d}x \qquad \forall r > 0.$$

Lemma 2.1. *Let v be a weak subsolution (supersolution) to (1.3) in Q_{2R}. Then, there exists a constant $c > 0$ depending only on d, s, Λ such that for any $l \in \mathbb{R}$ and any $0 < \rho_1 \leq r \leq r + \rho_1 \leq R$ and $0 < \rho_2 \leq r \leq r + \rho_2 \leq R$ such that $I_{r+\rho_1}^{\ominus} \times B_{r+\rho_2} \subset Q_{2R}$:*

$$(2.2)$$

$$\sup_{t \in I_r^{\ominus}} \int_{B_r} w_{\pm}^2(x, t) \mathrm{d}x + \int_{I_r^{\ominus}} \mathcal{E}_{B_r}^s(w_{\pm}(t), w_{\pm}(t)) \mathrm{d}t$$

$$\leq c[\rho_2^{-2s} \vee ((r + \rho_1)^{2s} - r^{2s})^{-1}] \int_{I_{r+\rho_1}^{\ominus}} \int_{B_{r+\rho_2}} w_{\pm}^2(t, x) \mathrm{d}x \mathrm{d}t$$

$$+ c \left(\frac{r + \rho_2}{\rho_2} \right)^d \rho_2^{-2s} \int_{I_{r+\rho_1}^{\ominus}} \left(\int_{B_{r+\rho_2}} w_{\pm}(t, x) \mathrm{d}x \right) \mathrm{Tail}(w_{\pm}(t); 0, r + \rho_2) \mathrm{d}t,$$

and

$$(2.3)$$

$$\sup_{t \in I_{r+\rho_1}^{\ominus}} \int_{B_r} w_{\pm}^2(x, t) \mathrm{d}x + \int_{I_{r+\rho_1}^{\ominus}} \mathcal{E}_{B_r}^s(w_{\pm}(t), w_{\pm}(t)) \mathrm{d}t$$

$$+ r^{-d-2s} \int_{I_{r+\rho_1}^{\ominus}} \int_{B_r} \int_{B_r} w_{\pm}(t, x) w_{\mp}(t, y) \mathrm{d}y \mathrm{d}x \mathrm{d}t$$

$$\leq \int_{B_{r+\rho_2}} w_{\pm}^2(t_0 - (r + \rho_1)^{2s}, x) \mathrm{d}x + c \rho_2^{-2s} \int_{I_{r+\rho_1}^{\ominus}} \int_{B_{r+\rho_2}} w_{\pm}^2(t, x) \mathrm{d}x \mathrm{d}t$$

$$+ c \left(\frac{r + \rho_2}{\rho_2} \right)^d \rho_2^{-2s} \int_{I_{r+\rho_1}^{\ominus}} \left(\int_{B_{r+\rho_2}} w_{\pm}(t, x) \mathrm{d}x \right) \mathrm{Tail}(w_{\pm}(t); 0, r + \rho_2) \mathrm{d}t,$$

where we denote $w = u - l$.

Remark 2.2. Note that in the following proof we will test the weak formulation of (1.3) with functions depending on u. While strictly speaking, a priori we only have $u \in L^2(I; H^s(\Omega))$, testing with the solution itself can be rigorously justified regardless. In fact, this can be achieved by perturbing the equation with an artificial diffusion term of the type $\varepsilon \Delta u$ and then passing to the limit as $\varepsilon \to 0$, see [CV10, Appendix C] for more details.

Proof. Let us explain how to prove the desired estimates for subsolutions. The proof of (2.2) and (2.3) is classical in case $b \equiv 0$ and goes by testing the weak formulation with the test function $\phi(t, x) = \eta^2(t)\tau^2(x)w_+(t, x)$ for some cut-off function $\tau \in C_c^\infty(B_{r+\frac{\rho_2}{2}})$ with $\tau \equiv 1$ in B_r, $0 \leq \tau \leq 1$, and $|\nabla \tau| \leq 4\rho_2^{-1}$, and $\eta \in C^\infty(\mathbb{R})$, which yields for any $t \in I_{r+\rho_1}^\ominus$:

(2.4)

$$
\int_{B_{r+\rho_2}} \eta^2(t)\tau^2 w_+^2(t)\mathrm{d}x + \int_{I_{r+\rho_1}^\ominus} \eta^2 \left(\int_{B_{r+\rho_2}} (b, \nabla[\tau^2 w_+])u\,\mathrm{d}x \right)\mathrm{d}t
$$
$$
+ \int_{I_{r+\rho_1}^\ominus} \eta^2 \mathcal{E}^{K_s}(u, \tau^2 w_+)\mathrm{d}t
$$
$$
\leq \eta^2(t_0 - (r + \rho_1)^{2s}) \int_{B_{r+\rho_2}} \tau^2 w_+^2(t_0 - (r + \rho_1)^{2s})\mathrm{d}x
$$
$$
+ 2\int_{I_{r+\rho_1}^\ominus} \eta|\eta'| \left(\int_{B_{r+\rho_2}} \tau^2 w_+^2\mathrm{d}x \right)\mathrm{d}t.
$$

Moreover, we recall the following estimate (see for instance [KW22b, Lemma 3.1])

(2.5)

$$
\int_{I_{r+\rho_1}^\ominus} \eta^2 \mathcal{E}_{B_{r+\rho_2}}^s(\tau w_+, \tau w_+)\mathrm{d}t - \int_{I_{r+\rho_1}^\ominus} \eta^2 \mathcal{E}_{B_{r+\rho_2}}^s(w_-, \tau w_+)\mathrm{d}t
$$
$$
\leq c\int_{I_{r+\rho_1}^\ominus} \eta^2 \mathcal{E}^{K_s}(u, \tau^2 w_+)\mathrm{d}t + c\rho_2^{-2s} \int_{I_{r+\rho_1}^\ominus} \eta^2 \left(\int_{B_{r+\rho_2}} w_+^2\mathrm{d}x \right)\mathrm{d}t
$$
$$
+ c\left(\frac{r + \rho_2}{\rho_2} \right)^d \rho_2^{-2s} \int_{I_{r+\rho_1}^\ominus} \eta^2 \left(\int_{B_{r+\rho_2}} w_+\mathrm{d}x \right)\mathrm{Tail}(w_+; 0, r + \rho_2)\mathrm{d}t,
$$

where we also used assumption (iii) on K_t. These two ingredients are sufficient to prove the energy estimates in case $b \equiv 0$. In order to treat the present case including a drift term, we observe

$$
(b, \nabla\phi)u = \eta^2 w_\pm^2(b, \nabla\tau^2) + \eta^2 l(b, \nabla(\tau^2)) + \frac{1}{2}\eta^2\tau^2(b, \nabla w_\pm^2),
$$

and that therefore, after integrating by parts and using that b is divergence free and bounded:

(2.6)
$$
\left| \int_{I_{r+\rho_1}^\ominus} \int_{B_{r+\rho_2}} (b, \nabla\phi)u\,\mathrm{d}x\mathrm{d}t \right| \leq \frac{3}{2}\int_{I_{r+\rho_1}^\ominus} \eta^2 \int_{B_{r+\rho_2}} |(b, \nabla\tau^2)|w_\pm^2\mathrm{d}x\mathrm{d}t
$$
$$
\leq c\rho_2^{-1} \int_{I_{r+\rho_1}^\ominus} \int_{B_{r+\rho_2}} w_\pm^2\mathrm{d}x\mathrm{d}t.
$$

The proof of (2.2) follows now by combining the estimates (2.4), (2.5), and (2.6), taking the supremum in $t \in I_{r+\rho_1}^\ominus$ upon choosing $\eta(t_0 - (r+\rho_1)^{2s}) = 0$, $\eta \equiv 1$ in $I_r^\ominus$, $|\eta'| \leq 4((r + \rho_1)^{2s} - r^{2s})^{-1}$, and $0 \leq \eta \leq 1$. To prove (2.3) we proceed in the exact same way, however we work with $\eta \equiv 1$. $\square$

2.2. **Local boundedness.** The following is the main result of this subsection. It states that weak solutions to (1.3) are locally bounded in their solution domain.

Lemma 2.2 (Local boundedness). *Let u be a weak solution to (1.3) in Q_{2R}. Then, there exists a constant $c > 0$, depending only on d, s, Λ, such that*

$$\sup_{Q_{R/2}} |u| \leq c \fint_{Q_R} |u(t,x)| \mathrm{d}x \mathrm{d}t + c \fint_{I_R^\ominus} \mathrm{Tail}(u(t); R) \mathrm{d}t.$$

The main ingredient in the proof is the following lemma:

Lemma 2.3. *Let u be a weak subsolution to (1.3) in Q_{2R}. Then, there exists a constant $c > 0$, depending only on d, s, Λ, such that*

$$\sup_{Q_{R/2}} u_+ \leq c \left(\fint_{Q_R} u_+^2(t,x) \mathrm{d}x \mathrm{d}t \right)^{1/2} + c \fint_{I_R^\ominus} \mathrm{Tail}(u_+(t); R) \mathrm{d}t.$$

Remark 2.3. In particular, the above lemma implies that weak solutions to (1.3) in Q are locally bounded, i.e. $u \in L_{loc}^\infty(Q)$.

Proof of Lemma 2.3. Having at hand the energy estimate (2.2), which is exactly the same as in case $b \equiv 0$, the proof follows by the same arguments as in [KW23a, Theorem 1.8]. Let us give a brief sketch: First of all, we observe that with the help of (2.2), we can derive the following estimates for weak subsolutions v to (1.3):

(2.7)
$$\sup_{Q_{R/2}} v_+ \leq c \left(\fint_{Q_R} v_+^2(t,x) \mathrm{d}x \mathrm{d}t \right)^{1/2} + c \sup_{t \in I_R^\ominus} \mathrm{Tail}(v_+(t); R),$$

(2.8)
$$\sup_{t \in I_{R/2}^\ominus} \left(\fint_{B_{R/2}} v_+^2(t,x) \mathrm{d}x \right)^{1/2} \leq \left(\fint_{Q_R} v_+^2(t,x) \mathrm{d}x \mathrm{d}t \right)^{1/2} + c \fint_{I_R^\ominus} \mathrm{Tail}(v_+(t); R) \mathrm{d}t.$$

The first estimate can be achieved by a De Giorgi iteration scheme, based on (2.2) (see [KW22b, Theorem 3.6]), without relying anymore on the equation. The second estimate is a direct consequence of (2.2) applied with $l = 0$ and an interpolation argument. It can be found in [KW23a, Lemma 3.1]. Moreover, note that if w is a subsolution to $\partial_t w + (b, \nabla w) + \mathcal{L}_t w \leq f$, for some $f \in L_{t,x}^{1,\infty}$, then $v(t,x) = w(t,x) - \int_{t_0}^t \|f_+(s)\|_{L^\infty(B_R)} \mathrm{d}s$ is a subsolution to (1.3). Therefore, we can deduce from (2.7) that:

(2.9)
$$\sup_{Q_{R/2}} w_+ \leq c \left(\fint_{Q_R} w_+^2(t,x) \mathrm{d}x \mathrm{d}t \right)^{1/2} + c \sup_{t \in I_R^\ominus} \mathrm{Tail}(w_+(t); R) + c \|f_+\|_{L_t^1 L_x^\infty(Q_R)}.$$

To prove the desired result, we will now take $\eta \in C_c^\infty(B_{3R/2})$ with $\eta = 1$ in B_R and observe that

$$\partial_t(\eta u) + (b, \nabla(\eta u)) + \mathcal{L}_t(\eta u) \leq -(b, \nabla((1-\eta)u)) - \mathcal{L}_t((1-\eta)u) =: f \quad \text{in } Q_{2R}.$$

Moreover, for $t \in I_{2R}^{\ominus}$ and $x \in B_{R/2}$ it holds $(b, \nabla((1 - \eta)u)) = 0$ and therefore

$$\|f_+\|_{L_{t,x}^{1,\infty}(Q_{R/2})} = \|[\mathcal{L}_t((1 - \eta)u)]_+\|_{L_{t,x}^{1,\infty}(Q_{R/2})} \leq \fint_{I_{R/2}^{\ominus}} \mathrm{Tail}(u_+(t); R)\mathrm{d}t.$$

Thus, by application of (2.9) to ηu, and estimating $\sup_{t \in I_R^{\ominus}} \mathrm{Tail}((\eta u)_+(t); R)$ with the help of (2.8), we obtain the desired result. $\qquad\square$

Proof of Lemma 2.2. An application of Lemma 2.3 yields

$$\sup_{Q_{R/2}} |u| \leq c \left(\fint_{Q_R} |u(t, x)|^2 \mathrm{d}x\mathrm{d}t \right)^{1/2} + c \fint_{I_R^{\ominus}} \mathrm{Tail}(u(t); R)\mathrm{d}t.$$

The exponent in the first term on the right hand side can be lowered by a standard covering and iteration argument (see [KW22b, Theorem 6.2]). $\quad\square$

2.3. Hölder regularity estimates.

The main goal of this subsection is to prove Theorem 1.1. We establish the oscillation decay by using a nonlocal modification of the parabolic De Giorgi iteration, as it has been carried out for instance in [Lia22] under the additional assumption that the tails are bounded in time (see also [APT22], [BK23]). As it was the case for the local boundedness, this proof also goes as in case $b \equiv 0$, since the equation will only be used through the energy estimates (2.2) and (2.3) from now on. However, some of the arguments need to be modified in order to deal with tails that are L^q in time for some $q \in (1, \infty)$. Therefore, we sketch some parts of the proof.

Lemma 2.4. *Let u be a weak supersolution to (1.3) in Q_{2R} and assume that $u \geq 0$ in Q_{2R}. Then, for any $\delta \in (0, 1)$ and $H > 0$ there exists $\nu \in (0, 1)$, depending only on d, s, Λ, q, such that the following holds true: If*

$$\left|\{u \leq H\} \cap I_{\delta R}^{\ominus} \times B_R\right| \leq \nu |I_{\delta R}^{\ominus} \times B_R|,$$

and

$$(2.10) \qquad \left(\fint_{I_{2R}^{\ominus}} \mathrm{Tail}(u_-(t); 2R)^q \mathrm{d}t \right)^{1/q} \leq \delta^{\frac{2s}{q}} H,$$

then it holds:

$$u \geq \frac{H}{2} \quad in \ I_{\delta R/2}^{\ominus} \times B_{R/2}.$$

Proof. The proof goes by a standard De Giorgi iteration argument. We treat the L^q-tails by a standard interpolation argument (see [DiB93]). We define

$$k_i = \frac{H}{2} + \frac{H}{2^i}, \ \hat{k}_i = \frac{k_i + k_{i+1}}{2}, \ w_i = (u - \hat{k}_i)_+, \ R_i = \frac{R}{2} + \frac{R}{2^i}, \ \hat{R}_i = \frac{R_i + R_{i+1}}{2}$$

$$A_i = \frac{\left|\{u \leq k_i\} \cap I_{\delta R_i}^{\ominus} \times B_{R_i}\right|}{|I_{\delta R_i}^{\ominus} \times B_{R_i}|}, \quad B_i = \left(\fint_{I_{\delta R_i}^{\ominus}} \left(\frac{|\{u(t) \leq k_i\} \cap B_{R_i}|}{|B_{R_i}|} \right)^{\frac{q}{q-1}} \mathrm{d}t \right)^{\frac{q-1}{q(1+\kappa)}},$$

where $\kappa := \frac{2s}{d}\frac{q-1}{q}$, and deduce from (2.2):

(2.11)

$$
\sup_{t \in I^{\ominus}_{\delta R_{i+1}}} \int_{B_{R_{i+1}}} w_i^2(x,t)\mathrm{d}x + \int_{I^{\ominus}_{R_{i+1}}} \mathcal{E}^s_{B_{R_{i+1}}}(w_i(t), w_i(t))\mathrm{d}t
$$

$$
\leq c2^{i2s}(\delta R)^{-2s} \int_{I^{\ominus}_{\delta \hat{R}_i}} \int_{B_{\hat{R}_i}} w_i^2(t,x)\mathrm{d}x\mathrm{d}t
$$

$$
+ c2^{i(d+2s)}\hat{R}_i^{-2s} \int_{I^{\ominus}_{\delta \hat{R}_i}} \int_{B_{\hat{R}_i}} w_i(t,x)\,\mathrm{Tail}(w_i(t); 0, \hat{R}_i)\mathrm{d}t
$$

$$
\leq c(2^{i2s} + 2^{id})R^{-2s}H^2 A_i|Q_{R_i}|
$$

$$
+ c2^{i(d+2s)}R^{-2s}H \int_{I^{\ominus}_{\delta \hat{R}_i}} \mathrm{Tail}(w_i(t); 0, 2R) \left(\int_{B_{\hat{R}_i}} \mathbb{1}_{\{u \leq k_i\}}(t,x)\mathrm{d}x \right)\mathrm{d}t
$$

$$
\leq c2^{i(d+2s)}R^{-2s}H^2 A_i|Q_{R_i}|
$$

$$
+ c2^{i(d+2s)}R^{-2s}\delta^{\frac{2s}{q}} \left(\fint_{I^{\ominus}_{2R}} \mathrm{Tail}(u_-(t); 2R)^q\mathrm{d}t \right)^{1/q} B_i^{1+\kappa}|Q_{R_i}|
$$

$$
\leq c2^{i(d+2s)}R^{-2s}H^2|Q_{R_i}| \left(A_i + B_i^{1+\kappa} \right),
$$

where we applied Hölder's inequality and used (2.10). Moreover, recall the following estimate, which follows from the fractional Sobolev inequality and Hölder interpolation:

(2.12)

$$
A_{i+1}|I^{\ominus}_{\delta R_{i+1}} \times B_{R_{i+1}}| \leq c2^{2i}H^{-2}(A_i|I^{\ominus}_{\delta R_i} \times B_{R_i}|)^{\frac{2s}{d+2s}} \left[\sup_{t \in I^{\ominus}_{R_{i+1}}} \int_{B_{R_{i+1}}} w_i^2(x,t)\mathrm{d}x \right.
$$

$$
\left. + \int_{I^{\ominus}_{R_{i+1}}} \mathcal{E}^s_{B_{R_{i+1}}}(w_i(t), w_i(t))\mathrm{d}t + 2^{i2s}R^{-2s} \int_{I^{\ominus}_{R_{i+1}}} \int_{B_{R_{i+1}}} w_i^2(t,x)\mathrm{d}x\mathrm{d}t \right].
$$

A combination of (2.11) and (2.12) yields for some $\gamma_1 > 1$ and $c_1 > 0$:

$$
(2.13) \qquad A_{i+1} \leq \frac{c_1}{\delta^{2s}} 2^{\gamma_1 i}\delta^{2s\iota} \left(A_i^{1+\frac{2s}{d+2s}} + A_i^{\frac{2s}{d+2s}}B_i^{1+\kappa} \right),
$$

where $\iota := \frac{2s}{d+2s} \wedge 1 - \frac{q-1}{q(1+\kappa)}$. Similar to the proof of (2.12), one obtains

(2.14)

$$
B_{i+1} \leq c2^{(d+2)i}\delta^{-\frac{2s(q-1)}{q(1+\kappa)}}R^{-d}H^{-2} \left[\sup_{t \in I^{\ominus}_{R_{i+1}}} \int_{B_{R_{i+1}}} w_i^2(x,t)\mathrm{d}x \right.
$$

$$
\left. + \int_{I^{\ominus}_{R_{i+1}}} \mathcal{E}^s_{B_{R_{i+1}}}(w_i(t), w_i(t))\mathrm{d}t + 2^{i2s}R^{-2s} \int_{I^{\ominus}_{R_{i+1}}} \int_{B_{R_{i+1}}} w_i^2(t,x)\mathrm{d}x\mathrm{d}t \right],
$$

which yields the following estimate after combination with (2.11) for some $\gamma_2 > 1$ and $c_2 > 0$:

$$
(2.15) \qquad B_{i+1} \leq \frac{c_2}{\delta^{2s}} 2^{\gamma_2 i}\delta^{2s\iota} \left(A_i + B_i^{1+\kappa} \right).
$$

Finally, note that by assumption, we have $A_0 \leq \nu < 1$, and therefore

$$A_0 + B_0^{1+\kappa} \leq A_0 + A_0^{\frac{q-1}{q}} \leq 2A_0^{\frac{q-1}{q}} \leq 2\nu^{\frac{q-1}{q}}.$$

Thus, the desired result follows by an application of a classical iteration lemma (see [DiB93, Chapter 1, Lemma 4.2]), upon choosing $\nu > 0$ small enough, depending only on $d, s, q, c_1, c_2, \gamma_1, \gamma_2$. $\qquad\square$

Lemma 2.5. *Let u be a weak supersolution to (1.3) in Q_{2R} and assume that $u \geq 0$ in Q_{2R}. Let $\alpha \in (0, 1]$. Then, there exist $\varepsilon, \delta \in (0, 1)$, depending only on d, s, Λ, q, α, such that for any $H > 0$ the following holds true: If*

$$(2.16) \qquad |\{u(t_0, \cdot) \geq H\} \cap B_R| \geq \alpha|B_R|$$

for some $t_0 \in I_{2R}^\ominus$, then either

$$(2.17) \qquad \left(\fint_{I_{2R}^\ominus} \mathrm{Tail}(u_-(t); 2R)^q \mathrm{d}t \right)^{1/q} > \delta^{\frac{2s}{q}} H,$$

or for all $t \in I_{\delta R}^\oplus(t_0)$ with $I_{\delta R}^\oplus(t_0) \subset I_{2R}^\ominus$:

$$|\{u(t, \cdot) \geq \varepsilon H\} \cap B_R| \geq \frac{\alpha}{2}|B_R|.$$

Proof. Let us assume that (2.17) does not hold true. Let $\sigma \in (0, 1)$. We apply (2.3) to obtain for any $t \in I_{\delta R}^\oplus(t_0)$

$$
\begin{aligned}
\int_{B_{(1-\sigma)R}} (u - H)_-^2(t, x)\mathrm{d}x &\leq \int_{B_R} (u - H)_-^2(t_0, x)\mathrm{d}x \\
&\quad + c(\sigma R)^{-2s} \int_{I_{\delta R}^\oplus(t_0)} \int_{B_R} (u - H)_-^2 \mathrm{d}x\mathrm{d}\tau \\
&\quad + c\sigma^{-d-2s} R^{-2s} \int_{I_{\delta R}^\oplus(t_0)} \int_{B_R} (u - H)_- \mathrm{Tail}((u - H)_-; R)\mathrm{d}x\mathrm{d}\tau \\
&\leq H^2|B_R| \left((1 - \alpha) + c\sigma^{-2s}\delta^{2s} + c\sigma^{-d-2s}\delta^{2s} \right),
\end{aligned}
$$

where we also used (2.16), and applied similar arguments to derive the second estimate, as in the previous proofs. From here, the proof is standard (see [Lia22, Lemma 3.3]), namely, for any $\varepsilon \in (0, 1)$, we have

$$
\begin{aligned}
|\{u(t, \cdot) \leq \varepsilon H\} \cap B_R| &\leq |\{u(t, \cdot) \leq \varepsilon H\} \cap B_{(1-\sigma)R}| + |B_R \setminus B_{(1-\sigma)R}| \\
&\leq ((1 - \varepsilon)H)^{-2} \int_{B_{(1-\sigma)}} (u - H)_-^2(t, x)\mathrm{d}x + d\sigma|B_R| \\
&\leq (1 - \varepsilon)^{-2} \left((1 - \alpha) + c\sigma^{-2s}\delta^{2s} + c\sigma^{-d-2s}\delta^{2s} + d\sigma \right) |B_R|,
\end{aligned}
$$

which yields the desired result upon choosing $\sigma, \varepsilon, \delta$ suitably, depending only on d, s, c, α. $\qquad\square$

Lemma 2.6. *Let u be a weak supersolution to (1.3) in Q_{2R} and assume that $u \geq 0$ in Q_{2R}. Let $\delta, \sigma \in (0, 1]$ and $H > 0$. Then, the following holds true: If*

$$(2.18) \qquad |\{u(t, \cdot) \geq H\} \cap B_R| \geq \alpha|B_R| \,\, \forall t \in I_{\delta R}^\ominus$$

for some $\alpha \in (0,1)$, then either

$$(2.19) \qquad \left(\fint_{I_{2R}^{\ominus}} \mathrm{Tail}(u_-(t);2R)^q \mathrm{d}t \right)^{1/q} > \delta^{\frac{2s}{q}} H,$$

or there exists a constant $c > 0$, depending only on d, s, Λ, q, such that

$$\left| \left\{ u \le \frac{\sigma H}{4} \right\} \cap I_{\delta R}^{\ominus} \times B_R \right| \le c \frac{\sigma}{\delta^{2s}\alpha} |I_{\delta R}^{\ominus} \times B_R|.$$

Proof. Let us assume that (2.19) does not hold true. We define $w = u - \frac{\sigma H}{2}$ and apply (2.3), which yields:

$$R^{-d-2s} \int_{I_{\delta R}^{\ominus}} \int_{B_R} \int_{B_R} w_-(t,x)w_+(t,y)\mathrm{d}y\mathrm{d}x\mathrm{d}t \le \int_{B_{2R}} w_-^2(t_0 - R^{2s}, x)\mathrm{d}x$$

$$+ cR^{-2s} \int_{I_{\delta R}^{\ominus}} \int_{B_{2R}} w_-^2 \mathrm{d}x\mathrm{d}t + cR^{-2s} \int_{I_{\delta R}^{\ominus}} \left(\int_{B_{2R}} w_-(t,x)\mathrm{d}x \right) \mathrm{Tail}(w_-;R)\mathrm{d}t$$

$$\le c(\sigma H)^2 (\delta R)^{-2s} |I_{\delta R}^{\ominus} \times B_R| + c(\sigma H)\delta^{-2s} R^d \fint_{I_R^{\ominus}} \mathrm{Tail}(w_-;2R)\mathrm{d}t$$

$$\le c(\sigma H)^2 (\delta R)^{-2s} |I_{\delta R}^{\ominus} \times B_R|$$

Here, we used the same arguments as in the proof of the previous lemmas to obtain the second estimate. Moreover, we used Hölder's inequality to estimate the tail term. Finally, by (2.18), we conclude

$$\left| \left\{ u \le \frac{\sigma H}{4} \right\} \cap I_{\delta R}^{\ominus} \times B_R \right| \le c(\sigma H)^{-1} \int_{I_{\delta R}^{\ominus}} \int_{B_R} w_-(t,x)\mathrm{d}x\mathrm{d}t$$

$$\le c(\sigma\alpha)^{-1} H^{-2} R^{-d} \int_{I_{\delta R}^{\ominus}} \int_{B_R} \int_{B_R} w_-(t,x)w_+(t,y)\mathrm{d}y\mathrm{d}x\mathrm{d}t$$

$$\le c \frac{\sigma}{\delta^{2s}\alpha} |I_{\delta R}^{\ominus} \times B_R|,$$

as desired. $\qquad \square$

Let us state the following growth lemma, which is an immediate consequence of Lemmas 2.4, 2.5, and 2.6:

Corollary 2.1 (Growth lemma). *Let u be a weak supersolution to (1.3) in Q_{2R} and assume that $u \ge 0$ in Q_{2R}. Let $\alpha \in (0,1]$ and $H > 0$. Then, the following holds true: If for some $t_0 \in I_R^{\ominus}$*

$$(2.20) \qquad |\{u \ge H\} \cap B_R| \ge \alpha|B_R|,$$

then there exist $\delta, \theta \in (0,1)$, depending only on d, s, Λ, q, α such that if $I_{\delta R}^{\oplus}(t_0) \subset Q_{2R}$, and

$$(2.21) \qquad \left(\fint_{I_{2R}^{\ominus}} \mathrm{Tail}(u_-(t);2R)^q \mathrm{d}t \right)^{1/q} \le \delta^{\frac{2s}{q}} \theta H,$$

then

$$u \ge \frac{\theta H}{8} \quad in \ I_{\delta R/2}^{\ominus}(t_0 + (\delta R)^{2s}) \times B_{R/2}.$$

Proof. Let $\xi, \alpha \in (0, 1]$. We choose $\delta, \varepsilon \in (0, 1)$ as in Lemma 2.5 and obtain for all $t \in I_{\delta R}^{\oplus}(t_0)$

$$|\{u(t, \cdot) \geq \varepsilon H\} \cap B_R| \geq \frac{\alpha}{2}|B_R|.$$

Next, we can apply Lemma 2.6 and obtain that for $\sigma = C^{-1}\nu\delta\alpha$ (where ν denotes the constant from Lemma 2.4) and $C > 0$ such that $c\frac{\sigma}{\delta^{2s}\alpha} \leq \nu$ (where $c > 0$ denotes the constant from 2.6)

$$\left|\left\{u \leq \frac{\sigma\varepsilon H}{4}\right\} \cap I_{\delta R}^{\oplus}(t_0) \times B_R\right| \leq c\frac{\sigma}{\delta^{2s}\alpha}|I_{\delta R}^{\oplus}(t_0) \times B_R| \leq \nu|I_{\delta R}^{\oplus}(t_0) \times B_R|,$$

Finally, we apply Lemma 2.4, which yields

$$u \geq \frac{\sigma\varepsilon H}{8} \quad \text{in } Q_{R/2} \quad \text{in } I_{\delta R/2}^{\ominus}(t_0 + (\delta R)^{2s}) \times B_{R/2},$$

as desired. $\qquad\square$

We are now ready to prove the Hölder regularity estimate:

Proof of Theorem 1.1. Having Corollary 2.1 and Lemma 2.2 at our disposal, the proof is standard (see for instance [KW23a, Lia22]): The proof goes by constructing sequences (M_j) and (m_j) that are non-decreasing and non-increasing, respectively, and to find a small $\gamma \in (0, 1)$ and a large $\nu > 1$ such that for every $j \in \mathbb{N}$:

(2.22) $\qquad m_j \leq u \leq M_j \quad \text{in } Q_{\nu^{-j}R}, \quad \text{and} \quad M_j - m_j = L\nu^{-\gamma j},$

where

$$L := C_0\|u\|_{L^\infty(Q_R)} + \left(\fint_{I_{2R}^{\ominus}} \text{Tail}(u(t); 2R)^q \mathrm{d}t\right)^{\frac{1}{q}}$$

for some constant C_0 to be determined later. Once this is achieved, we obtain

$$|u(t, x) - u(s, y)| \leq R^{-\gamma}L(|x - y| + |t - s|^{1/(2s)})^\gamma,$$

from which we conclude the desired result upon recalling Lemma 2.2.

We choose ν such that $Q_{\nu^{-1}R} \subset I_{\delta R/2}^{\ominus} \times B_{\delta R/2}$, where $\delta \in (0, 1]$ denotes the constant in Corollary 2.1 corresponding to $\alpha = 1/2$. To prove (2.22), we observe that upon setting $M_j = \nu^{-\gamma j}L/2$ and $m_j = -\nu^{-\gamma j}L/2$, (2.22) holds true for every $j \leq j_0$ if we choose $C_0 \geq 2\nu^{\gamma j_0}$, where j_0 can be chosen arbitrarily. To prove (2.22) for $j > j_0$, we proceed by induction. Let us assume that (2.22) holds true for some $j \geq j_0$, and assume without loss of generality that

(2.23) $\qquad |\{u(t_0-, \cdot) \geq m_j + (M_j - m_j)/2\} \cap B_{\nu^{-j}R}| \geq \frac{1}{2}|B_{\nu^{-j}R}|.$

We apply the growth lemma Corollary 2.1 to $v = u - m_j$, and deduce that $u \geq M_j + \frac{\theta L\nu^{-\gamma j}}{16}$ in $Q_{\nu^{-(j+1)}R}$. This yields (2.22) for $j + 1$ after defining $M_{j+1} = M_j$ and $m_{j+1} = m_j + \frac{\theta L\nu^{-\gamma j}}{16}$. Note that in case (2.23) fails, we proceed in the same way but upon defining $v = M_j - u$. Note that the verification of the assumptions of Corollary 2.1 is standard and goes exactly as in [KW23a], relying on appropriate choices of the parameters γ, ν, j_0. $\qquad\square$

2.4. **The case of BMO drifts.** Let us explain the necessary modifications in the aforementioned proofs which allow us to prove Hölder regularity estimates (see Proposition 1.1) for solutions to (1.1) with $b \in L^\infty(\mathbb{R}; \mathrm{BMO}(\mathbb{R}^d))$.

We state an adapted version of the energy estimate (see Lemma 2.1).

Lemma 2.7. *Let v be a weak subsolution (supersolution) to (1.3) in Q_{2R}. Then, in the same setting as in Lemma 2.1, if $s > 1/2$, the estimate (2.2) and (2.3) remain true. If $s = 1/2$, then*

(2.24)
$$\sup_{t \in I_r^\ominus} \int_{B_r} w_\pm^2(x,t)\mathrm{d}x + \int_{I_r^\ominus} \mathcal{E}_{B_r}^{1/2}(w_\pm(t), w_\pm(t))\mathrm{d}t$$

$$\leq c[\rho_2^{-1} \vee ((r + \rho_1) - r)^{-1} \vee |\log_2(r + \rho_2)|(r + \rho_2)\rho_2^{-2}] \int_{I_{r+\rho_1}^\ominus} \int_{B_{r+\rho_2}} w_\pm^2(t,x)\mathrm{d}x\mathrm{d}t$$

$$+ c\left(\frac{r + \rho_2}{\rho_2}\right)^d \rho_2^{-1} \int_{I_{r+\rho_1}^\ominus} \left(\int_{B_{r+\rho_2}} w_\pm(t,x)\mathrm{d}x\right) \mathrm{Tail}(w_\pm(t); 0, r + \rho_2)\mathrm{d}t,$$

and

(2.25)
$$\sup_{t \in I_{r+\rho_1}^\ominus} \int_{B_r} w_\pm^2(x,t)\mathrm{d}x + \int_{I_{r+\rho_1}^\ominus} \mathcal{E}_{B_r}^{1/2}(w_\pm(t), w_\pm(t))\mathrm{d}t$$

$$+ r^{-d-1} \int_{I_{r+\rho_1}^\ominus} \int_{B_r} \int_{B_r} w_\pm(t,x)w_\mp(t,y)\mathrm{d}y\mathrm{d}x\mathrm{d}t$$

$$\leq \int_{B_{r+\rho_2}} w_\pm^2(t_0 - (r + \rho_1), x)\mathrm{d}x$$

$$+ c\left[\rho_2^{-1} \vee |\log_2(r + \rho_2)|(r + \rho_2)\rho_2^{-2}\right] \int_{I_{r+\rho_1}^\ominus} \int_{B_{r+\rho_2}} w_\pm^2(t,x)\mathrm{d}x\mathrm{d}t$$

$$+ c\left(\frac{r + \rho_2}{\rho_2}\right)^d \rho_2^{-1} \int_{I_{r+\rho_1}^\ominus} \left(\int_{B_{r+\rho_2}} w_\pm(t,x)\mathrm{d}x\right) \mathrm{Tail}(w_\pm(t); 0, r + \rho_2)\mathrm{d}t,$$

where $c > 0$ also depends on C_1, C_2.

Proof. We only need to explain how to modify (2.6). Note that (1.10) implies

(2.26)
$$\sup_{t,x_0} \left(\fint_{B_R(x_0)} |b(t,x)|^{\frac{d}{s}}\mathrm{d}x\right)^{\frac{s}{d}} \leq -\log_2(R)C,$$

since we have the following computation, based on John-Nirenberg lemma:

$$\left(\fint_{B_R(x_0)} |b(t,x)|^{\frac{d}{s}}\mathrm{d}x\right)^{\frac{s}{d}} \leq \left(\fint_{B_R(x_0)} |b(t,x) - (b(t))_{B_R(x_0)}|^{\frac{d}{s}}\mathrm{d}x\right)^{\frac{s}{d}} + (b(t))_{B_R(x_0)}$$

$$\leq c(1 - \log_2(R))C_1 + C_2.$$

We compute using Hölder– and Sobolev inequality, and $\frac{d+2s}{2d} = \frac{s}{d} + \frac{1}{2}$:

$$\frac{3}{2} \int_{I_{r+\rho_1}^\ominus} \eta^2 \int_{B_{r+\rho_2}} |(b, \nabla\tau^2)|w_\pm^2\mathrm{d}x\mathrm{d}t$$

$$\leq \varepsilon \int_{I^{\ominus}_{r+\rho_1}} \eta^2 \left(\int_{B_{r+\rho_2}} (\tau w_{\pm})^{\frac{2d}{d-2s}} \mathrm{d}x \right)^{\frac{d-2s}{d}} \mathrm{d}t$$

$$+ c(\varepsilon) \int_{I^{\ominus}_{r+\rho_1}} \eta^2 \left(\int_{B_{r+\rho_2}} [(b, \nabla\tau)w_{\pm}]^{\frac{2d}{d+2s}} \mathrm{d}x \right)^{\frac{d+2s}{d}} \mathrm{d}t$$

$$\leq c\varepsilon \int_{I^{\ominus}_{r+\rho_1}} \eta^2 \mathcal{E}^s(\tau w_{\pm}, \tau w_{\pm}) \mathrm{d}t$$

$$+ c(\varepsilon) \int_{I^{\ominus}_{r+\rho_1}} \eta^2 \left(\int_{B_{r+\rho_2}} b^{\frac{d}{s}} \mathrm{d}x \right)^{\frac{2s}{d}} \left(\int_{B_{r+\rho_2}} |\nabla\tau|^2 w_{\pm}^2 \mathrm{d}x \right) \mathrm{d}t.$$

Now, the first term can be absorbed and the second term can be estimated from above.

Indeed, if $s > 1/2$, then using (2.26), we obtain:

$$c(\varepsilon) \int_{I^{\ominus}_{r+\rho_1}} \eta^2 \left(\int_{B_{r+\rho_2}} b^{\frac{d}{s}} \mathrm{d}x \right)^{\frac{2s}{d}} \left(\int_{B_{r+\rho_2}} |\nabla\tau|^2 w_{\pm}^2 \mathrm{d}x \right) \mathrm{d}t$$

$$\leq c(\varepsilon) \int_{I^{\ominus}_{r+\rho_1}} \eta^2 \left(\sup_{x_0 \in B_{r+\rho_2} \setminus B_r} \int_{B_{\rho_2}(x_0)} b^{\frac{d}{s}} \mathrm{d}x \right)^{\frac{2s}{d}} \left(\int_{B_{r+\rho_2}} |\nabla\tau|^2 w_{\pm}^2 \mathrm{d}x \right) \mathrm{d}t$$

$$\leq c(\varepsilon) \int_{I^{\ominus}_{r+\rho_1}} (-\log_2(\rho_2))\rho_2^{2s} \left(\int_{B_{r+\rho_2}} |\nabla\tau|^2 w_{\pm}^2 \mathrm{d}x \right) \mathrm{d}t$$

$$\leq -c(\varepsilon, C)\rho_2^{-2s} \int_{I^{\ominus}_{r+\rho_1}} \int_{B_{r+\rho_2}} w_{\pm}^2 \mathrm{d}x\mathrm{d}t,$$

where we used that $-\log_2(\rho_2)\rho_2^{2s-2} \leq c\rho_2^{-2s}$. If $s = 1/2$, then we estimate the second term as follows:

$$c(\varepsilon) \int_{I^{\ominus}_{r+\rho_1}} \eta^2 \left(\int_{B_{r+\rho_2}} b^{\frac{d}{s}} \mathrm{d}x \right)^{\frac{2s}{d}} \left(\int_{B_{r+\rho_2}} |\nabla\tau|^2 w_{\pm}^2 \mathrm{d}x \right) \mathrm{d}t$$

$$\leq c(\varepsilon, C) \int_{I^{\ominus}_{r+\rho_1}} (-\log_2(r+\rho_2))(r+\rho_2) \left(\int_{B_{r+\rho_2}} |\nabla\tau|^2 w_{\pm}^2 \mathrm{d}x \right) \mathrm{d}t$$

$$\leq -c(\varepsilon, C) \log_2(r+\rho_2)(r+\rho_2)\rho_2^{-2} \int_{I^{\ominus}_{r+\rho_1}} \int_{B_{r+\rho_2}} w_{\pm}^2 \mathrm{d}x\mathrm{d}t.$$

This observation concludes the proof. $\qquad\square$

Clearly, since Lemma 2.1 remains true for BMO-drifts if $s > 1/2$, the proof of the Hölder estimate (1.11) in this setting follows in the exact same way as described before. Thus, in the following, we restrict ourselves to proving Proposition (1.11) in the critical case $s = 1/2$.

2.4.1. *Local boundedness in case $s = 1/2$ with slanted cylinders.* Note that the pre-factors in (2.24) and (2.25) are different from the case of bounded drifts. When following the proof of the local boundedness estimate using

De Giorgi iteration, and tracking the pre-factor, one easily sees that the following estimate holds true on scale one:

$$(2.27) \qquad \sup_{Q_{1/2}} |u| \le c \fint_{Q_1} |u(t,x)| \mathrm{d}x \mathrm{d}t + c \fint_{I_1^\ominus} \mathrm{Tail}(u(t);1)\mathrm{d}t,$$

where $c > 0$ depends on d, Λ, C_1, C_2.

To get a local boundedness estimate on scale $R \in (0,1)$, we will make use of the following rescaling, which leaves the mean value of the drift invariant (see [NST23, page 1677-1678], [CV10]):

Lemma 2.8. *Let $s = 1/2$ and assume* (1.10)*. Let $r \in (0,1)$, and $z_r(t)$ be the solution to the ODE*

$$z_r'(t) = \fint_{B_1} b(rt, rx + rz_r(t))\mathrm{d}x, \quad z_r(0) = 0.$$

Then, there exists a constant $c > 0$, depending only on d such that

$$(2.28) \qquad \|z_r\|_{C^1(I_1)} \le c(C_1 + C_2|\log(r)|) =: \theta - 1.$$

Moreover, let u be a solution to (1.1) *in $Q_{\theta r}$ for some $r \in (0,1)$. Then, $u_r(t,x) := u(rt, rx + rz_r(t))$ satisfies*

$$\partial_t u_r + (b_r, \nabla u_r) + \mathcal{L}_r u_r = 0 \quad \text{in } Q_1,$$

where $\mathcal{L}_r$ is an operator of the form (1.2) *with kernel K_r given by*

$$K_r(x,y) := r^{d+1} K_t(rx - rz_r(t), ry - rz_r(t)),$$

satisfying the conditions (i), (ii), (iii). Moreover, the drift b_r is defined as $b_r(t,x) := b(rt, rx + rz_r(t)) - z_r'(t)$ and satisfies

$$\mathrm{div}(b_r) = 0, \quad \fint_{B_1} b_r(t,x)\mathrm{d}x = 0, \quad \|b_r\|_{L^\infty(\mathbb{R},\mathrm{BMO}(\mathbb{R}^d))} = \|b(rt,x)\|_{L^\infty(\mathbb{R},\mathrm{BMO}(\mathbb{R}^d))},$$

$$\|b_r\|_{L^\infty(I_1;L^q(B_1))} \le cC_2,$$

for any $q \ge 0$, where $c > 0$ depends only on d, q.

Moreover, given $t_0 \in \mathbb{R}$, $x_0 \in \mathbb{R}^d$ and $R > 0$, we recall the definition of the slanted cylinders $\tilde{Q}_R(t_0, x_0) = I_R^\ominus(t_0) \times \tilde{B}_{t,R}(x_0)$, which are given as follows

$$\tilde{Q}_R(t_0, x_0) = \{(t, x + Rz_R(t/R)) : (t,x) \in Q_R(t_0, x_0)\}.$$

By combination of the previous lemma with (2.27), we get the following local boundedness estimate on small scales in slanted cylinders:

Lemma 2.9. *Let $s = 1/2$ and assume* (1.10)*. Let u be a weak solution to* (1.3) *in $Q_{\theta R}$, where $\theta - 1 = c(C_1 + C_2|\log(R)|)$, as in Lemma* 2.8*. Then, there exists a constant $c > 0$, depending only on d, s, Λ, such that*

$$(2.29) \qquad \sup_{\tilde{Q}_R} u \le c \fint_{\tilde{Q}_{2R}} |u(t,x)| \mathrm{d}x \mathrm{d}t + c \fint_{I_2^\ominus} \widetilde{\mathrm{Tail}}(u(t);2)\mathrm{d}t.$$

Proof. Defining u_R as in 2.8 and applying (2.27) to u_R yields

$$\sup_{(t,x)\in Q_R} |u(t, x + Rz_R(t/R))| = \sup_{(t,x)\in Q_1} u_R$$

$$\le c \fint_{Q_2} |u_R(t,x)| \mathrm{d}x \mathrm{d}t + c \fint_{I_2^\ominus} \mathrm{Tail}(u_R(t);2)\mathrm{d}t$$

$$\leq c \fint_{Q_{2R}} |u(t, x + Rz_R(t/R))| \mathrm{d}x\mathrm{d}t + c \fint_{I_R^\ominus} \mathrm{Tail}(u(t, \cdot + Rz_R(t/R)); R)\mathrm{d}t$$

for some constant $c > 0$, depending only on d, Λ, C_1, C_2. This implies the desired result. $\square$

2.4.2. *Hölder regularity in case $s = 1/2$ with slanted cylinders.* Moreover, note that the preliminary growth lemmas from the previous section (see Lemma 2.4, Lemma 2.5, and Lemma 2.6) remain true for $R = 1$, and as a consequence, also the following counterpart of Corollary 2.1 holds true on scale one:

Corollary 2.2 (Growth lemma). *Let $s = 1/2$ and assume (1.10). Let u be a weak supersolution to (1.3) in Q_2 and assume that $u \geq 0$ in Q_2. Let $\alpha \in (0, 1]$ and $H > 0$. Then, the following holds true: If for some $t_0 \in I_1^\ominus$*

$$|\{u \geq H\} \cap B_1| \geq \alpha |B_1|,$$

then there exist $\delta, \theta \in (0, 1)$, depending only on $d, \Lambda, \alpha, q, C_1, C_2$, such that if $I_\delta^\oplus(t_0) \subset Q_2$, and

$$\left(\fint_{I_2^\ominus} \mathrm{Tail}(u_-(t); 2)^q \mathrm{d}t \right)^{1/q} \leq \delta^{\frac{2s}{q}} \theta H,$$

then

$$u \geq \frac{\theta H}{8} \quad \text{in } I_{\delta/2}^\ominus(t_0 + \delta^{2s}) \times B_{1/2}.$$

The proof of Corollary 2.2 goes in the same way as before.

We are now in position to explain the proof of the Hölder regularity estimate (1.11) in slanted cylinders:

Proof of (1.11). The proof is split into two steps:
Step 1: First, we mimic the proof of [NST23, Theorem 4.1], using Corollary 2.2 on scale one, instead of the growth lemma in [NST23] for the harmonic extension. Note that we also need to make use of the suitable rescaling of (1.1), which was introduced in Lemma 2.8 and preserves the quantities in (1.10) in the same way as in the proof of [NST23, Theorem 4.1]. The arguments from this article transferred to our setup (replacing also the L^∞ global bound by the L^q-tail, as it appears in Corollary 2.2) yield that there exists $\delta > 0$ such that for any $(t, x) \in Q_{1/2}$ it holds:

$$\|u\|_{C^\alpha(Q_{\frac{\delta}{C_1}}(t,x))} \leq c \left[\|u\|_{L^\infty(Q_{3/4})} + \left(\fint_{I_1^\ominus} \mathrm{Tail}(u(t); 1)^q \mathrm{d}t \right)^{1/q} \right],$$

where $c > 0$ depends only on d, Λ, q, C_2. From here,
Step 2: The next step is to prove the Hölder estimate on scale R. To do so, we first observe that a simple covering argument, together with the

local boundedness estimate (2.27), implies

$$
\|u\|_{C^\alpha(Q_{1/2})} \le c \left[\|u\|_{L^\infty(Q_{3/4})} + \left(\fint_{I_1^\ominus} \mathrm{Tail}(u(t);1)^q \mathrm{d}t \right)^{1/q} \right]
$$
(2.30)
$$
\le c \left[\fint_{Q_1} |u| \mathrm{d}x\mathrm{d}t + \left(\fint_{I_1^\ominus} \mathrm{Tail}(u(t);1)^q \mathrm{d}t \right)^{1/q} \right],
$$

where $c > 0$ depends only on d, Λ, q, C_1, C_2. Next, we apply Lemma 2.8 and obtain:

$$
\|u\|_{C^\alpha(\tilde{Q}_{R/2})} = \|u_R\|_{C^\alpha(Q_{1/2})}
$$
$$
\le c \left[\fint_{Q_1} |u_R| \mathrm{d}x\mathrm{d}t + \left(\fint_{I_1^\ominus} \mathrm{Tail}(u_R(t);1)^q \mathrm{d}t \right)^{1/q} \right]
$$
$$
= c \left[\fint_{\tilde{Q}_R} |u| \mathrm{d}x\mathrm{d}t + \left(\fint_{I_R^\ominus} \widetilde{\mathrm{Tail}}(u(t);1)^q \mathrm{d}t \right)^{1/q} \right],
$$

as desired. $\qquad\square$

3. Comparison estimates

The goal of this section is to prove the following comparison estimate, which compares a solution to (1.1) with a solution to the homogeneous equation (1.3). Note that in this section we do not use any other property of b except for $\mathrm{div}(b) = 0$. Thus, the same proof works for $L^\infty-$ and BMO-drifts. For a comment on how (3.2) might be additionally adapted in case $s = 1/2$ when $b \in \mathrm{BMO}$, we refer to the end of this section (see (3.4)).

Lemma 3.1 (Comparison estimate). *Let u be a weak solution to*

$$
\partial_t u + (b, \nabla u) + \mathcal{L}_t u = \mu \in \mathcal{M}(\mathbb{R}^{d+1}) \quad \text{in } Q_{2r},
$$

and v be the unique weak solution to

(3.1)
$$
\begin{cases}
\partial_t v + (b, \nabla v) + \mathcal{L}_t v & = 0 \quad \text{in } Q_r, \\
v & = u \quad \text{in } I_r^\ominus \times (\mathbb{R}^d \setminus B_r), \\
v & = u \quad \text{in } \{-r^{2s}\} \times \mathbb{R}^d.
\end{cases}
$$

We set $w := u - v$. Then, we have

(3.2)
$$
\sup_{t \in I_r^\ominus} \fint_{B_r} |w| \le c \left(\frac{|\mu|(Q_r)}{|B_r|} \right),
$$

where $c > 0$ is a constant depending only on d, s, Λ.

Remark 3.1. The existence of the weak solution v to the problem (3.1) can for instance be established by standard variational methods, see for instance [BLS21, Theorem A.3] for a similar existence proof in the setting of the parabolic fractional p-Laplacian.

Proof. Let us test the weak formulations for u and v with $\phi_\varepsilon^\pm = \pm \left(1 \wedge \frac{w_\pm}{\varepsilon}\right)$ and subtract the two resulting identities. This yields:

$$\int_{B_r} (\partial_t w)\phi_\varepsilon^\pm dx + \int_{B_r} (b(t,x), \nabla\phi_\varepsilon^\pm(t,x))w(t,x)dx$$

$$+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (w(t,x) - w(t,y))(\phi_\varepsilon^\pm(t,x) - \phi_\varepsilon^\pm(t,y))K_t(x,y)dydx$$

$$= \int_{B_r} \phi_\varepsilon d\mu \leq |\mu(t)|(B_r),$$

where we used that $|\phi_\varepsilon^\pm| \leq 1$. Note that since we want to estimate the left hand side from below, we can neglect the second term, since by definition

$$(w(t,x) - w(t,y))(\phi_\varepsilon^\pm(t,x) - \phi_\varepsilon^\pm(t,y)) \geq 0.$$

Moreover, the drift term can be treated as follows, using $\operatorname{div}(b) = 0$:

$$\int_{B_r} (b(t,x), \nabla\phi_\varepsilon^\pm(t,x))w(t,x)dx = \frac{1}{2}\int_{\mathbb{R}^d} (b(t,x), \nabla((\phi_\varepsilon^\pm)^2(t,x)))dx = 0.$$

Thus, we have

$$(3.3) \qquad \int_{B_r} (\partial_t w)\phi_\varepsilon^\pm dx \leq |\mu(t)|(B_r).$$

Moreover, let us compute

$$(\partial_t w)\phi_\varepsilon^\pm = \partial_t \int_0^{w_\pm} \left(1 \wedge \frac{\sigma}{\varepsilon}\right) d\sigma.$$

Now, given $\delta > 0$ and $\tau \in I_r^\ominus$, let us define

$$\eta_\delta(t) = \begin{cases} 1, \ t \leq \tau, \\ 1 - \varepsilon^{-1}(\tau - t), \ \tau \leq t \leq \tau + \varepsilon, \\ 0, \ t > \tau + \varepsilon. \end{cases}$$

We observe that after integration by parts:

$$\int_{I_r^\ominus} \eta_\delta \int_{B_r} (\partial_t w)\phi_\varepsilon^\pm dxdt$$

$$= \int_{I_r^\ominus} (-\partial_t\eta_\delta) \int_{B_r} \left(\int_0^{w_\pm} \left(1 \wedge \frac{\sigma}{\varepsilon}\right) d\sigma\right) dxdt$$

$$- \eta_\delta(-r^{2s}) \int_{B_r} \left(\int_0^{w_\pm(-r^{2s},x)} \left(1 \wedge \frac{\sigma}{\varepsilon}\right) d\sigma\right) dx$$

$$= \int_{I_r^\ominus} (-\partial_t\eta_\delta) \int_{B_r} \left(\int_0^{w_\pm} \left(1 \wedge \frac{\sigma}{\varepsilon}\right) d\sigma\right) dxdt.$$

In the last step we used that $w(-r^{2s}) \equiv 0$. Let us now multiply (3.3) on both sides with η_δ and integrate over $I_r^\ominus$. Then, by plugging in the previous observation, we obtain

$$\int_{I_r^\ominus} (-\partial_t\eta_\delta) \int_{B_r} \left(\int_0^{w_\pm} \left(1 \wedge \frac{\sigma}{\varepsilon}\right) d\sigma\right) dxdt \leq \int_{I_r^\ominus} \eta_\delta(t)|\mu(t)|(B_r)dt \leq |\mu|(Q_r),$$

where we also used that $|\eta_\delta| \leq 1$. Now, taking the limit $\varepsilon \searrow 0$, we observe that by the monotone convergence theorem:

$$\int_{I_r^\ominus} (-\partial_t \eta_\delta) \int_{B_r} w_\pm \mathrm{d}x \mathrm{d}t \leq |\mu|(Q_r).$$

Moreover, since $-\partial_t \eta_\delta \to \delta_\tau$, as $\delta \searrow 0$, we obtain from the dominated convergence theorem

$$\int_{B_r} w_\pm(\tau, x) \mathrm{d}x \leq |\mu|(Q_r),$$

and by adding the corresponding estimates for the positive and negative part, we obtain (3.2), as desired. $\square$

We close this section by commenting on the case $s = 1/2$ and $b \in \mathrm{BMO}$. By carefully tracking the proof and replacing Q_r by $\tilde{Q}_r$, it becomes apparent that the following result holds true instead of (3.2) when u solves the equation in $\tilde{Q}_{2r}$ and v is a solution in $\tilde{Q}_r$:

$$(3.4) \qquad \sup_{t \in I_r^\ominus} \fint_{\tilde{B}_{t,r}} |w| \leq c \left(\frac{|\mu|(\tilde{Q}_r)}{|B_r|} \right).$$

Here the balls $\tilde{B}_{t,r}$ are defined as in (4.5) below.

4. Proof of potential estimates

We are now in position to prove the excess decay lemma given by Lemma 1.1, which turns out to be the crucial ingredient for the proof of the potential estimates given by Theorem 1.2.

Proof of Lemma 1.1. Let v be the comparison solution as defined in Section 3 with respect to $r = R$. Splitting into annuli and using the parabolic Hölder estimate (see Theorem 1.1) yields

$E(v, t_0, x_0, 2^{-m}R)$

$$\leq \left(\fint_{I_{2^{-m}R}^\ominus(t_0)} \left[\fint_{B_{2^{-m}R}(x_0)} |v - (v)_{Q_{2^{-m}R}(t_0,x_0)}| \mathrm{d}x \right]^q \mathrm{d}t \right)^{1/q}$$

$$+ (2^{-m}R)^{2s} \sum_{k=1}^m \left(\fint_{I_{2^{-m}R}^\ominus(t_0)} \left[\int_{B_{2^{k-m}R}(x_0) \setminus B_{2^{k-m-1}R}(x_0)} \frac{|v - (v)_{Q_{2^{-m}R}(t_0,x_0)}|}{|x_0 - y|^{d+2s}} \mathrm{d}y \right]^q \mathrm{d}t \right)^{1/q}$$

$$+ (2^{-m}R)^{2s(1-1/q)} \left(\int_{I_{2^{-m}R}^\ominus(t_0)} \left[\int_{\mathbb{R}^d \setminus B_R(x_0)} \frac{|v - (v)_{Q_{2^{-m}R}(t_0,x_0)}|}{|x_0 - y|^{d+2s}} \mathrm{d}y \right]^q \mathrm{d}t \right)^{1/q}$$

$$\leq C \sum_{k=0}^{m-1} 2^{-2sk} \left(\fint_{I_{2^{-m}R}^\ominus(t_0)} \left[\fint_{B_{2^{k-m}R}(x_0)} |v - (v)_{Q_{2^{-m}R}(t_0,x_0)}| \mathrm{d}y \right]^q \mathrm{d}t \right)^{1/q}$$

$$+ C 2^{-2s(1-1/q)m} \left(\fint_{I_R^\ominus(t_0)} \left[\fint_{B_R(x_0)} |v - (v)_{Q_{2^{-m}R}(t_0,x_0)}| \mathrm{d}y \right]^q \mathrm{d}t \right)^{1/q}$$

$$+ 2^{-2s(1-1/q)m} \| \mathrm{Tail}(v(\cdot) - (v)_{Q_{2^{-m}R}(t_0,x_0)}; x_0, R) \|_{L^q(I_R^\ominus(t_0))}$$

$$\leq C \sum_{k=0}^{m-1} 2^{-2sk} \operatorname*{osc}_{Q_{2^{k-m}R}(t_0,x_0)} v + C 2^{-2s(1-1/q)m} E(v,t_0,x_0,R)$$

$$\leq C 2^{-\alpha m} \sum_{k=0}^{\infty} 2^{(\alpha-2s)k} E(v,t_0,x_0,R) + C 2^{-2s(1-1/q)m} E(v,t_0,x_0,R)$$

$$\leq C 2^{-\alpha m} E(v,t_0,x_0,R).$$

Together with the comparison estimate from Lemma 3.1, we obtain that

$$E(u,t_0,x_0,2^{-m}R)$$
$$\leq E(v,t_0,x_0,2^{-m}R) + E(u-v,t_0,x_0,2^{-m}R)$$
$$\leq C 2^{-\alpha m} E(v,t_0,x_0,R) + C 2^{((d+2s)/q)m} E(u-v,t_0,x_0,R)$$
$$\leq C 2^{-\alpha m} E(u,t_0,x_0,R) + C 2^{((d+2s)/q)m} R^{-d} |\mu|(Q_R(t_0,x_0)),$$

finishing the proof. $\qquad\square$

Proof of Theorem 1.2. Let $m \geq 1$ to be chosen large enough in a way such that m only depends on d, s, Λ. In particular, we require m to be large enough such that

(4.1) $$C_0 2^{-\alpha m/2} \leq \frac{1}{2}.$$

Observe that this in particular implies that for any integer $i \geq 0$,

(4.2) $$C_0 2^{-\alpha m(i+1)/2} \leq 2^{-(i+1)}.$$

Moreover, for any integer $i \geq 0$, set

$$E_i := E(u,t_0,x_0,2^{-im}R).$$

By Lemma 1.1 and (4.1), (4.2), for any integer $i \geq 0$ we have

(4.3) $$E_{i+1} \leq \frac{1}{2} E_i + C_0 2^{((d+2s)/q)m} (2^{-im}R)^{-d} |\mu|(Q_{2^{-im}R}(t_0,x_0)).$$

For $l \in \mathbb{N}$, summing (4.3) over $i \in \{0,...,l-1\}$ leads to

$$\sum_{i=1}^{l} E_i \leq \frac{1}{2} \sum_{i=0}^{l-1} E_i + C_0 2^{((d+2s)/q)m} \sum_{i=0}^{l-1} (2^{-im}R)^{-d} |\mu|(Q_{2^{-im}R}(t_0,x_0)),$$

so that by reabsorbing the first term on the right-hand side of the previous inequality we deduce

$$\sum_{i=1}^{l} E_i \leq 2E_0 + 2C_0 2^{((d+2s)/q)m} \sum_{i=0}^{\infty} (2^{-im}R)^{-d} |\mu|(Q_{2^{-im}R}(t_0,x_0)).$$

Thus, we have

$$|(u)_{Q_{2^{-(l+1)m}R}(t_0,x_0)}|$$
$$\leq \sum_{i=0}^{l} (|(u)_{Q_{2^{-(i+1)m}R}(t_0,x_0)} - (u)_{Q_{2^{-im}R}(t_0,x_0)}|) + |(u)_{Q_R(t_0,x_0)}|$$
$$\leq C_1 \sum_{i=1}^{l} E_i + \fint_{Q_R(t_0,x_0)} |u| \mathrm{d}x\mathrm{d}t$$

$$\leq C\Bigg(\bigg(\fint_{I_R^\ominus(t_0)}\Big[\fint_{B_R(t_0,x_0)}|u|\mathrm{d}x\Big]^q\,dt\bigg)^{1/q}$$
$$+\,R^{-2s/q}\|\operatorname{Tail}(u(\cdot)-(u)_{Q_R(t_0,x_0)};x_0,R)\|_{L^q(I_R^\ominus(t_0))}$$
$$+\,\sum_{i=0}^\infty (2^{-im}R)^{-d}|\mu|(Q_{2^{-im}R}(t_0,x_0))\Bigg),$$

where C depends only on d,s,Λ,q, which is in particular because m only depends on the aforementioned quantities. Since the last term on the right-hand side can be controlled by the parabolic potential of order $2s$, the proof is finished by applying the Lebesgue differentiation theorem.

$\square$

We end this section by commenting on how to obtain the parabolic potential estimates in case $b\equiv 0$ for the full range $s\in(0,1)$ (see Remark 1.2).

Remark 4.1. In case $b\equiv 0$, and $s\in(0,1)$, the Hölder estimate Proposition 1.1 was obtained in [KW23a]. Moreover, it is easy to see that Lemma 3.1 can be proved in the exact same way in this case. Having at hand these two results, one can follow the proofs of Lemma 1.1 and Theorem 1.2 line by line and thereby deduce the potential estimate for parabolic equations without drift and $s\in(0,1)$.

4.1. **The case of BMO drifts.** Clearly, since Lemma 2.1 remains true for BMO drifts if $s>1/2$, the proof of the potential estimate Theorem 1.2 in this setting follows in the exact same way as described before. Thus, in the following, we restrict ourselves to proving (1.12) in the critical case $s=1/2$.

Lemma 4.1 (Excess decay under BMO drifts). *Let $s=1/2$, assume that (1.10) holds, let $R>0$, $t_0\in I$, $x_0\in\mathbb{R}^n$, $q>1$ and $\mu\in\mathcal{M}(\mathbb{R}^{d+1})$. In addition, let $\alpha\in(0,s(1-1/q))$ be given by (1.11). Let u be a weak solution to (1.1) in $I\times\Omega$ and assume that $Q_{\theta R}(t_0,x_0)\subset I\times\Omega$, where $\theta-1=c(C_1+C_2|\log(R)|)$ for some $c>0$, depending only on d. Then for any integer $m\geq 1$, we have*

$$(4.4)\qquad \widetilde{E}(u,t_0,x_0,2^{-m}R)\leq C_0 2^{-\alpha m}\widetilde{E}(u,t_0,x_0,R)$$
$$+\,C_0 2^{((d+2s)/q)m}R^{-d}|\mu|(\tilde{Q}_R(t_0,x_0)),$$

where $C_0\geq 1$ depends only on d,s,Λ,α,q and for any $r>0$, the slanted excess is defined by

$$\widetilde{E}(u,t_0,x_0,r):=\bigg(\fint_{I_r^\ominus(t_0)}\Big[\fint_{\tilde{B}_{t,r}(x_0)}|u-(u)_{\tilde{Q}_r(t_0,x_0)}|\mathrm{d}x\Big]^q\,\mathrm{d}t\bigg)^{1/q}$$
$$+\,r^{-2s/q}\|\widetilde{\operatorname{Tail}}(u(\cdot)-(u)_{\tilde{Q}_r(t_0,x_0)};x_0,r)\|_{L^q(I_r^\ominus(t_0))},$$

where for any $t\in\mathbb{R}$,

$$(4.5)\qquad \tilde{B}_{t,r}(x_0)=\{x+rz_r(t/r):x\in B_r(x_0)\}.$$

Proof. Let v be the comparison solution as defined in Section 3 with respect to $r = R$. Splitting into annuli and using the parabolic Hölder estimate in the case of BMO drifts (see (1.11)) yields

$$\widetilde{E}(v, t_0, x_0, 2^{-m}R)$$

$$\leq \left(\fint_{I_{2^{-m}R}^{\ominus}(t_0)} \left[\fint_{\tilde{B}_{t,2^{-m}R}(x_0)} |v - (v)_{\tilde{Q}_{2^{-m}R}(t_0,x_0)}| dx \right]^q dt \right)^{1/q}$$

$$+ (2^{-m}R)^{2s}$$

$$\times \sum_{k=1}^{m} \left(\fint_{I_{2^{-m}R}^{\ominus}(t_0)} \left[\int_{B_{2^{k-m}R}(x_0) \setminus B_{2^{k-m-1}R}(x_0)} \frac{|v(t, y + Rz_R(t/R)) - (v)_{\tilde{Q}_{2^{-m}R}(t_0,x_0)}|}{|x_0 - y|^{d+2s}} dy \right]^q dt \right)^{1/q}$$

$$+ (2^{-m}R)^{2s(1-1/q)} \left(\int_{I_{2^{-m}R}^{\ominus}(t_0)} \left[\int_{\mathbb{R}^d \setminus B_R(x_0)} \frac{|v(t, y + Rz_R(t/R)) - (v)_{\tilde{Q}_{2^{-m}R}(t_0,x_0)}|}{|x_0 - y|^{d+2s}} dy \right]^q dt \right)^{1/q}$$

$$\leq C \sum_{k=0}^{m-1} 2^{-2sk} \left(\fint_{I_{2^{-m}R}^{\ominus}(t_0)} \left[\fint_{\tilde{B}_{t,2^{k-m}R}(x_0)} |v - (v)_{\tilde{Q}_{2^{-m}R}(t_0,x_0)}| dy \right]^q dt \right)^{1/q}$$

$$+ C 2^{-2s(1-1/q)m} \left(\fint_{I_R^{\ominus}(t_0)} \left[\fint_{\tilde{B}_{t,R}(x_0)} |v - (v)_{\tilde{Q}_{2^{-m}R}(t_0,x_0)}| dy \right]^q dt \right)^{1/q}$$

$$+ 2^{-2s(1-1/q)m} ||\widetilde{\mathrm{Tail}}(v(\cdot) - (v)_{\tilde{Q}_{2^{-m}R}(t_0,x_0)}; x_0, R)||_{L^q(I_R^{\ominus}(t_0))}$$

$$\leq C \sum_{k=0}^{m-1} 2^{-2sk} \operatorname*{osc}_{\tilde{Q}_{2^{k-m}R}(t_0,x_0)} v + C 2^{-2s(1-1/q)m} \widetilde{E}(v, t_0, x_0, R)$$

$$\leq C 2^{-\alpha m} \sum_{k=0}^{\infty} 2^{(\alpha-2s)k} \widetilde{E}(v, t_0, x_0, R) + C 2^{-2s(1-1/q)m} \widetilde{E}(v, t_0, x_0, R)$$

$$\leq C 2^{-\alpha m} \widetilde{E}(v, t_0, x_0, R).$$

Together with the slanted comparison estimate from (3.4), we obtain that

$$\widetilde{E}(u, t_0, x_0, 2^{-m}R)$$
$$\leq \widetilde{E}(v, t_0, x_0, 2^{-m}R) + \widetilde{E}(u - v, t_0, x_0, 2^{-m}R)$$
$$\leq C 2^{-\alpha m} \widetilde{E}(v, t_0, x_0, R) + C 2^{((d+2s)/q)m} \widetilde{E}(u - v, t_0, x_0, R)$$
$$\leq C 2^{-\alpha m} \widetilde{E}(u, t_0, x_0, R) + C 2^{((d+2s)/q)m} R^{-d} |\mu|(\tilde{Q}_R(t_0, x_0)),$$

finishing the proof. □

We are now ready to prove (1.12).

Proof of (1.12). The proof easily follows after replacing the standard excess $E(u, t_0, x_0, r)$ in the proof of Theorem 1.2 by the slanted excess $\widetilde{E}(u, t_0, x_0, r)$. □

5. UPPER HEAT KERNEL ESTIMATES FOR DRIFT-DIFFUSION EQUATIONS

Let $\mathcal{L}_t$ and b be as before with $||b||_{L^\infty(\mathbb{R}^{d+1})} \leq \Lambda$, and assume that K_t satisfies (i), (ii), (iii) for some $s \in [1/2, 1)$ and $\Lambda > 0$. Given $\eta \in [0, \infty)$

and $y \in \mathbb{R}^d$, we define $(t, x) \mapsto p(\eta, x; t, y)$ as the solution to the following problem:

$$(5.1) \qquad \begin{cases} \partial_t u + (b, \nabla u) + \mathcal{L}_t u & = 0 \text{ in } [\eta, \infty) \times \mathbb{R}^d, \\ u(\eta) & = \delta_y. \end{cases}$$

We call p the heat kernel of the operator $\partial_t + (b, \nabla \cdot) + \mathcal{L}_t$.

The goal of this section is to prove Theorem 1.3, i.e. to show that for any $\eta \le t \le T$ it holds:

$$(5.2) \qquad p(\eta, x; t, y) \le c \left((t - \eta)^{-\frac{d}{2s}} \wedge \frac{t - \eta}{|x - y|^{d + 2s}} \right) \quad \forall x, y \in \mathbb{R}^d$$

for some constant $c > 0$, depending only on d, s, Λ, T.

Note that given $\eta \in [0, \infty)$ and $y \in \mathbb{R}^d$, the function $(t, x) \mapsto p(\eta, y; t, x)$ solves the following dual problem:

$$(5.3) \qquad \begin{cases} \partial_t u - (b, \nabla u) + \mathcal{L}_t u & = 0 \text{ in } [\eta, \infty) \times \mathbb{R}^d, \\ u(0) & = \delta_y. \end{cases}$$

This is a simple consequence of the fact that b is divergence free. Let us also introduce the notation $\widehat{p}(\eta, x; t, y) = p(\eta, y; t, x)$ and define the corresponding semigroups (P_t) and $(\widehat{P}_t)$ as follows:

$$P_{\eta, t} f(x) = \int_{\mathbb{R}^d} p(\eta, x; t, y) f(y) \mathrm{d}y, \ \ \widehat{P}_{\eta, t} f(x) = \int_{\mathbb{R}^d} \widehat{p}(\eta, x; t, y) f(y) \mathrm{d}y, \ \ f \in L^2(\mathbb{R}^d).$$

Let us first collect a few basic properties of the heat kernel and the associated semigroup:

Lemma 5.1. *The heat kernel $p(\eta, x; t, y)$ exists under the assumptions of Theorem 1.3, and it holds:*

$$(5.4) \qquad \int_{\mathbb{R}^d} p(\eta, x; t, y) \mathrm{d}y = \int_{\mathbb{R}^d} p(\eta, x; t, y) \mathrm{d}x = 1,$$

$$(5.5) \qquad 0 \le p(\eta, x; t, y) \le c(t - \eta)^{-\frac{d}{2s}},$$

$$(5.6) \qquad p(\eta, x; t, y) = \int_{\mathbb{R}^d} p(\eta, x; \tau, z) p(\tau, z; t, y) \mathrm{d}z \ \forall \eta < \tau < t.$$

Moreover, for any $u_0 \in L^2(\mathbb{R}^d)$, $P_{\eta, t} f(x) = \int_{\mathbb{R}^d} u_0(y) p(\eta, x; t, y) \mathrm{d}y$ and $\widehat{P}_{\eta, t} u_0(x) = \int_{\mathbb{R}^d} u_0(y) p(\eta, x; t, y) \mathrm{d}x$ solve (5.1) with $P_{\eta, \eta} u_0 = u_0$, and (5.3) with $\widehat{P}_{\eta, \eta} u_0 = u_0$, respectively.

Proof. The proof of all properties is standard. For more details, we refer to [MM13a]. Let us just mention that the upper bound in (5.5) is a direct consequence of Lemma 2.2 applied with $Q_{R/2}^{\ominus}(t, x)$, where $R = (t - \eta)^{1/(2s)}$, and (5.4), i.e.

$$p(\eta, x; t, y) \le c(t - \eta)^{-\frac{d}{2s}} \sup_{\tau \in (\eta, t)} \int_{\mathbb{R}^d} p(\eta, z; \tau, y) \mathrm{d}y \le c(t - \eta)^{-\frac{d}{2s}}.$$

$\square$

Before we start with the actual proof of Theorem 1.3, we need to introduce corresponding truncated objects. First, given $\rho > 0$, we define

$$\mathcal{L}_t^\rho u(x) = \text{p.v.} \int_{B_\rho(x)} (u(x) - u(y)) K_t(x,y) \mathrm{d}y.$$

The heat kernel and semigroup corresponding to the problem

$$\partial_t u + (b, \nabla u) + \mathcal{L}_t^\rho u = 0 \text{ in } [\eta, \infty) \times \mathbb{R}^d,$$

are denoted as $p^\rho, \widehat{p}^\rho, (P_t^\rho), (\widehat{P}_t^\rho)$, and are defined in the same way as their non-truncated counterparts, replacing $\mathcal{L}_t$ by $\mathcal{L}_t^\rho$ in all definitions. Moreover, Lemma 5.1 remains true with p replaced by p^ρ and $P_{\eta,t}$ replaced by $P_{\eta,t}^\rho$. Moreover, we introduce the truncated "carré du champ" operator

$$\Gamma_s^\rho(v,v)(x) = \int_{B_\rho(x)} (v(x) - v(y))^2 |x-y|^{-d-2s} \mathrm{d}y.$$

We need the following truncated version of the local boundedness-type estimate Lemma 2.2:

Lemma 5.2. *There exists a constant $C > 0$, depending only on d, s, Λ, such that for every $t_0 \in (0, \infty)$, $x_0 \in \mathbb{R}^d$, and $\rho, R > 0$ with $R \leq \rho/2 \wedge t_0^{1/(2s)}$, and every solution u to $\partial_t u + (b, \nabla u) + \mathcal{L}_t^\rho u = 0$ in $Q_R(t_0, x_0)$ it holds:*

$$(5.7) \qquad \sup_{Q_{R/2}(t_0,x_0)} u \leq C \left(\frac{\rho}{R}\right)^{\frac{d}{2}} R^{-\frac{d}{2}} \sup_{t \in I_R^\ominus(t_0)} \left(\int_{B_{2\rho}(x_0)} u^2(t,x) \mathrm{d}x\right)^{1/2}.$$

Proof. Recall that we have the same energy estimate (2.2) as in case $b \equiv 0$. The energy estimate remains true if $\mathcal{L}_t$ is replaced by $\mathcal{L}_t^\rho$ with $\text{Tail}(w_+; 0, r + \rho_2)$ replaced by $\int_{B_\rho \setminus B_{r+\rho_2}} w_+(y)|y|^{-d-2s} \mathrm{d}y$. From here, the proof follows along the exact same arguments as the proof of [KW23b, Lemma 2.4]. $\square$

5.1. **Aronson's weighted estimate.** The following lemma is an adaptation of [KW23b, Lemma 3.1]. It is a nonlocal adaptation of Aronson's weighted L^2-estimate, which is the main ingredient in the proof of the upper heat kernel estimates. Due to the presence of the divergence free drift term, we get the additional summand $|\nabla H|$ in the first condition on H, compared to [KW23b].

Lemma 5.3 (Aronson's weighted estimate). *Assume that $\rho > 0$, $0 \leq \eta < T$. Let $u \in L^\infty((\eta, T) \times \mathbb{R}^d)$ be a solution to the ρ-truncated problem $\partial_t u + (b, \nabla u) + \mathcal{L}_t^\rho u = 0$ in $(\eta, T) \times \mathbb{R}^d$ with $u(\eta) = u_0$. Then there exists a constant $C > 0$, depending only on d, s, Λ, such that for every bounded function $H : [\eta, t] \times \mathbb{R}^d \to [0, \infty)$ satisfying*

- $C\left(|\nabla H| + \Gamma_s^\rho(H^{1/2}, H^{1/2})\right) \leq -\partial_t H$ *in* $(\eta, t) \times \mathbb{R}^d$,
- $H^{1/2} \in L^2((\eta, t); H^s(\mathbb{R}^d))$,

the following estimate holds true:

$$(5.8) \qquad \sup_{\tau \in (\eta,t)} \int_{\mathbb{R}^d} H(\tau, x) u^2(\tau, x) \mathrm{d}x \leq \int_{\mathbb{R}^d} H(\eta, x) u_0^2(x) \mathrm{d}x.$$

Proof. The proof is very similar to the one of [KW23b, Lemma 3.1]. As in [KW23b], let us take $\gamma_R \in C_c^\infty(\mathbb{R}^d)$ is such that $\gamma_R \equiv 1$ in $B_{R-1}(0)$, $\gamma_R \equiv 0$ in $\mathbb{R}^d \setminus B_R(0)$, $0 \le \gamma_R \le 1$, $|\nabla \gamma_R| \le 2$ for $R > 0$.
Note that

$$\int_\eta^t \int_{\mathbb{R}^d} (b, \nabla u(\tau, x)) \gamma_R u H(\tau, x) \mathrm{d}x \mathrm{d}\tau = \frac{1}{2} \int_\eta^t \int_{\mathbb{R}^d} (b, \nabla(u^2)) \gamma_R H(\tau, x) \mathrm{d}x \mathrm{d}\tau$$

$$= \frac{1}{2} \int_\eta^t \int_{\mathbb{R}^d} (b, \nabla \gamma_R) u^2 H(\tau, x) \mathrm{d}x \mathrm{d}\tau + \frac{1}{2} \int_\eta^t \int_{\mathbb{R}^d} (b, \nabla H) u^2 \gamma_R \mathrm{d}x \mathrm{d}\tau.$$

We test the weak formulation with $\phi = \gamma_R^2 H u$, and obtain from the previous computation, as well as the same arguments as in [KW23b]:

$$\sup_{\tau \in (\eta, t)} \int_{\mathbb{R}^d} u^2(\tau, x) \gamma_R^2(x) H(\tau, x) \mathrm{d}x \le \int_{\mathbb{R}^d} u_0^2(x) \gamma_R^2(x) H(\eta, x) \mathrm{d}x$$

$$+ c_2 \int_\eta^t \int_{\mathbb{R}^d} \nabla(u^2) \gamma_R H \mathrm{d}x \mathrm{d}\tau$$

$$+ \int_\eta^t \int_{\mathbb{R}^d} u^2 \gamma_R^2 \left(2c_2 \Gamma_s^\rho(H^{1/2}, H^{1/2}) + 2c_2 |\nabla H| + \partial_t H \right) \mathrm{d}x \mathrm{d}\tau$$

$$+ 2c_2 \int_\eta^t \int_{\mathbb{R}^d} \int_{B_\rho(x)} u^2(\tau, x) \gamma_R^2(z) (H^{1/2}(\tau, x) - H^{1/2}(\tau, z))^2 |x - y|^{-d-2s} \mathrm{d}z \mathrm{d}x \mathrm{d}\tau$$

$$+ 2c_2 \int_\eta^t \int_{\mathbb{R}^d} u^2 H \Gamma_s^\rho(\gamma_R, \gamma_R) \mathrm{d}x \mathrm{d}\tau$$

$$+ 2c_2 \int_\eta^t \int_{\mathbb{R}^d} \int_{B_\rho(x)} u^2(\tau, z) H(\tau, x) (\gamma_R(x) - \gamma_R(z))^2 |x - y|^{-d-2s} \mathrm{d}z \mathrm{d}x \mathrm{d}\tau$$

$$+ 2c_2 \int_\eta^t \int_{\mathbb{R}^d} u^2 H |\nabla \gamma_R| \mathrm{d}x \mathrm{d}\tau$$

$$\le \int_{\mathbb{R}^d} u_0^2(x) \gamma_R^2(x) H(\eta, x) \mathrm{d}x$$

$$+ \int_\eta^t \int_{\mathbb{R}^d} u^2 \gamma_R^2 \left(4c_2 \Gamma_s^\rho(H^{1/2}, H^{1/2}) + 4c_2 |\nabla H| + \partial_t H \right) \mathrm{d}x \mathrm{d}\tau$$

$$+ 2c_2 \|u\|_\infty^2 \int_\eta^t \int_{\mathbb{R}^d \setminus B_{\frac{R-1}{2}}(0)} \Gamma_s^\rho(H^{1/2}, H^{1/2}) \mathrm{d}x \mathrm{d}\tau$$

$$+ 4c_2 \|u\|_\infty^2 \int_\eta^t \int_{\mathbb{R}^d} \left(\Gamma_s^\rho(\gamma_R, \gamma_R) + |\nabla \gamma_R| \right) H \mathrm{d}x \mathrm{d}\tau.$$

By the definition of γ_R, and since $H \in L^1((\eta, t) \times \mathbb{R}^d)$ by assumption, we have that the last term in the previous estimate converges to zero, as $R \to \infty$. The second term goes to zero by the same arguments as in [KW23b, Lemma 3.1]. This implies the desired result. $\qquad\square$

5.2. Gluing lemma.
The goal of this section is to prove the following relation between the truncated heat kernel p^ρ and p:

Lemma 5.4 (Gluing lemma). *Let $\rho > 0$ and $0 \le \eta \le t$. Then,*

$$p(\eta, x; t, y) \le p^\rho(\eta, x; t, y) + c(t - \eta)\rho^{-d-2s} \quad \forall x, y \in \mathbb{R}^d.$$

Moreover, it holds

$$p^\rho(\eta, x; t, y) \leq e^{C(t-\eta)\rho^{-2s}} p(\eta, x; t, y) \quad \forall x, y \in \mathbb{R}^d.$$

The constants $c, C > 0$ depend only on d, s, Λ.

A central ingredient in its proof is the following parabolic maximum principle, which remains true in the presence of a divergence free drift b. For $\Omega \subset \mathbb{R}^d$ open, we define $H_\Omega^s(\mathbb{R}^d) = \{u \in H^s(\mathbb{R}^d) : u \equiv 0 \text{ in } \mathbb{R}^d \setminus \Omega\}$.

Lemma 5.5. *Let $\Omega \subset \mathbb{R}^d$ be open. Assume that u solves*

$$(5.9) \qquad \begin{cases} \partial_t u + (b, \nabla u) + \mathcal{L}_t u \leq 0, \text{ in } (\eta, \infty) \times \Omega, \\ u_+(t) \in H_\Omega^s(\mathbb{R}^d) \ \forall t \in (\eta, \infty), \\ u_+(t) \to 0 \text{ in } L^2(\Omega), \text{ as } t \searrow \eta. \end{cases}$$

Then $u \leq 0$ a.e. in $(\eta, \infty) \times \Omega$. The same result holds for subsolutions to $\partial_t u + (b, \nabla u) + \mathcal{L}_t^\rho u \leq 0$ in $(\eta, \infty) \times \Omega$.

Proof. First, we observe that

$$\int_{\mathbb{R}^d} (b, \nabla u) u_+ \mathrm{d}x = \frac{1}{2} \int_{\mathbb{R}^d} (b, \nabla(u^2)) \mathrm{d}x = 0$$

since b is divergence free. Moreover, note that $\mathcal{E}^{K_t}(u(t), u_+(t)) \geq 0$. Thus, testing the weak formulation for u with $\phi = u_+$, we obtain for any $\eta < t_1 < t_2$:

$$\int_\Omega u_+^2(t_2, x) \mathrm{d}x - \int_\Omega u_+(t_1, x) \mathrm{d}x \leq 0.$$

The desired result follows by taking the limit $t_1 \searrow \eta$. $\qquad\square$

We are now in the position to give the

Proof of Lemma 5.4. Let $\Omega_n \subset \mathbb{R}^d$, $n \in \mathbb{N}$, be an increasing sequence of domains, such that $\Omega_n \nearrow \mathbb{R}^d$. Let $f \in L^1(\mathbb{R}^d)$ with $\|f\|_{L^1(\mathbb{R}^d)} \leq 1$. We define $(t, x) \mapsto P_{\eta,t}^{\Omega_n} f(x)$ to be the solution to $\partial_t u + (b, \nabla u) + \mathcal{L}_t u = 0$ with $u(\eta) = f$, and $(t, x) \mapsto P_{\eta,t}^{\Omega_n, \rho} f(x)$ to solve the corresponding problem with $\mathcal{L}_t$ replaced by $\mathcal{L}_t^\rho$.

The proof of the result goes as in [KW23b, Lemma 2.2] and is based on an application of the parabolic maximum principle Lemma 5.5 to the function $(t, x) \mapsto P_{\eta,t}^{\Omega_n} f(x) - u(t, x)$, where

$$u(t, x) = P_{\eta,t}^{\Omega_n, \rho} f(x) + C(t - \eta)\rho^{-d-2s} \phi_n(x),$$

and $\phi_n \in C_c^\infty(\mathbb{R}^d)$ with $0 \leq \phi_n \leq 1$, $\phi_n \equiv 1$ in Ω_n.

In order to apply Lemma 5.5, it remains to prove that u solves $\partial_t u + (b, \nabla u) + \mathcal{L}_t u \leq 0$ in $(\eta, \infty) \times \Omega_n$. To prove it, we take an arbitrary test function $\psi \in H_{\Omega_n}^s(\mathbb{R}^d)$. First, note

$$\mathcal{E}^{K_t}(P_{\eta,t}^{\rho, \Omega_n} f, \psi) - \mathcal{E}^{K_t^\rho}(P_{\eta,t}^{\rho, \Omega_n} f, \psi) \geq -c\rho^{-d-2s} \|\psi\|_{L^1(\mathbb{R}^d)} \|P_{\eta,t}^{\rho, \Omega_n} f\|_{L^1(\mathbb{R}^d)}$$

$$\geq -c\rho^{-d-2s} \|\psi\|_{L^1(\mathbb{R}^d)},$$

where $\mathcal{E}^{K_t}$ was defined in (2.1), we wrote $K_t^\rho(x,y) = K_t(x,y)\mathbb{1}_{|x-y|\leq\rho}(x,y)$, and we used:

$$\|P_{\eta,t}^{\rho,\Omega_n}f\|_{L^1(\mathbb{R}^d)} \leq \|P_{\eta,t}^\rho f\|_{L^1(\mathbb{R}^d)} \leq \|\widehat{P}_{\eta,t}^\rho 1\|_{L^\infty(\mathbb{R}^d)}\|f\|_{L^1(\mathbb{R}^d)} \leq 1.$$

Note that in the first estimate, we applied Lemma 5.5, and in the last estimate, we used (5.4). Moreover, we have

$$\int_{\mathbb{R}^d} \phi_n \psi \mathrm{d}x \geq \|\psi\|_{L^1(\mathbb{R}^d)}, \ \int_{\mathbb{R}^d}(b,\nabla\phi_n)\psi\mathrm{d}x = 0, \ \mathcal{E}^{K_t}(\phi_n,\psi) \geq 0$$

by the definitions of ϕ_n, ψ. Therefore,

$$\int_{\mathbb{R}^d} \partial_t u(t,x)\psi(x)\mathrm{d}x + \int_{\mathbb{R}^d}(b,\nabla u)\psi\mathrm{d}x + \mathcal{E}^{K_t}(u,\psi)$$
$$= \left[\int_{\mathbb{R}^d} \partial_t P_{\eta,t}^{\rho,\Omega_n}f(x)\psi(x)\mathrm{d}x + \int_{\mathbb{R}^d}(b,\nabla P_{\eta,t}^{\rho,\Omega_n}f)\psi\mathrm{d}x + \mathcal{E}^{K_t}(P_{\eta,t}^{\rho,\Omega_n}f,\psi)\right]$$
$$+ C\rho^{-d-2s}\left[\int_{\mathbb{R}^d}\phi_n\psi\mathrm{d}x + (t-\eta)\int_{\mathbb{R}^d}(b,\nabla\phi_n)\psi\mathrm{d}x + (t-\eta)\mathcal{E}^{K_t}(\phi_n,\psi)\right]$$
$$\geq -c\rho^{-d-2s}\|\psi\|_{L^1(\mathbb{R}^d)} + C\rho^{-d-2s}\|\psi\|_{L^1(\mathbb{R}^d)}$$
$$\geq \rho^{-d-2s}\|\psi\|_{L^1(\mathbb{R}^d)}(C-c).$$

Thus, we can choose $C > c$ large enough such that, indeed $(t,x) \mapsto P_{\eta,t}^{\Omega_n}f(x) - u(t,x)$ is a subsolution.

Therefore, by the maximum principle Lemma 5.5, we obtain:

$$(5.10) \qquad P_{\eta,t}^{\Omega_n}f(x) \leq P_{\eta,t}^{\Omega_n,\rho}f(x) + C(t-\eta)\rho^{-d-2s}\|f\|_{L^1(\mathbb{R}^d)}.$$

Taking $n \to \infty$ in (5.10) implies by the same arguments as in [KW23b, Lemma 5.3]

$$P_{\eta,t}f(x) \leq P_{\eta,t}^\rho f(x) + C(t-\eta)\rho^{-d-2s}\|f\|_{L^1(\mathbb{R}^d)}.$$

This yields the desired result by recalling the relation between $P_{\eta,t}f$ and p in Lemma 5.1, and choosing $f = \mathbb{1}_A$ for any measurable set $A \subset \mathbb{R}^d$.

The proof of the second claim follows by an application of the parabolic maximum principle to $u(t,x) = P_{\eta,t}^{\rho,\Omega}f - P_t^{\Omega_n}f(x)e^{(t-\eta)\rho^{-2s}}$. $\square$

5.3. Upper estimates for the truncated heat kernel.
The goal of this section is to deduce upper estimates for the heat kernel p^ρ corresponding to the truncated operator $\mathcal{L}_t^\rho$:

Proposition 5.1 (Truncated heat kernel estimate). *Let $T > 0$. There exist $c > 0$, $\nu > 1$, depending only on d, s, Λ, T, such that for every $\rho > 0$, $0 \leq \eta \leq t \leq T$, and $x, y \in \mathbb{R}^d$ with $t - \eta \leq \frac{1}{4\nu}\rho^{2s}$:*

$$(5.11) \qquad p^\rho(\eta,x;t,y) \leq c(t-\eta)^{-\frac{d}{2s}}2^{\frac{|x-y|}{12\rho}}\left(\frac{\rho^{2s}}{\nu(t-\eta)}\right)^{-\frac{|x-y|}{12\rho}+\frac{1}{2}+\frac{d}{4s}}.$$

Given $y \in \mathbb{R}^d$, $\rho > 0$, $0 \le \eta \le t \le T$, with $t - \eta \le \frac{1}{4\nu}\rho^{2s}$, where $\nu > 1$, we set

(5.12)

$$
H(\tau, x) := \left(\frac{\rho^{2s}}{\nu[2(t-\eta)-(\tau-\eta)]} \right)^{-1} \wedge \left(\frac{\rho^{2s}}{\nu[2(t-\eta)-(\tau-\eta)]} \right)^{-\frac{|x-y|}{3\rho}}
$$

$$
= e^{-\log\left(\frac{\rho^{2s}}{\nu[2(t-\eta)-(\tau-\eta)]} \right)\left(\frac{|x-y|}{3\rho} \vee 1 \right)}.
$$

The following lemma shows that H satisfies the assumption of Lemma 5.3 if $\nu > 1$ is chosen appropriately.

Lemma 5.6. *For every $C, T > 0$, there exists $\nu = \nu(d, s, C, T) > 0$ such that for every $\rho > 0$, $y \in \mathbb{R}^d$ and $0 \le \eta \le t \le T$ with $t - \eta \le \frac{1}{4\nu}\rho^{2s}$, the function H defined above satisfies:*

$$
C\left[\Gamma_s^\rho(H^{1/2}, H^{1/2}) + |\nabla H| \right] \le -\partial_t H \quad \text{in } (\eta, t) \times \mathbb{R}^d, \qquad H^{1/2} \in L^2((\eta, t); H^s(\mathbb{R}^d)).
$$

Proof. Note that H is the same function as in [KW23b]. Therefore, we refer to [KW23b, Lemma 3.3] for the computations involving $\partial_t H$ and $\Gamma_s^\rho(H^{1/2}, H^{1/2})$. In case $|x - y| \le 2\rho$, we have

$$
-\partial_t H(\tau, x) = \nu\rho^{-2s}, \ \Gamma_s^\rho(H^{1/2}, H^{1/2})(\tau, x) = |\nabla H(\tau, x)| = 0.
$$

In case $2\rho \le |x - y| \le 3\rho$, we have

$$
-\partial_t H(\tau, x) = \nu\rho^{-2s}, \ \Gamma_s^\rho(H^{1/2}, H^{1/2})(\tau, x) \le c\rho^{-2s}, \ |\nabla H(\tau, x)| = 0.
$$

In case $|x - y| > 3\rho$, it holds

$$
-\partial_t H(\tau, x) = \frac{|x-y|}{3\rho[2(t-\eta)-(\tau-\eta)]} e^{-\log\left(\frac{\rho^{2s}}{\nu[2(t-\eta)-(\tau-\eta)]} \right)\left(\frac{|x-y|}{3\rho} \right)},
$$

$$
\Gamma_s^\rho(H^{1/2}, H^{1/2}) \le c\frac{|x-y|}{3\rho\nu[2(t-\eta)-(\tau-\eta)]} e^{-\log\left(\frac{\rho^{2s}}{\nu[2(t-\eta)-(\tau-\eta)]} \right)\left(\frac{|x-y|}{3\rho} \right)},
$$

$$
|\nabla H(\tau, x)| \le c\frac{1}{3\rho} e^{-\log\left(\frac{\rho^{2s}}{\nu[2(t-\eta)-(\tau-\eta)]} \right)\left(\frac{|x-y|}{3\rho} \right)}
$$

$$
\le c\frac{|x-y|}{3\rho\nu^{\frac{1}{2s}}[2(t-\eta)-(\tau-\eta)]} e^{-\log\left(\frac{\rho^{2s}}{\nu[2(t-\eta)-(\tau-\eta)]} \right)\left(\frac{|x-y|}{3\rho} \right)}.
$$

In the last step, we used that since $s \in [1/2, 1)$:

$$
[2(t-\eta)-(\tau-\eta)] \le c(s, T)(t-\eta)^{\frac{1}{2s}} \le \frac{c(s, T)}{\nu^{\frac{1}{2s}}}\rho \le \frac{c(s, T)}{\nu^{\frac{1}{2s}}}|x-y|,
$$

where $c(s, T) > 0$ is a constant. This proves the desired result upon choosing $\nu > 1$ accordingly. $\square$

Next, we apply Lemma 5.3 to H. This yields the following lemma:

Lemma 5.7. *Let $y \in \mathbb{R}^d$, $\sigma, \rho > 0$ and $0 \le \eta \le T$. Let $u_0 \in L^2(\mathbb{R}^d)$ be such that $u_0 \equiv 0$ in $B_\sigma(y)$. Assume that $u \in L^\infty((\eta, T) \times \mathbb{R}^d)$ is a weak solution to $\partial_t u \pm (b, \nabla u) + \mathcal{L}_t^\rho u = 0$ in $(\eta, T) \times \mathbb{R}^d$ with $u(\eta) = u_0$. Then there exist*

$\nu > 1, C > 0$, *depending only on* d, s, Λ, T, *such that for every* $t \in (\eta, T)$ *with* $t - \eta \leq \frac{1}{4\nu}\rho^{2s}$:

$$|u(t, y)| \leq C(t - \eta)^{-\frac{d}{4s}} 2^{\frac{\sigma}{6\rho}} \left(\frac{\rho^{2s}}{\nu(t - \eta)}\right)^{-\frac{\sigma}{6\rho} + \frac{1}{2} + \frac{d}{4s}} \|u_0\|_{L^2(\mathbb{R}^d)}.$$

Proof. We apply Lemma 5.3 with H as in (5.12). This is possible due to Lemma 5.6. It follows that for every $y \in \mathbb{R}^d$, $0 \leq \eta \leq t \leq T$ with $t - \eta \leq \frac{1}{4\nu}\rho^{2s}$, applying also Lemma 5.2 with $R = (t - \eta)^{\frac{1}{2s}}$, $t_0 = t$, $x_0 = y$:

$$|u(t, y)| \leq c_1(t - \eta)^{-\frac{d}{4s}} \left(\frac{\rho^{2s}}{t - \eta}\right)^{\frac{d}{4s}} \sup_{\tau \in (\eta, t)} \left(\int_{B_{2\rho}(y)} u^2(\tau, x) \mathrm{d}x\right)^{1/2}$$

$$\leq c_2(t - \eta)^{-\frac{d}{4s}} \left(\frac{\rho^{2s}}{t - \eta}\right)^{\frac{d}{4s}} \left(\frac{\sup_{x \in \mathbb{R}^d \setminus B_\sigma(y)} H(\eta, x)}{\inf_{\tau \in (\eta, t), x \in B_{2\rho}(y)} H(\tau, x)}\right)^{1/2} \|u_0\|_{L^2(\mathbb{R}^d)}$$

for some constants $c_1, c_2 > 0$.

Note that there exist $c_3, c_4 > 0$ such that for $x \in \mathbb{R}^d \setminus B_\sigma(y)$ it holds

$$H(\eta, x) \leq c_3 \left(\frac{\rho^{2s}}{2\nu(t - \eta)}\right)^{-\frac{\sigma}{3\rho}}$$

and for $(\tau, x) \in [\eta, t] \times B_{2\rho}(y)$ we have:

$$H(\tau, x) \geq c_4 \left(\frac{\rho^{2s}}{\nu(t - \eta)}\right)^{-1}.$$

This follows directly from the definition of H and $t - \eta \leq \frac{1}{4\nu}\rho^{2s}$. Together, we obtain

$$|u(t, y)| \leq c_5(t - \eta)^{-\frac{d}{4s}} 2^{\frac{\sigma}{6\rho}} \left(\frac{\rho^{2s}}{\nu(t - \eta)}\right)^{-\frac{\sigma}{6\rho} + \frac{1}{2} + \frac{d}{4s}} \|u_0\|_{L^2(\mathbb{R}^d)}$$

for some constant $c_5 > 0$, as desired. $\square$

Now, we are in the position to deduce off-diagonal heat kernel estimates for the truncated heat kernel p^ρ:

Proof of Proposition 5.1. First, observe the following on-diagonal bound for p^ρ, which follows from (5.5) combined with the second estimate in Lemma 5.4:

(5.13) $$p^\rho(t, x, y) \leq e^{C(t-\eta)\rho^{-2s}}(t - \eta)^{-\frac{d}{2s}}.$$

The on-diagonal bound (5.13) and (5.4) immediately imply for every $0 \leq \eta \leq t \leq T$ with $t - \eta \leq \frac{1}{4\nu}\rho^{2s}$ and $x, y \in \mathbb{R}^d$:

(5.14) $$\left(\int_{\mathbb{R}^d} p^\rho(\eta, x; t, z)^2 \mathrm{d}z\right)^{1/2} \leq c_1(t - \eta)^{-\frac{d}{4s}},$$

(5.15) $$\left(\int_{\mathbb{R}^d} p^\rho(\eta, z; t, x)^2 \mathrm{d}z\right)^{1/2} \leq c_1(t - \eta)^{-\frac{d}{4s}}$$

for some $c_1 > 0$. On the other hand, from Lemma 5.7, it follows for every $0 \leq \eta \leq t \leq T$ with $t - \eta \leq \frac{1}{4\nu}\rho^{2s}$ and $y \in \mathbb{R}^d$:

(5.16)
$$\left(\int_{\mathbb{R}^d \setminus B_\sigma(y)} p^\rho(\eta, y; t, z)^2 \mathrm{d}z\right)^{1/2} \leq c_2(t-\eta)^{-\frac{d}{4s}} 2^{\frac{\sigma}{6\rho}} \left(\frac{\rho^{2s}}{\nu(t-\eta)}\right)^{-\frac{\sigma}{6\rho}+\frac{1}{2}+\frac{d}{4s}},$$

(5.17)
$$\left(\int_{\mathbb{R}^d \setminus B_\sigma(y)} p^\rho(\eta, z; t, y)^2 \mathrm{d}z\right)^{1/2} \leq c_2(t-\eta)^{-\frac{d}{4s}} 2^{\frac{\sigma}{6\rho}} \left(\frac{\rho^{2s}}{\nu(t-\eta)}\right)^{-\frac{\sigma}{6\rho}+\frac{1}{2}+\frac{d}{4s}}$$

for some $c_2 > 0$. To prove (5.16), one observes that

$$u(\tau, x) = \int_{\mathbb{R}^d \setminus B_\sigma(y)} p^\rho(\eta, y; t, z) p^\rho(\eta, x; \tau, z) \mathrm{d}z$$

solves $\partial_t u + (b, \nabla u) + \mathcal{L}_t^\rho u = 0$ and satisfies the assumptions of Lemma 5.7 with

$$u_0(x) = p^\rho(\eta, y; t, x) \mathbb{1}_{\{|x-y|>\sigma\}}(x), \qquad u(t, y) = \int_{\mathbb{R}^d \setminus B_\sigma(y)} p^\rho(\eta, y; t, z)^2 \mathrm{d}z.$$

From here, (5.17) follows by the same arguments, choosing

$$\widehat{u}(\tau, x) = \int_{\mathbb{R}^d \setminus B_\sigma(y)} p^\rho(\eta, z; t, y) p^\rho(\eta, z; \tau, x) \mathrm{d}z,$$

which solves the dual equation $\partial_t u - (b, \nabla u) + \mathcal{L}_t^\rho u = 0$ with $\widehat{u}_0(x) = p^\rho(\eta, x; t, y) \mathbb{1}_{\{|x-y|>\sigma\}}(x)$.

Note that boundedness of u and $\widehat{u}$, which is required in order to apply Lemma 5.7, is a direct consequence of the on-diagonal bound (5.13) and integrability of the truncated heat kernel, which follows from (5.4) and the second property in Lemma 5.4.

To prove (5.11), let us fix $0 \leq \eta \leq t \leq T$ with $t - \eta \leq \frac{1}{4\nu}\rho^{2s}$, and $x, y \in \mathbb{R}^d$. Then we define $\sigma = \frac{1}{2}|x - y|$ and compute, using the semigroup property:

$$p^\rho(\eta, x; t, y) = \int_{\mathbb{R}^d} p^\rho(\eta, x; (t-\eta)/2, z) p^\rho((t-\eta)/2, z; t, y) \mathrm{d}z$$

$$= \int_{\mathbb{R}^d \setminus B_\sigma(y)} p^\rho(\eta, x; (t-\eta)/2, z) p^\rho((t-\eta)/2, z; ty) \mathrm{d}z$$

$$+ \int_{B_\sigma(y)} p^\rho(\eta, x; (t-\eta)/2, z) p^\rho((t-\eta)/2, z; ty) \mathrm{d}z$$

$$=: J_1 + J_2.$$

For J_1, we compute, using (5.14), (5.17)

$$J_1 \leq \left(\int_{\mathbb{R}^d \setminus B_\sigma(y)} p^\rho(\eta, x; (t-\eta)/2, z)^2 \mathrm{d}z\right)^{1/2} \left(\int_{\mathbb{R}^d \setminus B_\sigma(y)} p^\rho((t-\eta)/2, z; t, y)^2 \mathrm{d}z\right)^{1/2}$$

$$\leq c_3(t-\eta)^{-\frac{d}{2s}} 2^{\frac{|x-y|}{12\rho}} \left(\frac{\rho^{2s}}{\nu(t-\eta)}\right)^{-\frac{|x-y|}{12\rho}+\frac{1}{2}+\frac{d}{4s}}$$

for some $c_3 > 0$. For J_2, observe that $B_\sigma(y) \subset \mathbb{R}^d \setminus B_\sigma(x)$, and therefore by (5.15), (5.16):

$$J_2 \leq \left(\int_{\mathbb{R}^d \setminus B_\sigma(x)} p^\rho(\eta, x; (t-\eta)/2, z)^2 \mathrm{d}z \right)^{1/2} \left(\int_{\mathbb{R}^d \setminus B_\sigma(x)} p^\rho((t-\eta)/2, z; t, y)^2 \mathrm{d}z \right)^{1/2}$$

$$\leq c_3 (t-\eta)^{-\frac{d}{2s}} 2^{\frac{|x-y|}{12\rho}} \left(\frac{\rho^{2s}}{\nu(t-\eta)} \right)^{-\frac{|x-y|}{12\rho} + \frac{1}{2} + \frac{d}{4s}}.$$

Together, we obtain the desired result. $\qquad\qquad\square$

5.4. Proof of the upper heat kernel estimate. Now, we are in the position to prove Theorem 1.3:

Proof of Theorem 1.3. Let $x, y \in \mathbb{R}^d$ be fixed. By the on-diagonal estimate (5.5) it suffices to show that for some constants $c_0, c_1, c_2 > 0$ and $t - \eta \leq c_0 |x - y|^{2s}$ it holds

$$p(\eta, x; t, y) \leq c_1 \frac{t-\eta}{|x-y|^{d+2s}}.$$

By Lemma 5.4, we know that for every $\rho > 0$ and $0 \leq \eta \leq t \leq T$ with $t - \eta \leq \frac{1}{4\nu} \rho^{2s}$:

$$(5.18) \qquad p(\eta, x; t, y) \leq p^\rho(\eta, x; t, y) + C(t-\eta)\rho^{-d-2s}$$

for some $C > 0$. We choose $\rho = \frac{|x-y|}{12} \left(\frac{d+2s}{2s} + \frac{1}{2} + \frac{d}{4s} \right)^{-1}$ and define $c_2 = 12 \left(\frac{d+2s}{2s} + \frac{1}{2} + \frac{d}{4s} \right)$. Then by combination of (5.11) and (5.18), we deduce for $t - \eta \leq \frac{1}{4\nu} \rho^{2s}$:

$$p(\eta, x; t, y) \leq c_1 (t-\eta) |x-y|^{-d-2s},$$

where $c_1 > 0$. This proves the desired result. $\qquad\qquad\square$

ACKNOWLEDGMENTS

REFERENCES

[APT22] Karthik Adimurthi, Harsh Prasad, and Vivek Tewary. Hölder regularity for fractional p-Laplace equations. *arXiv:2203.13082*, 2022.

[Aro68] Donald Aronson. Non-negative solutions of linear parabolic equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 22(4):607–694, 1968.

[BCD+18] Dominic Breit, Andrea Cianchi, Lars Diening, Tuomo Kuusi, and Sebastian Schwarzacher. Pointwise Calderón-Zygmund gradient estimates for the *p*-Laplace system. *J. Math. Pures Appl. (9)*, 114:146–190, 2018.

[BDGO97] Lucio Boccardo, Andrea Dall'Aglio, Thierry Gallouët, and Luigi Orsina. Nonlinear parabolic equations with measure data. *J. Funct. Anal.*, 147(1):237–258, 1997.

[BK23] Sun-Sig Byun and Kyeongbae Kim. Hölder estimate with an optimal tail for nonlocal parabolic p-Laplace equations. 2023. To appear in Ann. Mat. Pura Appl.

[BLS21] Lorenzo Brasco, Erik Lindgren, and Martin Strömqvist. Continuity of solutions to a nonlinear fractional diffusion equation. *J. Evol. Equ.*, 21(4):4319–4381, 2021.

[BY19] Sun-Sig Byun and Yeonghun Youn. Potential estimates for elliptic systems with subquadratic growth. *J. Math. Pures Appl. (9)*, 131:193–224, 2019.

[CCV11] Luis Caffarelli, Chi Hin Chan, and Alexis Vasseur. Regularity theory for parabolic nonlinear integral operators. *J. Amer. Math. Soc.*, 24(3):849–869, 2011.

[Cia11] Andrea Cianchi. Nonlinear potentials, local solutions to elliptic equations and rearrangements. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 10(2):335–361, 2011.

[CKS87] Eric Carlen, Shigeo Kusuoka, and Daniel W Stroock. Upper bounds for symmetric Markov transition functions. In *Ann. Inst. H. Poincaré. Probab. Statist.*, volume 23, pages 245–287, 1987.

[CMT94] Peter Constantin, Andrew Majda, and Esteban Tabak. Formation of strong fronts in the 2-D quasigeostrophic thermal active scalar. *Nonlinearity*, 7(6):1495–1533, 1994.

[Coz17] Matteo Cozzi. Regularity results and Harnack inequalities for minimizers and solutions of nonlocal problems: a unified approach via fractional De Giorgi classes. *J. Funct. Anal.*, 272(11):4762–4837, 2017.

[CV10] Luis Caffarelli and Alexis Vasseur. Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation. *Ann. of Math. (2)*, 171(3):1903–1930, 2010.

[Dav87] E. Davies. Explicit constants for Gaussian upper bounds on heat kernels. *Amer. J. Math.*, 109(2):319–333, 1987.

[Dav89] E. Davies. *Heat kernels and spectral theory*. Number 92. Cambridge university press, 1989.

[DCKP16] Agnese Di Castro, Tuomo Kuusi, and Giampiero Palatucci. Local behavior of fractional *p*-minimizers. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 33(5):1279–1299, 2016.

[DF22] Cristiana De Filippis. Quasiconvexity and partial regularity via nonlinear potentials. *J. Math. Pures Appl. (9)*, 163:11–82, 2022.

[DiB93] Emmanuele DiBenedetto. *Degenerate parabolic equations*. Springer Science & Business Media, 1993.

[DM11] Frank Duzaar and Giuseppe Mingione. Gradient estimates via non-linear potentials. *Amer. J. Math.*, 133(4):1093–1149, 2011.

[DN23] Lars Diening and Simon Nowak. Calderón-Zygmund estimates for the fractional *p*-Laplacian. *arXiv:2111.05768*, 2023.

[DS18] Matías Delgadino and Scott Smith. Hölder estimates for fractional parabolic equations with critical divergence free drifts. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 35(3):577–604, 2018.

[DZ21] Hongjie Dong and Hanye Zhu. Gradient estimates for singular *p*-Laplace type equations with measure data. *arXiv:2102.08584; to appear in J. Eur. Math. Soc.*, 2021.

[FK13] Matthieu Felsinger and Moritz Kassmann. Local regularity for parabolic nonlocal operators. *Comm. Partial Differential Equations*, 38(9):1539–1573, 2013.

[Kas09] Moritz Kassmann. A priori estimates for integro-differential operators with measurable kernels. *Calc. Var. Partial Differential Equations*, 34(1):1–21, 2009.

[Kis10]   A. Kiselev. Regularity and blow up for active scalars. *Math. Model. Nat. Phe-nom.*, 5(4):225–255, 2010.

[KM94]   Tero Kilpeläinen and Jan Malý. The Wiener test and potential estimates for quasilinear elliptic equations. *Acta Math.*, 172(1):137–161, 1994.

[KM14]   Tuomo Kuusi and Giuseppe Mingione. Riesz potentials and nonlinear parabolic equations. *Arch. Ration. Mech. Anal.*, 212(3):727–780, 2014.

[KM18]   Tuomo Kuusi and Giuseppe Mingione. Vectorial nonlinear potential theory. *J. Eur. Math. Soc.*, 20(4):929–1004, 2018.

[KMS15]   Tuomo Kuusi, Giuseppe Mingione, and Yannick Sire. Nonlocal equations with measure data. *Comm. Math. Phys.*, 337(3):1317–1368, 2015.

[KMS18]   Tuomo Kuusi, Giuseppe Mingione, and Yannick Sire. Regularity issues involving the fractional p-Laplacian. In *Recent developments in nonlocal theory*, pages 303–334. De Gruyter, Berlin, 2018.

[KN09]   A. Kiselev and F. Nazarov. A variation on a theme of Caffarelli and Vasseur. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 370(Kraevye Zadachi Matematicheskoǐ Fiziki i Smezhnye Voprosy Teorii Funktsiǐ. 40):58–72, 220, 2009.

[KNS22]   Tuomo Kuusi, Simon Nowak, and Yannick Sire. Gradient regularity and first-order potential estimates for a class of nonlocal equations. *arXiv:2212.01950*, 2022.

[KNV07]   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.

[KW22a]   Moritz Kassmann and Marvin Weidner. Nonlocal operators related to non-symmetric forms I: Hölder estimates. *arXiv:2203.07418*, 2022.

[KW22b]   Moritz Kassmann and Marvin Weidner. Nonlocal operators related to nonsymmetric forms II: Harnack inequalities. *arXiv:2205.05531, to appear in Anal. PDE*, 2022.

[KW23a]   Moritz Kassmann and Marvin Weidner. The parabolic Harnack inequality for nonlocal equations. *arXiv:2303.05975*, 2023.

[KW23b]   Moritz Kassmann and Marvin Weidner. Upper heat kernel estimates for nonlocal operators via Aronson's method. *Calc. Var. Partial Differential Equations*, 62(2):68, 2023.

[Lia22]   Naian Liao. Hölder regularity for parabolic fractional p-Laplacian. *arXiv:2205.10111*, 2022.

[Min11]   Giuseppe Mingione. Gradient potential estimates. *J. Eur. Math. Soc.*, 13(2):459–486, 2011.

[MM13a]   Yasunori Maekawa and Hideyuki Miura. On fundamental solutions for nonlocal parabolic equations with divergence free drift. *Adv. Math.*, 247:123–191, 2013.

[MM13b]   Yasunori Maekawa and Hideyuki Miura. Upper bounds for fundamental solutions to non-local diffusion equations with divergence free drift. *J. Funct. Anal.*, 264(10):2245–2268, 2013.

[Ngu14]   Quoc-Hung Nguyen. Potential estimates and quasilinear parabolic equations with measure data. *arXiv:1405.2587, to appear in Mem. Amer. Math. Soc.*, 2014.

[NST23]   Quoc-Hung Nguyen, Yannick Sire, and Le Xuan Truong. Hölder continuity of solutions for a class of drift-diffusion equations. *Discrete Contin. Dyn. Syst.*, 43(3-4):1657–1685, 2023.

[Sil11]   Luis Silvestre. On the differentiability of the solution to the Hamilton-Jacobi equation with critical fractional diffusion. *Adv. Math.*, 226(2):2020–2039, 2011.

[Sil12a]   Luis Silvestre. Hölder estimates for advection fractional-diffusion equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 11(4):843–855, 2012.

[Sil12b]   Luis Silvestre. On the differentiability of the solution to an equation with drift and fractional diffusion. *Indiana Univ. Math. J.*, 61(2):557–584, 2012.

[TW02]   Neil Trudinger and Xu-Jia Wang. On the weak continuity of elliptic operators and applications to potential theory. *Amer. J. Math.*, 124(2):369–410, 2002.

Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, PR China.
Email address: qhnguyen@amss.ac.cn

Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany
Email address: simon.nowak@uni-bielefeld.de

Department of Mathematics, Johns Hopkins University, Baltimore, MD, 21218
Email address: ysire1@jhu.edu

Departament de Matemàtiques i Informàtica, Universitat de Barcelona, Gran Via de les Corts Catalanes 585, 08007 Barcelona, Spain
Email address: mweidner@ub.edu