Anisotropic nonlocal diffusion equations with singular forcing

Arturo de Pablo [a], Fernando Quirós [b,*], Ana Rodríguez [c]

[a] *Departamento de Matemáticas, Universidad Carlos III de Madrid, 28911 Leganés, Spain*
[b] *Departamento de Matemáticas, Universidad Autónoma de Madrid, 28049 Madrid, Spain*
[c] *Departamento de Matemática Aplicada, Universidad Politécnica de Madrid, 28040 Madrid, Spain*

Abstract

We prove existence, uniqueness and regularity of solutions of nonlocal heat equations associated to anisotropic stable diffusion operators. The main features are that the right-hand side has very little regularity and that the spectral measure can be singular in some directions. The proofs require having good enough estimates for the corresponding heat kernels and their derivatives.
© 2020 L'Association Publications de l'Institut Henri Poincaré. Published by Elsevier B.V. All rights reserved.

1. Introduction and main results

The aim of this paper is to study existence, uniqueness and regularity of solutions u to a nonlocal parabolic problem with a nonstandard forcing term,

$$\begin{cases} \partial_t u + \mathcal{L}u = \mathcal{L}f & \text{in } Q := \mathbb{R}^N \times \mathbb{R}_+, \\ u(\cdot, 0) = u_0 & \text{in } \mathbb{R}^N, \end{cases} \tag{1.1}$$

where $f = f(x, t)$ is bounded and Hölder continuous, u_0 is in some L^p space, $1 \le p \le \infty$, and $\mathcal{L}$ is a pseudo-differential operator corresponding to a symmetric stable process of order $\sigma \in (0, 2)$. In general the right-hand side $\mathcal{L}f$ is singular since it is only known to be a distribution in some negative Hölder space. As a consequence, u is not an energy solution. Thus, though the equation is linear and invariant under translations and scalings, de Giorgi or Moser-like approaches to regularity cannot be applied.

---

* Corresponding author.
E-mail addresses: arturo.depablo@uc3m.es (A. de Pablo), fernando.quiros@uam.es (F. Quirós), ana.rodriguez@upm.es (A. Rodríguez).

Parabolic equations with singular forcing terms have attracted a lot of attention in recent years; see for instance [9, 12]. Problem (1.1) has already been considered in [13,17] in the special case in which $\mathcal{L}$ is the fractional Laplacian as a tool to obtain higher regularity for nonlinear problems of the form $\partial_t u + (-\Delta)^{\sigma/2}\phi(u) = 0$. Hence, we expect that the results in this paper will allow to obtain higher regularity for nonlinear problems associated to more general nonlocal operators.

The nonlocal operator $\mathcal{L}$ is defined, for regular functions which do not grow too much at infinity, by

$$\mathcal{L}u(x) = \int_{\mathbb{R}^N} \left(u(x) - \frac{u(x+y) + u(x-y)}{2} \right) d\nu(y), \tag{1.2}$$

where the nonnegative Lévy measure ν has the polar decomposition

$$d\nu(y) = \frac{d(|y|)}{|y|^{N+\sigma}} d\mu\left(\frac{y}{|y|} \right), \qquad \text{with } 0 < \sigma < 2. \tag{1.3}$$

General operators of the form (1.2)–(1.3) arise as the infinitesimal generators of symmetric stable Lévy processes $X = \{X_t\}_{t\geq 0}$, which satisfy

$$\lambda X_t = X_{\lambda^\sigma t}, \qquad \lambda > 0, \ t \geq 0.$$

These processes appear in Physics, Mathematical Finance and Biology, among other applications, and have been the subject of intensive research in the last years from the point of view both of Probability and Analysis; see for instance the survey [15] and the references therein.

The measure μ on the sphere $\mathbb{S}^{N-1}$, called the *spectral measure*, is assumed to be finite, $\mu(\mathbb{S}^{N-1}) = \Lambda < \infty$, and to satisfy the "ellipticity" (non-degeneracy) condition

$$\inf_{\zeta \in \mathbb{S}^{N-1}} \int_{\mathbb{S}^{N-1}} |\zeta \cdot \theta|^\sigma \, d\mu(\theta) \geq \lambda > 0. \tag{1.4}$$

That is, we require that the spectral measure is not supported in any proper subspace of $\mathbb{R}^N$. We remark that we do not impose any symmetry to the measure μ, since symmetry of the operator (and thus of the process) comes directly from the way we write it in (1.2) using second differences. If μ were symmetric the operator would take the more familiar form

$$\mathcal{L}u(x) = P.V. \int_{\mathbb{R}^N} (u(x) - u(x+y)) \, d\nu(y).$$

By (1.3), the operator $\mathcal{L}$ can be expressed in the form

$$\mathcal{L}u(x) = \int_{\mathbb{S}^{N-1}} \int_0^\infty \left(u(x) - \frac{u(x+r\theta) + u(x-r\theta)}{2} \right) \frac{dr}{r^{1+\sigma}} \, d\mu(\theta). \tag{1.5}$$

The spectral measure is allowed to be *anisotropic*. Hence, we cannot use radial arguments as the ones employed in [17] to deal with the isotropic case $d\mu(\theta) = d\theta$, for which the operator reduces to (a multiple of) the well known fractional Laplacian, $\mathcal{L} = (-\Delta)^{\sigma/2}$. Note, however, that our anisotropic operator is still homogeneous of order σ, which will turn out to be an important tool in our proofs. We also allow the Lévy measure to be singular in some directions. It may even be concentrated on a set of directions of Lebesgue measure zero, as in example (1.8).

The functions u and f in equation (1.1) do not possess in general the required regularity to give $\mathcal{L}u$ and $\mathcal{L}f$ a pointwise meaning. Moreover, u will not even belong to the energy space associated to the operator. We therefore have to work with solutions in a very weak sense; see formulas (3.1) and (4.1) below. Our first result shows that the problem is well-posed in this very weak formulation.

Theorem 1.1. *If $f \in C^\alpha(Q) \cap L^\infty(Q)$ for some $0 < \alpha < 1$, and $u_0 \in L^p(\mathbb{R}^N)$ for some $1 \leq p \leq \infty$, then there exists a unique very weak solution of problem* (1.1)*, which is moreover bounded for positive times.*

The solution is given explicitly by Duhamel's type formula

$$u(x,t) = \int\limits_{\mathbb{R}^N} P(x - \overline{x}, t) u_0(\overline{x}) \, d\overline{x} + \int\limits_0^t \int\limits_{\mathbb{R}^N} \mathcal{L}P(x - \overline{x}, t - \overline{t}) f(\overline{x}, \overline{t}) \, d\overline{x} d\overline{t}, \tag{1.6}$$

where P is the fundamental solution for the homogeneous problem. Note that in the second term the operator is applied to P, in contrast with the case in which the right-hand side is not singular. Since P is smooth for positive times, the first term, corresponding to the initial datum, is smooth. On the contrary, giving a meaning to the second integral in (1.6) in some principal value sense, see (4.5), requires some effort, since $\mathcal{L}P$ has a nonintegrable singularity at $(x - \bar{x}, t - \bar{t}) = (0, 0)$. However, thanks to the homogeneity of the operator, the kernel and its derivatives are known to have a self-similar structure. In particular, $P(x,t) = t^{-\frac{N}{\sigma}} \Phi(xt^{-\frac{1}{\sigma}})$, $\mathcal{L}P(x,t) = t^{-\frac{N}{\sigma}-1} \mathcal{L}\Phi(xt^{-\frac{1}{\sigma}})$ for some positive profile Φ. The singularity of $\mathcal{L}P$ at the origin can then be controlled in terms of the decay of the profile Φ at infinity, which is next combined with the Hölder regularity of f to obtain integrability at the origin. The same decay estimates account for integrability at space infinity.

In the isotropic case, $\mathcal{L} = (-\Delta)^{\sigma/2}$, we have the pointwise estimates

$$\Phi(r) \sim r^{-N-\sigma}, \quad |\mathcal{L}\Phi(r)| \le cr^{-N-\sigma} \quad \text{for } r \text{ large}, \tag{1.7}$$

which have been known for a long time [2]. However, for anisotropic processes the pointwise decay can be much slower in some directions, as observed in [14] for the case

$$\mathcal{L} = \sum_{j=1}^N (-\partial_{x_j x_j})^{\sigma/2}; \tag{1.8}$$

see also Section 2.2. Nevertheless, the decay for Φ *on average*

$$\int\limits_{\mathbb{S}^{N-1}} \Phi(r\theta) \, d\theta \le cr^{-N-\sigma} \tag{1.9}$$

holds for all stable processes [14, Theorem 2]. On the other hand, for any $k \in \mathbb{N}$ and $\beta \in (0, \sigma)$ there is a function $\Omega_{k,\beta} \in L^1(\mathbb{S}^{N-1})$ such that

$$|\mathcal{L}^k \Phi(r\theta)| \le \Omega_{k,\beta}(\theta) r^{-N-\beta}; \tag{1.10}$$

see [7, Theorem 5.1]. Hence,

$$\int\limits_{\mathbb{S}^{N-1}} |\mathcal{L}\Phi(r\theta)| \, d\theta \le c_\beta r^{-N-\beta} \quad \text{for every } \beta < \sigma. \tag{1.11}$$

The estimates on average (1.9) and (1.11) suffice to show the well-posedness of our problem.

Remark. When $k = 0$ an estimate like (1.10) with $\beta = \sigma$ is only true if the spectral measure is absolutely continuous and its density belongs to a certain integrability class [8]. However, the threshold decay is reached on average; see (1.9). We expect to have also such limit decay on average for $k \ge 1$.

In order to prove that u is Hölder continuous we cannot use de Giorgi or Moser approaches, as done for instance in [5,10], since the solution does not lie in general in the energy space. Hence we have chosen a different approach, that requires to estimate further derivatives of $\mathcal{L}P$, a subject that has independent interest. The pointwise estimate (1.10) provides the required decay on average for $\mathcal{L}^2\Phi$. This corresponds to estimating $\partial_t \mathcal{L}P$. A similar result can be obtained for the radial derivatives, using the equation satisfied by the profile. But this is not enough to estimate the standard spatial derivatives, which involve variations of angles. Hence we make the following extra assumption on the behaviour of $\mathcal{L}\Phi$ on average,

$$\int\limits_{\mathbb{S}^{N-1}} |\mathcal{L}\Phi(r\theta) - \mathcal{L}\Phi(r\theta - s\varphi)| \, d\theta \le c\delta r^{-N-\beta} \quad \text{for some } \beta > 0, \tag{1.12}$$

whenever $0 < s < \delta \leq 1$, $\varphi \in \mathbb{S}^{N-1}$, $r \geq 1$, which roughly speaking means estimating $\nabla \mathcal{L} P$ on average. The required smoothness for the function f will be given in terms of a topology adapted to the scaling of the equation, through the Hölder spaces $C_\sigma^\alpha(Q)$ defined in Section 4.1.

Theorem 1.2. *Assume that the profile Φ of the fundamental solution for the operator $\partial_t + \mathcal{L}$ satisfies (1.12) for some $\beta > 0$. If u is a very weak solution to problem (1.1) with $f \in C_\sigma^\alpha(Q) \cap L^\infty(Q)$ for some $0 < \alpha < \min\{\beta, \sigma, 1\}$, then $u \in C_\sigma^\alpha(Q)$.*

A slight modification of the proof of this result allows to improve the regularity of the solution at each point where f is more regular, which is stated in Theorem 4.3. This will be used in a separate work to prove regularity for the nonlinear equation $\partial_t u + \mathcal{L}\phi(u) = 0$; see [17] for the case of the fractional Laplacian.

The key hypothesis (1.12) holds, with $\beta = \sigma$, in the important special case of stable operators given by sums of fractional Laplacians (of order σ) of smaller dimensions

$$\mathcal{L} = \sum_{k=1}^M (-\Delta_{x^k})^{\sigma/2}, \qquad x^k \in \mathbb{R}^{n_k}, \quad \sum_{k=1}^M n_k = N, \quad 1 \leq M \leq N, \tag{1.13}$$

the simplest example being (1.8); see Section 2.2. In this special situation we also improve estimate (1.11) to include the critical value $\beta = \sigma$.

Theorem 1.3. *Conditions (1.11) and (1.12) with $\beta = \sigma$ hold for operators $\mathcal{L}$ of the form (1.13).*

The proof follows by observing that in this case the derivatives of the kernel P can be estimated by P itself, combined with estimate (1.9). We conjecture that this property is true for the profile of any stable process. In the case of isotropic processes, even depending on time, such estimates for the spatial derivatives have been obtained in [11].

Corollary 1.1. *Let $\mathcal{L}$ be given by (1.13). If u is a very weak solution to problem (1.1) with $f \in C_\sigma^\alpha(Q) \cap L^\infty(Q)$ for some $0 < \alpha < \min\{\sigma, 1\}$, then $u \in C_\sigma^\alpha(Q)$.*

The obtention of estimates for the derivatives of heat kernels of stable Lévy processes has been the subject of intensive research in the last years. To this aim an auxiliary smoothness scale of Hausdorff-type, which we describe next, was introduced in [3]; see also [18]. A measure μ on $\mathbb{S}^{N-1}$ is said to be a γ–measure if there is a constant $c > 0$ such that

$$\mu(B(\theta, r) \cap \mathbb{S}^{N-1}) \leq cr^{\gamma-1} \quad \text{for all } \theta \in \mathbb{S}^{N-1} \text{ and } 0 < r < 1/2.$$

It is easy to see that necessarily $\gamma \leq N$ and that any finite measure is at least a 1–measure. The case $\gamma = N$ holds if and only if μ is absolutely continuous with respect to the Lebesgue measure and has a density function which is bounded. This does not mean that the measure is comparable to that of the isotropic case, since it may degenerate in some directions. If $\gamma > 1$, the measure has no atoms. If $\gamma < N$, it is singular. For instance, a spectral measure satisfying $d\mu(\theta) = a(\theta)d\theta$ where $a(\theta)$ has a singularity of the form $a(\theta) \sim |\theta - \theta_0|^{\gamma-N}$, $\gamma \in (1, N)$, is a γ–measure.

When the spectral measure μ is a γ–measure with $\sigma + \gamma - N > 0$, it was proved in [4, Lemma 2.7] that

$$|\nabla \mathcal{L} \Phi(x)| \leq C(1 + |x|^2)^{-(\sigma+\gamma)/2}, \tag{1.14}$$

which implies the estimate on average (1.12) for $\beta = \sigma + \gamma - N$. Thus, when μ is an N–measure, the only restriction in Theorem 1.2 is $\alpha < \min\{\sigma, 1\}$, as in the case in which $\mathcal{L} = (-\Delta)^{\sigma/2}$ studied in [17].

Let us remark that, though the pointwise estimate (1.14) is optimal, the integral version that is derived from it seems far from being so if $\gamma < N$, since it does not take into account that the measure of the set of directions in which the derivatives decay slowly is small. Thus, we get a restriction on β, which cannot be arbitrarily close to σ. We believe that this restriction is technical, since it does not appear in the "worst" case (1.8), for which $\gamma = 1$; see Corollary 1.1.

When the right-hand side is standard, $\partial_t u + \mathcal{L}u = g$, with $g \in L^\infty(Q) \cap C_\sigma^\alpha(Q)$, $\alpha < \min\{\sigma, 1\}$, very weak solutions satisfy $\partial_t u$, $\mathcal{L}u \in C_\sigma^\alpha(Q)$. This was proved through a blowup argument combined with a Liouville type theorem in [6]. This result follows from ours whenever hypothesis (1.12) holds for every $\beta < \sigma$, which is the case when the operator

is given by (1.13), or when it comes from an N–measure. Let us emphasize that the result in [6] holds for general stable operators, without any restriction on the spectral measure. However, the blowup argument used there requires some regularity of the right-hand side term, which is not available for problem (1.1). This is in fact the main difficulty in the present work.

It is also worth mentioning the papers [5,10], where the authors show Hölder regularity when $g = 0$ for a class of operators $\mathcal{L}$, which are not necessarily translation invariant, that include the special case (1.8). Though their proof may perhaps be adapted to consider $g \in L^\infty(Q)$, it assumes that the solution lies in the energy space, and hence cannot be used to deal with solutions of problem (1.1) when f is not smooth enough.

Observe finally that if f depends only on x or only on t then problem (1.1) becomes trivial. In the application to nonlinear problems that we have in mind the right-hand sides that arise depend typically on both variables.

ORGANIZATION OF THE PAPER. We start with the discussion of the required estimates for the kernel and its derivatives in Section 2, devoting a separate subsection to the special case of operators of the form (1.13). Section 3 deals with the homogeneous case, $f = 0$, which yields uniqueness also when the right-hand side is nontrivial. Finally, we consider the problem with a singular forcing in Section 4, proving existence and regularity.

2. Properties of the heat kernel

The aim of this section is to obtain estimates for the heat kernel and its derivatives allowing to apply Theorem 1.2 to some families of stable operators. We start by describing estimates which are valid for general stable operators, and pass then to consider the case of operators of the form (1.13), for which much better estimates are available.

2.1. General stable operators

Taking Fourier transform in (1.5) we get that the multiplier m of the operator $\mathcal{L}$, defined by $\widehat{\mathcal{L}u}(\xi) = m(\xi)\widehat{u}(\xi)$, satisfies

$$
\begin{aligned}
m(\xi) &= \int_{\mathbb{S}^{N-1}} \int_0^\infty (1 - \cos(\xi \cdot (r\theta))) \frac{dr}{r^{1+\sigma}} \, d\mu(\theta) = |\xi|^\sigma g(\xi/|\xi|), \quad \text{where} \\
g(\zeta) &= c_{N,\sigma} \int_{\mathbb{S}^{N-1}} |\zeta \cdot \theta|^\sigma \, d\mu(\theta), \quad \text{and} \\
c_{N,\sigma} &= \int_0^\infty (1 - \cos t) \frac{dt}{t^{1+\sigma}} = \frac{\pi^{1/2}\Gamma(1 - \sigma/2)}{2^\sigma \sigma \Gamma((1+\sigma)/2)}.
\end{aligned}
\tag{2.1}
$$

In particular m is homogeneous of order σ and $m(\xi) \sim |\xi|^\sigma$, since by the finiteness of the measure and the non-degeneracy condition (1.4) we have

$$
\lambda c_{N,\sigma} \leq g(\zeta) \leq \Lambda c_{N,\sigma}.
\tag{2.2}
$$

The homogeneity of the multiplier implies the homogeneity of the operator,

$$
\phi(x) = u(\lambda x) \quad \Rightarrow \quad \mathcal{L}\phi(x) = \lambda^\sigma \mathcal{L}u(\lambda x).
$$

On the other hand, by [16, Theorem 2.4.3], any symmetric stable process $X = \{X_t\}_{t \geq 0}$ defined on a probability space $(\Omega, \mathbb{F}, \mathbb{P})$ has a characteristic function

$$
E[e^{i\xi \cdot X}] = e^{-m(\xi)},
$$

where $m(\xi)$ is given by the Lévy-Khintchine formula (2.1). We have therefore a one-to-one correspondence between our family of operators $\mathcal{L}$ and the family of symmetric stable processes X. If we now consider the family of probability measures $\{\mathfrak{m}_t\}_{t \geq 0}$ on $\mathbb{R}^N$, such that for every Borel set $A \subset \mathbb{R}^N$

$$
\int_A d\mathfrak{m}_t \equiv \mathbb{P}(\{\omega \in \Omega : X_t(\omega) \in A\}),
$$

we have that $d\mathfrak{m}_t = P(\cdot, t)dx$, and P satisfies the problem

$$\partial_t P + \mathcal{L}P = 0 \quad \text{in } Q, \qquad P(\cdot, 0) = \delta \quad \text{in } \mathbb{R}^N.$$

The density function P is usually known as the *transition probability density*, the *Gauss kernel* associated to $\mathcal{L}$, or the *fundamental solution* for the operator $\partial_t + \mathcal{L}$. The homogeneity of the multiplier m gives that this kernel is self-similar,

$$P(x, t) = t^{-N/\sigma} \Phi(xt^{-1/\sigma}), \qquad \widehat{\Phi}(\xi) = e^{-m(\xi)}. \tag{2.3}$$

Clearly, since $m(\xi) \geq c|\xi|^\sigma$ we have $\Phi \in C^\infty(\mathbb{R}^N)$, $0 \leq \Phi \leq \int_{\mathbb{R}^N} e^{-m(\xi)} \leq c$, and $\int_{\mathbb{R}^N} \Phi = 1$. Moreover it is also easy to see that Φ is strictly positive.

As mentioned in the Introduction, in the isotropic case $\mathcal{L} = (-\Delta)^{\sigma/2}$ the profile Φ of the kernel is radial, with a decay $\Phi(r) \sim r^{-N-\sigma}$ for r large. In the anisotropic case an estimate like the previous one is not true in general. However, as proved in [14], this rate of decay holds on average; see (1.9). Following the proof of that paper it is not difficult to obtain a decay estimate, on average, of the derivatives of Φ.

Theorem 2.1. *For any* $k, l \in \mathbb{N} \cup \{0\}$ *we have*

$$\int_{\mathbb{S}^{N-1}} \partial_r^l \mathcal{L}^k \Phi(r\theta) \, d\theta = O(r^{-N-(1+(k-1)_+)\sigma-l}) \quad \text{as } r \to \infty. \tag{2.4}$$

Proof. Consider first the case $l = 0$. Since $\widehat{\mathcal{L}^k \Phi}(\xi) = m^k(\xi)e^{-m(\xi)}$, we may write

$$\int_{\mathbb{S}^{N-1}} \mathcal{L}^k \Phi(r\theta) \, d\theta = \int_{\mathbb{S}^{N-1}} \int_{\mathbb{R}^N} m^k(\xi)e^{-m(\xi)}e^{ir\theta\cdot\xi} \, d\xi d\theta = \int_{\mathbb{R}^N} m^k(\xi)e^{-m(\xi)} \int_{\mathbb{S}^{N-1}} e^{ir\theta\cdot\xi} \, d\theta d\xi.$$

The inner integral is computed, using spherical coordinates, in [14],

$$\int_{\mathbb{S}^{N-1}} e^{ir\theta\cdot\xi} \, d\theta = c(r|\xi|)^{\frac{2-N}{2}} J_{\frac{N-2}{2}}(r|\xi|),$$

where J_ω is the Bessel function of the first kind of order ω. We thus get

$$\int_{\mathbb{S}^{N-1}} \mathcal{L}^k \Phi(r\theta) \, d\theta = cr^{\frac{2-N}{2}} \int_{\mathbb{R}^N} m^k(\xi)e^{-m(\xi)}|\xi|^{\frac{2-N}{2}} J_{\frac{N-2}{2}}(r|\xi|) \, d\xi$$

$$= cr^{\frac{2-N}{2}} \int_{\mathbb{S}^{N-1}} g^k(\eta) \int_0^\infty e^{-g(\eta)s^\sigma} s^{\frac{N}{2}+k\sigma} J_{\frac{N-2}{2}}(rs) \, ds d\eta.$$

We conclude, using [14, Lemma 1] and (2.2), the behaviour

$$\int_{\mathbb{S}^{N-1}} \mathcal{L}^k \Phi(r\theta) \, d\theta = \begin{cases} O(r^{-N-\sigma}) & \text{if } k = 0, \\ O(r^{-N-k\sigma}) & \text{if } k \geq 1. \end{cases}$$

To estimate the usual derivatives we use the equation for the profile,

$$\sigma \mathcal{L}\Phi = N\Phi + r\partial_r \Phi,$$

which differentiated gives, for each $k \geq 0$, $l \geq 0$,

$$\sigma \partial_r^l \mathcal{L}^{k+1} \Phi = (N + k\sigma + l)\partial_r^l \mathcal{L}^k \Phi + r\partial_r^{l+1} \mathcal{L}^k \Phi,$$

and obtain (2.4) by induction in l. $\quad\square$

Unfortunately the estimates needed in our regularity arguments throughout this paper require taking absolute value before taking the average. For the fractional derivatives, the pointwise estimate

$$|\mathcal{L}^k \Phi(r\theta)| \leq \Omega_{\beta,k}(\theta) r^{-N-\beta} \quad \text{a.e. } \theta \in \mathbb{S}^{N-1}, \tag{2.5}$$

was obtained in [7] for every $r \geq 1$, $k \geq 1$ and every $0 < \beta < \sigma$, where $\Omega_{\beta,k} \in L^1(\mathbb{S}^{N-1})$. This implies in particular the decay that is enough for our purposes.

As we have commented upon in the Introduction, pointwise estimates for $\nabla \mathcal{L} \Phi$ are not available, except for γ–measures with $\gamma > N - \sigma$, for which we have (1.14).

2.2. The sum of fractional Laplacians in lower dimensions

We now turn our attention to the interesting model of stable operators (1.5) of the form (1.13). Our aim is to show that condition (1.12) holds, so that Theorem 1.2 can be applied.

For the reader's convenience we perform the calculations in detail. We thus consider sums of fractional Laplacians $(-\Delta_{x^k})^{\sigma/2}$, whose action on functions of $x = (x', x^k, x'')$, $x' \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_{k-1}}$, $x'' \in \mathbb{R}^{n_{k+1}} \times \cdots \times \mathbb{R}^{n_M}$, is defined by

$$(-\Delta_{x^k})^{\sigma/2} u(x) = c_{n_k,\sigma} \int_{\mathbb{R}^{n_k}} \left(u(x', x^k, x'') - \frac{u(x', x^k + y^k, x'') + u(x', x^k - y^k, x'')}{2} \right) |y^k|^{-n_k - \sigma} \, dy^k.$$

The normalization constant $c_{n_k,\sigma}$ is chosen so that the symbol of that operator is $m_k(\xi) = |\xi^k|^\sigma$, see (2.1), and thus the symbol of $\mathcal{L}$ is

$$m(\xi) = \sum_{k=1}^{M} |\xi^k|^\sigma, \qquad \xi = (\xi^1, \cdots, \xi^M), \quad \xi^k \in \mathbb{R}^{n_k}.$$

The spectral measure of $\mathcal{L}$ is

$$d\mu(\theta) = \sum_{k=1}^{M} c_{n_k,\sigma} \delta_{\Omega_k}(\theta), \qquad \Omega_k = \mathbb{S}_+^{N-1} \cap \mathbb{R}^{n_k}.$$

The most relevant case is when $\mathcal{L}$ is the sum of fractional Laplacians of dimension one, cf. (1.8), for which the spectral measure is $d\mu(\theta) = c_{1,\sigma} \sum_{j=1}^{N} \delta_{e_j}(\theta)$, where $\{e_j\}_{j=1}^{N}$ is the canonical basis in $\mathbb{R}^N$. Actually we have

$$\int_{\mathbb{S}^{N-1}} \int_0^\infty \left(u(x) - \frac{u(x+r\theta) + u(x-r\theta)}{2} \right) \frac{dr}{r^{1+\sigma}} \, d\mu(\theta)$$

$$= c_{1,\sigma} \sum_{j=1}^{N} \int_0^\infty \left(u(x) - \frac{u(x+re_j) + u(x-re_j)}{2} \right) \frac{dr}{r^{1+\sigma}} = \sum_{j=1}^{N} (-\partial_{x_j x_j}^2)^{\sigma/2} u(x).$$

The operator (1.13) is the infinitesimal generator of the Lévy process in $\mathbb{R}^N$ given by $X = \{X_t\}_{t \geq 0}$, with $X_t = (X_t^1, \ldots, X_t^M)$, and X_t^k being independent symmetric stable processes in dimension n_k. The kernel associated to these processes has a profile in separated variables,

$$\Phi(x) = \prod_{k=1}^{M} \Psi_k(x^k), \tag{2.6}$$

where Ψ_k is the profile of the kernel corresponding to $(-\Delta_{x^k})^{\sigma/2}$. This kernel is explicit only when $\sigma = 1$, $\Psi_k(w) = d_{n_k}(1 + |w|^2)^{-\frac{n_k+1}{2}}$. In this particular case, if we let x tend to infinity along one of the axes $x_j \in x^k$ (see notation below), then $\Phi(x) \sim \Psi_k(x^k) \sim |x^k|^{-n_k-1}$. Thus, the first estimate in (1.7) is not satisfied. The same happens for any $\sigma \in (0, 2)$. This example motivates the use of estimates on average on $\mathbb{S}^{N-1}$; see [14].

The proof of (1.12) when $\mathcal{L}$ is given by (1.13) relies on an explicit calculation and an estimate of the kernels Ψ_k.

Proposition 2.1. *The profile Φ in* (2.6) *satisfies*

$$|\mathcal{L}\Phi(x)|, \ |\nabla\Phi(x)|, \ |\nabla\mathcal{L}\Phi(x)| \le c\Phi(x). \tag{2.7}$$

Proof. First of all we observe that each Ψ_k is radial, so that by Fourier transform as in [14], see also [17], we have for every $w \in \mathbb{R}^{n_k}$,

$$c_1(1+|w|^2)^{-\frac{n_k+\sigma}{2}} \le \Psi_k(w) \le c_2(1+|w|^2)^{-\frac{n_k+\sigma}{2}},$$

$$|L_k\Psi_k(w)| \le c(1+|w|^2)^{-\frac{n_k+\sigma}{2}}, \tag{2.8}$$

$$|\nabla\Psi_k(w)|, \ |\nabla L_k\Psi_k(w)| \le c(1+|w|^2)^{-\frac{n_k+\sigma+1}{2}}.$$

We have denoted $L_k = (-\Delta_{x^k})^{\sigma/2}$. Therefore (2.7) holds for each factor in the product (2.6). We also use the following convention

$$x = (x_1, x_2, \cdots, x_N) = (x^1, x^2, \cdots, x^M), \qquad x^j = (x_1^j, x_2^j, \cdots, x_{n_j}^j),$$

$$x_m \in x^j \ \Leftrightarrow \ x_m = x_\ell^j \text{ for some } \ell \in \{1, 2, \cdots, n_j\}.$$

We now calculate

$$\partial_{x_m}\Phi(x) = \partial_{x_m}\Psi_j(x^j)\prod_{i \ne j}\Psi_i(x^i) = \frac{\partial_{x_m}\Psi_j(x^j)}{\Psi_j(x^j)}\Phi(x) \quad \text{if } x_m \in x^j,$$

$$\mathcal{L}\Phi(x) = \sum_{k=1}^M L_k\Psi_k(x^k)\prod_{i \ne k}\Psi_i(x^i) = \sum_{k=1}^M \frac{L_k\Psi_k(x^k)}{\Psi_k(x^k)}\Phi(x),$$

$$\partial_{x_m}\mathcal{L}\Phi(x) = \partial_{x_m}\left(\frac{L_j\Psi_j(x^j)}{\Psi_j(x^j)}\right)\Phi(x) + \sum_{k=1}^M \frac{L_k\Psi_k(x^k)}{\Psi_k(x^k)}\partial_{x_m}\Phi(x)$$

$$= \left(\frac{\partial_{x_m}L_j\Psi_j(x^j)}{\Psi_j(x^j)} - L_j\Psi_j(x^j)\frac{\partial_{x_m}\Psi_j(x_j)}{\Psi_j^2(x^j)} + \frac{\partial_{x_m}\Psi_j(x^j)}{\Psi_j(x^j)}\sum_{k=1}^M \frac{L_k\Psi_k(x^k)}{\Psi_k(x^k)}\right)\Phi(x).$$

Therefore,

$$\nabla\mathcal{L}\Phi = \left(\left\{\frac{\nabla L_j\Psi_j}{\Psi_j} - \frac{L_j\Psi_j\nabla\Psi_j}{\Psi_j^2} + \frac{\nabla\Psi_j}{\Psi_j}\sum_{j=1}^M \frac{L_k\Psi_k}{\Psi_k}\right\}\bigg|_{j=1}^M\right)\Phi.$$

We conclude that each coefficient of Φ in the above derivatives is bounded. $\quad\square$

Remark. Actually, estimate (2.8) implies a sharper estimate for the gradient of Φ,

$$|\nabla\Phi(x)|^2 = \sum_{j=1}^M \frac{|\nabla\Psi_j(x^j)|^2}{\Psi_j^2(x^j)}\Phi^2(x) \le c\sum_{j=1}^M \frac{1}{(1+|x^j|^2)}\Phi^2(x).$$

This gives $|\nabla\Phi(x)| \le c|x|^{-N-\sigma-1}$, as in the radial case, provided $|x|$ is large with $|x^j| \sim |x^k|$ for every j, k. In the same way $|\nabla\mathcal{L}\Phi(x)| \le c|x|^{-N-\sigma-1}$ for those directions.

We have now the ingredients to prove Theorem 1.3.

Proof of Theorem 1.3. We estimate the difference within the integral (1.12) by the Mean Value Theorem. Thanks to Proposition 2.1 this amounts to estimate $\int_{\mathbb{S}^{N-1}}\Phi(r\theta - \lambda s\varphi)\,d\theta$, $0 \le \lambda \le 1$, where λ may depend on φ. In order to use now the estimate on average (1.9), which would conclude the proof, we must check that we can replace $r\theta - \lambda s\varphi$ in the integral by $r\theta$. If this were the case

$$\int\limits_{\mathbb{S}^{N-1}} |\mathcal{L}\Phi(r\theta) - \mathcal{L}\Phi(r\theta - s\varphi)| \, d\theta \; = s \int\limits_0^1 \int\limits_{\mathbb{S}^{N-1}} |\nabla\mathcal{L}\Phi(r\theta - \lambda s\varphi)| \, d\theta d\lambda \le c\delta \int\limits_0^1 \int\limits_{\mathbb{S}^{N-1}} |\Phi(r\theta - \lambda s\varphi)| \, d\theta d\lambda$$

$$\le c\delta \int\limits_{\mathbb{S}^{N-1}} |\Phi(r\theta)| \, d\theta d \le c\delta r^{-N-\sigma}.$$

So it is enough to prove that $\Phi(z) \le c\Phi(x)$ whenever $|x - z| < \delta \le 1$ and $|x| > 2\delta$. In fact, since $|x - z|^2 = \sum_{k=1}^M |x^k - z^k|^2$, we have $|x^k - z^k| < \delta$ for every $k = 1, \cdots, M$. If $|x^k| > 2\delta$ this implies $|z^k| > \frac{|x^k|}{2}$, and thus $\Psi_k(z^k) \le \frac{2^{n_k+\sigma} c_2}{c_1} \Psi(x^k)$ by (2.8). On the other hand, if $|x^k| \le 2\delta \le 2$ we have, again by (2.8), $\Psi_k(z^k) \le \frac{c_2 5^{\frac{n_k+\sigma}{2}}}{c_1} \Psi(x^k)$. The claim is proved by multiplying all the factors in k, and so is the theorem. $\square$

Remark. For general operators of the form (1.5), even if we had an estimate like $|\nabla\mathcal{L}\Phi(x)| \le c\Phi(x)$, this would not imply directly (1.12) as in the previous case, where the special form (2.6) was used, and further investigation would be needed.

3. The homogeneous problem

We consider in this section problem (1.1) with $f \equiv 0$ and prove existence and uniqueness of a very weak solution for every initial datum $u_0 \in L^p(\mathbb{R}^N)$ for some $1 \le p \le \infty$. To define such concept of solution we consider the weighted space $L_\rho = L^1(\mathbb{R}^N, \rho \, dx)$ with weight $\rho(x) = (1 + |x|)^{-(N+\sigma)}$. We say that $u \in L^1_{loc}((0, \infty) : L_\rho)$ is a *very weak solution* to problem (1.1) with $f \equiv 0$ if

$$\int\limits_Q u(\partial_t \zeta - \mathcal{L}\zeta) \, dx dt + \int\limits_{\mathbb{R}^N} u_0(x)\zeta(x, 0) \, dx = 0 \tag{3.1}$$

for all $\zeta \in C_0^\infty(\overline{Q})$. The introduction of the weighted space L_ρ allows for the term $\int_Q u\mathcal{L}\zeta$ to be well defined, due to the decay of $\mathcal{L}\zeta$. In fact by a classical result on Fourier Analysis, $m(\xi) \sim |\xi|^\sigma$ implies $\mathcal{L}\zeta(x) = O(|x|^{-(N+\sigma)})$ for large $|x|$.

We will show that test functions which are not compactly supported are also admissible, provided they have a minimal decay at infinity. In order to prove this assertion we will use the formula contained in the next proposition, which follows easily from a direct computation.

Proposition 3.1. *For every pair* $v, w \in L_\rho$,

$$\mathcal{L}(vw) = v\mathcal{L}(w) + w\mathcal{L}(v) - E(v, w),$$

where

$$E(v, w)(x) = \frac{1}{2} \int\limits_{\mathbb{R}^N} \Big(v(x + y) - v(x)\Big)\Big(w(x + y) - w(x)\Big) d\nu(y)$$

$$+ \frac{1}{2} \int\limits_{\mathbb{R}^N} \Big(v(x) - v(x - y)\Big)\Big(w(x) - w(x - y)\Big) d\nu(y). \tag{3.2}$$

Observe that if the measure ν were symmetric this expression would simplify to

$$E(v, w)(x) = \int\limits_{\mathbb{R}^N} \Big(v(x + y) - v(x)\Big)\Big(w(x + y) - w(x)\Big) d\nu(y),$$

formula that appears in [1] for the case of the fractional Laplacian.

Proposition 3.2. *Let u be a very weak solution to problem* (1.1) *with* $f \equiv 0$. *Let also* $\varphi \in C^\infty(Q)$ *be a function that vanishes for* $t > t_0$ *for some* $t_0 > 0$, *and satisfies* φ, $|\nabla \varphi| \le c\rho$ *in* $\mathbb{R}^N \times [0, t_0]$. *Then identity* (3.1) *holds with* ζ *replaced by* φ.

Proof. We multiply φ by a sequence of cut-off functions, use identity (3.1) with these admissible test functions and pass to the limit.

Let then $\phi \in C_0^\infty(\mathbb{R}_+)$ be a nonincreasing function such that $\phi \equiv 1$ for $0 \le s \le 1/2$ and $\phi \equiv 0$ for $s \ge 1$, and define the function $\phi_R(x) = \phi(R^{-1}|x|)$. We are done if we show that

$$\lim_{R \to \infty} \int_{\mathbb{R}^N} u \mathcal{L}(\varphi \phi_R) = \int_{\mathbb{R}^N} u \mathcal{L} \varphi \tag{3.3}$$

for each fixed time $0 < t < t_0$. In order to do that we need to compute the action of $\mathcal{L}$ on the product $\varphi \phi_R$. Since the bilinear form $E(\varphi, \phi_R)$ only involves products of differences, see (3.2), using the same proof as in [1] we obtain

$$\lim_{R \to \infty} \int_{\mathbb{R}^N} u E(\varphi, \phi_R) = 0.$$

The main point is the hypothesis φ, $|\nabla \varphi| \le c\rho$. Recall finally that we have $\|\mathcal{L}\phi\|_\infty \le c(\|\phi\|_\infty + \|D^2\phi\|_\infty)$, so that by homogeneity, and the fact that $u\varphi \in L^1(\mathbb{R}^N)$,

$$\lim_{R \to \infty} \int_{\mathbb{R}^N} u\varphi \mathcal{L} \phi_R = \lim_{R \to \infty} R^{-\sigma} \int_{\mathbb{R}^N} u\varphi \mathcal{L}\phi = 0.$$

We therefore get (3.3). □

Theorem 3.1. *If* $u_0 \in L^p(\mathbb{R}^N)$ *for some* $1 \le p \le \infty$, *then problem* (1.1) *with* $f \equiv 0$ *has a unique very weak solution. The solution is bounded and* C^∞ *smooth for every* $t > 0$ *and satisfies the equation in the classical sense.*

Proof. Existence follows easily by convolution with the heat kernel, $u = u_0 * P$. Thus we deduce the same standard smoothing effect as for the solutions of the fractional heat equation, or even the local heat equation: $u(\cdot, t) \in C^\infty(\mathbb{R}^N) \cap L^q(\mathbb{R}^N)$ for every $p \le q \le \infty$ and any $t > 0$, with

$$\|u(\cdot, t)\|_q \le ct^{-\frac{N}{\sigma}(\frac{1}{p} - \frac{1}{q})} \|u_0\|_p.$$

In order to prove uniqueness we just consider the case $u_0 \equiv 0$, then take R_0, $t_0 > 0$ arbitrary and show that

$$\int_Q u(x, t) F(x, t) \, dx dt = 0 \quad \text{for all } F \in C_0^\infty(\{|x| < R_0, \, 0 < t < t_0\}). \tag{3.4}$$

We use Hilbert's duality method by considering as test function in the definition of very weak solution the unique solution φ to the nonhomogeneous backward problem

$$\partial_t \varphi - \mathcal{L}\varphi = F, \quad x \in \mathbb{R}^N, \, 0 < t < t_0, \qquad \varphi = 0, \quad x \in \mathbb{R}^N, \, t \ge t_0. \tag{3.5}$$

This would yield (3.4) once we check that φ is a good test function. Though the fact that $\mathcal{L}$ is a non-local operator implies that φ does not have compact support in $\overline{Q}$, Proposition 3.2 allows to use it as a test function provided $\varphi \le c\rho$ and $|\nabla \varphi| \le c\rho$.

Using Duhamel's formula, a solution to (3.5) can be written using the heat kernel P in the form

$$\varphi(x, t) = \int_0^{t_0 - t} \int_{\mathbb{R}^N} P(x - y, t_0 - t - s) F(y, s) \, dy ds.$$

By Young's inequality we have $\|\varphi(\cdot, t)\|_\infty \le \|F\|_\infty$. On the other hand, since F has compact support, taking $|x| > 2R_0$, we have, using the self-similar form of P and (1.9),

$$|\varphi(x,t)| \quad \le \|F\|_\infty \int_0^{t_0-t} \int_{|y|<R_0} P(x-y, t_0-t-s)\, dy ds \le \|F\|_\infty \int_0^{t_0-t} \int_0^{R_0} \int_{\mathbb{S}^{N-1}} P(x-r\theta, s) r^{N-1}\, dr d\theta ds$$

$$\le c\|F\|_\infty \int_0^{t_0-t} \int_0^{R_0} s^{-\frac{N}{\sigma}} \left(|x-r\theta|s^{-\frac{1}{\sigma}}\right)^{-N-\sigma} r^{N-1}\, dr ds \le c\|F\|_\infty (t_0-t)^2 \int_0^{R_0} \frac{r^{N-1}}{|x-r\theta|^{N+\sigma}}\, dr$$

$$\le \frac{c\|F\|_\infty (t_0-t)^2 R_0^N}{(1+|x|^2)^{\frac{N+\sigma}{2}}},$$

since $|x-r\theta| > |x|/2 > c(1+|x|^2)^{\frac{1}{2}}$. In the same way we estimate $|\nabla\varphi|$, this time in terms of $\|\nabla F\|_\infty$. We end the proof as follows: use identity (3.1) with $u_0 \equiv 0$ and test function $\zeta = \varphi$ solution to problem (3.5), which gives (3.4) and thus $u \equiv 0$. $\quad\square$

4. The problem with reaction

We consider here the Cauchy problem (1.1) with a nontrivial right-hand side. Since the equation is linear, thanks to the previous section we may assume without loss of generality that $u_0 \equiv 0$. We define a *very weak solution* to problem (1.1) (with $u_0 \equiv 0$) as a function $u \in L^1_{loc}([0,\infty) : L_\rho)$ such that

$$\int_Q u \partial_t \zeta = \int_Q (u-f)\mathcal{L}\zeta \quad \text{for all } \zeta \in C_0^\infty(\overline{Q}). \tag{4.1}$$

4.1. σ–parabolic distance

We introduce now a topology adapted to the equation, in terms of which the estimates are easier to write. This notation has already been used in the literature; see for instance [17]. In order to reflect the different influence of the variables in the equation, we use a σ-parabolic "distance" $|Y_1 - Y_2|_\sigma$ between points $Y_1, Y_2 \in Q$, derived from the σ-parabolic "norm" defined by

$$|Y|_\sigma := \left(|x|^2 + |t|^{2/\sigma}\right)^{1/2} = |t|^{1/\sigma}(|z|^2+1)^{1/2}, \qquad Y = (x,t) \in \mathbb{R}^N \times \mathbb{R}, \quad z = x|t|^{-1/\sigma}.$$

We clearly have $|Y| \le |Y|_\sigma^{\min\{\sigma,\,1\}}$. The Hölder space $C_\sigma^\alpha(Q)$, $\alpha \in (0,\nu)$, will consist of functions u defined in Q such that for some constant $c > 0$

$$|u(Y_1) - u(Y_2)| \le c|Y_1 - Y_2|_\sigma^\alpha \quad \text{for every } Y_1, Y_2 \in Q.$$

The σ-parabolic ball is defined as $B_R := \{Y \in \mathbb{R}^{N+1} : |Y|_\sigma < R\}$. It is also useful to write each point $x \in \mathbb{R}^N$ in polar coordinates, $x = r\theta$, $r = |x|$, $\theta = x/r$. In that way, to each point $Y = (x,t) = (r\theta, t) \in \mathbb{R}^N \times \mathbb{R}$ we associate the point $\widetilde{Y} = (r,t) \in \mathbb{R}_+^2$ and write, by abuse of notation, $Y = (\widetilde{Y}, \theta)$. Observe that, again abusing notation,

$$|Y|_\sigma = |\widetilde{Y}|_\sigma = (r^2 + |t|^{2/\sigma})^{1/2}.$$

Let us also consider the ball $\widetilde{B}_R := \{\widetilde{Y} \in \mathbb{R}_+^2 : |\widetilde{Y}|_\sigma < R\}$. We have that the integrals in σ-parabolic balls can be decomposed as

$$\int_{B_R} w(Y)\, dY = \int_{\mathbb{S}^{N-1}} \int_{\{r^2+|t|^{2/\sigma}<R^2\}} w(r,t,\theta) r^{N-1}\, dr dt d\theta = \int_{\mathbb{S}^{N-1}} \int_{\widetilde{B}_R} w(\widetilde{Y},\theta) r^{N-1}\, d\widetilde{Y} d\theta.$$

For instance, by using the change of variables

$$s = r|t|^{-1/\sigma}, \qquad \rho = (r^2 + |t|^{2/\sigma})^{1/2}, \tag{4.2}$$

we can obtain

$$\int_{B_R} g(|Y|_\sigma)\,dY = 2\omega_N \int_0^\infty \int_0^R g(\rho)\left(\frac{s\rho}{(s^2+1)^{1/2}}\right)^{N-1} \frac{\sigma\rho^\sigma}{(s^2+1)^{\frac{\sigma+1}{2}}}\,d\rho\,ds = c\int_0^R g(\rho)\rho^{N+\sigma-1}\,d\rho.$$

In particular, the volume of the ball B_R is proportional to $R^{N+\sigma}$.

We finally write, in terms of the new distance, the estimates for the time derivatives of the Gauss kernel P that can be deduced from the decay estimates for the profile Φ, see Section 2.1. We use self-similarity and the fact that Φ is bounded, so any estimate $cr^{-N-\beta}$ for r large can be written as $c(1+r^2)^{\frac{-N-\beta}{2}}$ for $r>0$. For every $k\in\mathbb{N}$ it holds

$$\begin{aligned}
\int_{\mathbb{S}^{N-1}} |\partial_t^k P(r\theta,t)|\,d\theta &= \int_{\mathbb{S}^{N-1}} |\mathcal{L}^k P(r\theta,t)|\,d\theta = t^{-\frac{N}{\sigma}-k}\int_{\mathbb{S}^{N-1}} |\mathcal{L}^k \Phi(r\theta t^{-\frac{1}{\sigma}})|\,d\theta \\
&\le ct^{-\frac{N}{\sigma}-k}(1+r^2 t^{-\frac{2}{\sigma}})^{\frac{-N-\beta}{2}} \le ct^{-k+\frac{\beta}{\sigma}}|\widetilde{Y}|_\sigma^{-N-\beta},
\end{aligned}$$

(4.3)

for every $0<\beta<\sigma$, $\widetilde{Y}=(r,t)$, $r,t>0$. If $k=0$ it is also true with $\beta=\sigma$.

4.2. A cancellation property

We next show a cancellation property for $\mathcal{L}P$ crucial in later regularity arguments.

Theorem 4.1. *Let P be the Gauss kernel* (2.3). *Then for every $0<a<b$,*

$$\int_{B_{a,b}^+} \mathcal{L}P(Y)\,dY = 0,$$

(4.4)

where $B_{a,b}^+ = \{a<|Y|_\sigma<b,\ t>0\}$.

Proof. Using as before the change of variables (4.2) we get,

$$\begin{aligned}
\int_{B_{a,b}^+} \mathcal{L}P(Y)\,dY &= -\int_{B_{a,b}^+} \partial_t P(Y)\,dY \\
&= \frac{1}{\sigma}\int_{\mathbb{S}^{N-1}}\int_{\widetilde{B}_{a,b}^+} t^{-\frac{N+\sigma}{\sigma}}(N\Phi(r\theta t^{-\frac{1}{\sigma}})+rt^{-1/\sigma}\theta\nabla\Phi(r\theta t^{-1/\sigma}))r^{N-1}\,dr\,dt\,d\theta \\
&= \log(b/a)\int_{\mathbb{S}^{N-1}}\int_0^\infty s^{N-1}(N\Phi(s\theta)+s\theta\partial_s\Phi(s\theta))\,ds\,d\theta \\
&= \log(b/a)\int_{\mathbb{S}^{N-1}}\int_0^\infty \partial_s(s^N\Phi(s\theta))\,ds\,d\theta \\
&= \log(b/a)\lim_{R\to\infty} R^N\int_{\mathbb{S}^{N-1}} \Phi(R\theta))\,d\theta = 0.
\end{aligned}$$

The last limit uses the behaviour (1.9). $\square$

4.3. Existence

We formally write the solution using Duhamel's formula:

$$u(x,t) = \int_0^t \int_{\mathbb{R}^N} P(x-\overline{x}, t-\overline{t})\mathcal{L}f(\overline{x},\overline{t})\,d\overline{x}\,d\overline{t}, \quad (x,t)\in Q.$$

Now integrate by parts and consider the integral in principal value sense (in σ–parabolic topology). We prove that what we obtain is in fact the unique solution to our problem.

Theorem 4.2. *If $f \in C_\sigma^\alpha(Q) \cap L^\infty(Q)$ for some $0 < \alpha < 1$, then the function*

$$u(x,t) = \lim_{\varepsilon \to 0} \int_{\Omega_\varepsilon(x,t)} \mathcal{L}P(x - \overline{x}, t - \overline{t}) f(\overline{x}, \overline{t}) \, d\overline{x} d\overline{t}, \tag{4.5}$$

where $\Omega_\varepsilon(x,t) = \{\varepsilon^2 < |\overline{x} - x|^2 + |\overline{t} - t|^{2/\sigma}, \ 0 < \overline{t} < t\}$, is the unique very weak solution of problem (1.1) with $u_0 \equiv 0$, which is moreover bounded.

Proof. Uniqueness follows from the previous section. Let us show that the function in (4.5) is well defined. Let $Y = (x,t) \in Q$ be fixed and take $0 < \varepsilon < t^{1/\sigma}/2$. We decompose the integral as

$$\int_{\Omega_\varepsilon(x,t)} = \int_{\Omega_\varepsilon(x,t) - \Omega_{t^{1/\sigma}}(x,t)} + \int_{\Omega_{t^{1/\sigma}}(x,t)} = I_1 + I_2.$$

The cancellation property (4.4) implies

$$\int_{\Omega_\varepsilon(x,t) - \Omega_{t^{1/\sigma}}(x,t)} \mathcal{L}P(x - \overline{x}, t - \overline{t}) \, d\overline{x} d\overline{t} = 0.$$

Therefore, the Hölder regularity of f together with estimate (4.3) with $k = 1$ and $\beta < \sigma$ allow us to estimate the inner integral,

$$
\begin{aligned}
|I_1| &= \left| \int_{\Omega_\varepsilon(Y) - \Omega_{t^{1/\sigma}}(Y)} \mathcal{L}P(Y - \overline{Y}) f(\overline{Y}) \, d\overline{Y} \right| \\
&= \left| \int_{\Omega_\varepsilon(Y) - \Omega_{t^{1/\sigma}}(Y)} \mathcal{L}P(Y - \overline{Y}) \Big(f(\overline{Y}) - f(Y) \Big) \, d\overline{Y} \right| \\
&\leq [f]_{C^\alpha} \int_{|Y - \overline{Y}|_\sigma^\sigma < t} |\mathcal{L}P(Y - \overline{Y})| |\overline{Y} - Y|_\sigma^\alpha \, d\overline{Y} = c \int_{\mathbb{S}^{N-1}} \int_{|\widetilde{Y}|_\sigma^\sigma < t} |\mathcal{L}P(\widetilde{Y}, \theta)| |\widetilde{Y}|_\sigma^\alpha r^{N-1} \, d\widetilde{Y} d\theta \\
&\leq c \int_{|\widetilde{Y}|_\sigma^\sigma < t} \tau^{-1 + \beta/\sigma} |\widetilde{Y}|_\sigma^{-N - \beta + \alpha} r^{N-1} \, d\widetilde{Y} = c \int_0^\infty \frac{s^{N-1}}{(s^2 + 1)^{\frac{N+\beta}{2}}} \, ds \int_0^{t^{1/\sigma}} \rho^{\alpha - 1} \, d\rho = c t^{\alpha/\sigma}.
\end{aligned}
$$

We have put $Y - \overline{Y} = (\widetilde{Y}, \theta) = (r, \tau, \theta)$, integrated in the sphere and then used the change of variables (4.2).

We next prove that the outer integral is bounded by using the boundedness of f. Here we integrate first in the sphere, then in the radial variable and finally in time,

$$
\begin{aligned}
|I_2| &\leq \|f\|_\infty \int_{\Omega_{t^{1/\sigma}}(x,t)} |\mathcal{L}P(Y - \overline{Y})| \, d\overline{Y} \leq c \int_{\mathbb{S}^{N-1}} \int_{|\widetilde{Y}|_\sigma^\sigma > t} |\mathcal{L}P(\widetilde{Y}, \theta)| r^{N-1} \mathbb{1}_{\{0 < \tau < t\}} \, d\widetilde{Y} d\theta \\
&\leq c \int_0^t \int_{\sqrt{t^{2/\sigma} - \tau^{2/\sigma}}}^\infty \tau^{-1 + \beta/\sigma} r^{-\beta - 1} \, dr d\tau = c \int_0^t \tau^{-1 + \beta/\sigma} (t^{2/\sigma} - \tau^{2/\sigma})^{-\beta/2} \, d\tau = c.
\end{aligned}
$$

This also gives that the solution is bounded for every bounded interval of times. The fact that u is a very weak solution is immediate. $\square$
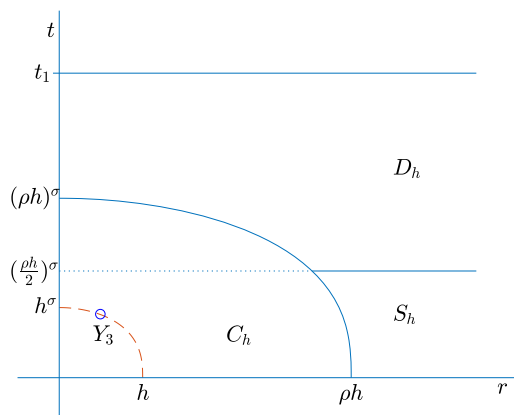
Fig. 1. Integration regions for $Y = (r\theta, t)$ for each $\theta \in \mathbb{S}^{N-1}$ fixed.

4.4. Hölder regularity

We study here the regularity of the function given by formula (4.5), using the notation $Y = (x, t) \in Q$,

$$u(Y) = \int_{\mathbb{R}^{N+1}} A(Y - \overline{Y}) \mathbb{1}_{\{0 < \overline{t} < t\}} f(\overline{Y}) \, d\overline{Y},$$

where $A = \mathcal{L}P$. We omit the principal value sense of the integral for simplicity.

Proof of Theorem 1.2. Let $Y_1 = (x_1, t_1)$, $Y_2 = (x_2, t_2) \in Q$ be two points with $|Y_1 - Y_2|_\sigma = h > 0$ small, and assume for instance $t_1 \geq t_2$. By substracting $f(Y_1)$ to f we may assume without loss of generality that $f(Y_1) = 0$. We must estimate the difference

$$|u(Y_1) - u(Y_2)| = \left| \int_{\mathbb{R}_+^{N+1}} \left(A(Y_1 - \overline{Y}) \mathbb{1}_{\{\overline{t} < t_1\}} - A(Y_2 - \overline{Y}) \mathbb{1}_{\{\overline{t} < t_2\}} \right) f(\overline{Y}) \, d\overline{Y} \right|$$

$$= \left| \int_0^{t_1} \int_{\mathbb{R}^N} \left(A(Y) \mathbb{1}_{\{t > 0\}} - A(Y - Y_3) \mathbb{1}_{\{t > t_3\}} \right) f(Y_1 - Y) \, dY \right|.$$

We have made the change of variables $Y = Y_1 - \overline{Y} = (x, t)$ and put $Y_3 = Y_1 - Y_2 = (x_3, t_3)$, so that $|Y_3|_\sigma = h$, $t_3 = t_1 - t_2 \in [0, h^\sigma]$. Observe that $|f(Y_1 - Y)| \leq c|Y|_\sigma^\alpha$. We decompose $Q_1 = \mathbb{R}^N \times (0, t_1)$ into three regions, depending on the sizes of $|x|$ and t, see Fig. 1,

$$\int_{Q_1} = \int_{C_h} + \int_{S_h} + \int_{D_h}.$$

(i) *The small "semiball"* $C_h = B_{\rho h}^+ = \{|Y|_\sigma < \rho h\} \cap \{t > 0\}$, where $\rho > 2$ is a constant to be fixed later. We take h small enough $((\rho h)^\sigma < \min\{t_1, 1\})$ so that $C_h \subset Q_1$. The difficulty in this region is the non-integrable singularity of $A(Y)$ at $Y = 0$, which is to be compensated by the regularity of f. We first have, repeating the computations of the proof of Theorem 4.2,

$$\int_{C_h} |A(Y)| |f(Y_1 - Y)| \, dY \leq \int_{C_h} |A(Y)| |Y|_\sigma^\alpha \, dY \leq ch^\alpha.$$

As to the second term in $\int_{C_h}$, we use the cancellation property (4.4) in order to counteract the singularity at $Y = Y_3$. Thus, taking $\rho \geq \max\{2, 2^{1/\sigma}\}$ we have

$$B_h^+(Y_3) = \{|Y - Y_3|_\sigma < h, \, t > t_3\} \subset C_h \subset B_{\rho^2 h}^+(Y_3),$$

so that,

$$\left| \int\limits_{C_h} A(Y - Y_3) \mathbb{1}_{\{t > t_3\}} f(Y_1 - Y) \, dY \right|$$

$$\leq \int\limits_{C_h} |A(Y - Y_3) \mathbb{1}_{\{t > t_3\}}| \, \left| f(Y_1 - Y) - f(Y_1 - Y_3) \right| dY + |f(Y_1 - Y_3)| \int\limits_{C_h - B_h(Y_3)} \left| A(Y - Y_3) \mathbb{1}_{\{t > t_3\}} \right| dY \qquad (4.6)$$

$$\leq c \int\limits_{B_{\rho^2 h}^+(Y_3)} |A(Y - Y_3)| \, |Y - Y_3|^\alpha \, dY + ch^\alpha \int\limits_{B_{\rho^2 h}^+(Y_3) - B_h^+(Y_3)} |A(Y - Y_3)| \, dY = I_1 + I_2.$$

The first integral satisfies again $I_1 \leq ch^\alpha$. We now show that I_2 can be controlled since we are far from the singularity. Putting $Z = Y - Y_3$, and using as always the notation in polar coordinates/time, $Z = (s\varphi, \tau) = (\widetilde{Z}, \varphi)$, we have

$$I_2 \leq ch^\alpha \int\limits_{\{\rho h < |Z|_\sigma < \rho^2 h, \, \tau > 0\}} |A(Z)| \, dZ = ch^\alpha \int\limits_{\{\rho h < |\widetilde{Z}|_\sigma < \rho^2 h, \, \tau > 0\}} \int\limits_{\mathbb{S}^{N-1}} |A(\widetilde{Z}, \varphi)| \, d\varphi \, d\widetilde{Z}$$

$$\leq ch^\alpha \int\limits_{\{\rho h \leq |\widetilde{Z}|_\sigma \leq \rho^2 h, \, \tau > 0\}} \tau^{-1 + \frac{\beta}{\sigma}} |\widetilde{Z}|_\sigma^{-N - \beta} \, d\widetilde{Z}.$$

Changing $(s, \tau) \to (w, \xi)$ following (4.2) we end up with the estimate

$$I_2 \leq ch^\alpha \int\limits_0^\infty \frac{w^{N-1}}{(w^2 + 1)^{\frac{N+\beta}{2}}} \, dw \int\limits_{\rho h}^{\rho^2 h} \frac{d\xi}{\xi} = ch^\alpha.$$

(ii) *Outside the ball $B_{\rho h}$ for small times,* $S_h = \{|Y|_\sigma \geq \rho h, \, 0 < t < (\rho h/2)^\sigma\}$. Since in this region we have $|Y|_\sigma \leq \rho_1 |Y - Y_3|_\sigma$ for some positive constant ρ_1 depending only on σ, both integrals in S_h are of the same order

$$\int\limits_{S_h} \left(|A(Y)| + |A(Y - Y_3)| \mathbb{1}_{\{t > t_3\}} \right) |f(Y_1 - Y)| \, dY \leq c \int\limits_0^{(\frac{\rho h}{2})^\sigma} \int\limits_0^\infty t^{-1 + \frac{\beta}{\sigma}} (r^2 + t^{\frac{2}{\sigma}})^{\frac{-N - \beta + \alpha}{2}} r^{N-1} \, dr \, dt$$

$$\leq c \int\limits_0^{(\frac{\rho h}{2})^\sigma} t^{-1 + \frac{\alpha}{\sigma}} \, dt \int\limits_0^\infty \frac{w^{N-1}}{(w^2 + 1)^{\frac{N + \beta - \alpha}{2}}} \, dw = ch^\alpha.$$

Notice that $\alpha < \sigma$ so the last integral is convergent.

(iii) *Outside the ball $B_{\rho h}$ for not so small times,* $D_h = \{|Y|_\sigma \geq \rho h, \, (\rho h/2)^\sigma < t < t_1\}$. Since in that set it is $t > t_3$, we have

$$A(Y) - A(Y - Y_3) \mathbb{1}_{\{t > t_3\}} = A(Y) - A(Y - Y_3),$$

and there will be some cancellation. We put $Y_3^* = (0, t_3)$ and decompose this difference as

$$|A(Y) - A(Y - Y_3)| \leq |A(Y) - A(Y - Y_3^*)| + |A(Y - Y_3^*) - A(Y - Y_3)|.$$

For the first term,

$$|A(Y) - A(Y - Y_3^*)| \leq h^\sigma \|\partial_t A(Z_1)\|,$$

where $Z_1 = (x, t - \lambda t_3)$ for some $\lambda \in (0, 1)$. Observe also that $t - \lambda t_3 \sim t$. Using now (4.3) we have, denoting as always, $Y = (r\theta, t)$,

$$\int\limits_{D_h} |A(Y) - A(Y - Y_3^*)||f(Y_1 - Y)|\, dY \le h^\sigma \int\limits_{D_h} |\partial_t A(Z_1)||f(Y_1 - Y)|\, dY$$

$$\le ch^\sigma \int\limits_{(\rho h/2)^\sigma}^{t_1} (t - \lambda t_3)^{-\frac{N}{\sigma} - 2} \int\limits_0^\infty \int\limits_{\mathbb{S}^{N-1}} |\mathcal{L}^2 \Phi(r\theta(t - \lambda t_3)^{-\frac{1}{\sigma}})|(r^2 + t^{\frac{2}{\sigma}})^{\frac{\alpha}{2}} r^{N-1}\, d\theta dr dt$$

$$\le ch^\sigma \int\limits_{(\rho h/2)^\sigma}^{t_1} t^{-2 + \frac{\beta}{\sigma}} \int\limits_0^\infty (r^2 + t^{\frac{2}{\sigma}})^{\frac{-N-\beta+\alpha}{2}} r^{N-1}\, dr dt = ch^\sigma \int\limits_{(\rho h/2)^\sigma}^{t_1} t^{-2 + \frac{\alpha}{\sigma}} \int\limits_0^\infty \frac{w^{N-1}}{(w^2 + 1)^{\frac{N+\beta-\alpha}{2}}}\, dw dt \le ch^\alpha.$$

The last integral is convergent provided $\alpha < \beta$. We now estimate the spatial difference,

$$\int\limits_{D_h} |A(Y - Y_3^*) - A(Y - Y_3)||f(Y_1 - Y)|\, dY \le c \int\limits_{(\rho h/2)^\sigma}^{t_1} (t - t_3)^{-\frac{N}{\sigma} - 1} \int\limits_0^\infty I(r, t)(r^2 + t^{\frac{2}{\sigma}})^{\frac{\alpha}{2}} r^{N-1}\, dr dt$$

where

$$I(r, t) = \int\limits_{\mathbb{S}^{N-1}} |\mathcal{L}\Phi(r\theta(t - t_3)^{-\frac{1}{\sigma}}) - \mathcal{L}\Phi((r\theta - s\varphi)(t - t_3)^{-\frac{1}{\sigma}})|\, d\theta, \qquad s\varphi = x_3.$$

Using hypothesis (1.12),

$$I(r, t) \le cs(t - t_3)^{-\frac{1}{\sigma}} (r(t - t_3)^{-\frac{1}{\sigma}})^{-N-\beta} \le cht^{\frac{N+\beta-1}{\sigma}} (r^2 + t^{\frac{2}{\sigma}})^{-\frac{N+\beta}{2}}.$$

Therefore

$$\int\limits_{D_h} |A(Y - Y_3^*) - A(Y - Y_3)||f(Y_1 - Y)|\, dY \le ch \int\limits_{(\rho h/2)^\sigma}^{t_1} \int\limits_0^\infty t^{-1 - \frac{1-\beta}{\sigma}} (r^2 + t^{\frac{2}{\sigma}})^{\frac{-N-\beta+\alpha}{2}} r^{N-1}\, dr dt$$

$$= ch \int\limits_{(\rho h/2)^\sigma}^{t_1} t^{-1 - \frac{1-\alpha}{\sigma}} \int\limits_0^\infty \frac{w^{N-1}}{(w^2 + 1)^{\frac{N+\beta-\alpha}{2}}}\, dw dt \le ch^\alpha.$$

As before we need $\alpha < \beta$. The proof is finished. $\quad\square$

We end with a modification of the previous proof by assuming that the datum f is $C^{\alpha+\epsilon}$ Hölder continuous at some point Y_1 and only C^α, but with a small coefficient, at the rest of the points, thus getting $C^{\alpha+\epsilon}$ regularity at Y_1.

Theorem 4.3. *In the hypotheses of Theorem 1.2, assume moreover that there exist $c > 0$, $\delta_0 > 0$ and $\epsilon > 0$, $\alpha + \epsilon < \max\{\sigma, 1\}$, such that*

$$|f(Y) - f(Y_1)| \le c|Y - Y_1|_\sigma^{\alpha+\epsilon}, \tag{4.7}$$

$$|f(Y) - f(\overline{Y})| \le c\delta^\epsilon |Y - \overline{Y}|_\sigma^\alpha, \tag{4.8}$$

for all $0 < \delta < \delta_0$, $Y, \overline{Y} \in B_\delta(Y_1)$. Then,

$$|u(Y) - u(Y_1)| \le c'|Y - Y_1|_\sigma^{\alpha+\epsilon},$$

for all $Y \in B_{\delta_0/2}(Y_1)$, where c' depends on c.

Proof. Since f is bounded, condition (4.7) holds for every $Y \in Q$. This is enough to make all the estimates used to prove Theorem 1.2 work, yielding terms which are $O(h^{\alpha+\epsilon})$, except that for the integral I_1 in (4.6). To estimate this term, take $\rho h < \delta_0$ and observe that (4.8) gives

$$I_1 = \int\limits_{C_h} |A(Y - Y_3)\mathbb{1}_{\{t > t_3\}}| \big|f(Y_1 - Y) - f(Y_1 - Y_3)\big|\, dY \le ch^\epsilon \int\limits_{C_h} \frac{1}{|Y - Y_3|_\sigma^{N+\sigma}} |Y - Y_3|_\sigma^\alpha\, dY \le ch^{\alpha+\epsilon}. \quad\square$$

This theorem will be used somewhere else to study the regularity of solutions to nonlinear nonlocal equations.

Declaration of competing interest

There is no competing interest.

Acknowledgements

References

[1] B. Barrios, I. Peral, F. Soria, E. Valdinoci, A Widder's type theorem for the heat equation with nonlocal diffusion, Arch. Ration. Mech. Anal. 213 (2) (2014) 629–650.

[2] R.M. Blumenthal, R.K. Getoor, Some theorems on stable processes, Trans. Am. Math. Soc. 95 (2) (1960) 263–273.

[3] K. Bogdan, P. Sztonyk, Estimates of the potential kernel and Harnack's inequality for the anisotropic fractional Laplacian, Stud. Math. 181 (2) (2007) 101–123.

[4] K. Bogdan, P. Sztonyk, V. Knopova, Heat kernel of anisotropic nonlocal operators, preprint, arXiv:1704.03705.

[5] M. Felsinger, M. Kassmann, Local regularity for parabolic nonlocal operators, Commun. Partial Differ. Equ. 38 (9) (2013) 1539–1573.

[6] X. Fernández-Real, X. Ros-Oton, Regularity theory for general stable operators: parabolic equations, J. Funct. Anal. 272 (10) (2017) 4165–4221.

[7] P. Glowacki, Lipschitz continuity of densities of stable semigroups of measures, Colloq. Math. 66 (1) (1993) 29–47.

[8] P. Glowacki, W. Hebisch, Pointwise estimates for densities of stable semigroups of measures, Stud. Math. 104 (3) (1993) 243–258.

[9] M. Hairer, Rough stochastic PDEs, Commun. Pure Appl. Math. 64 (11) (2011) 1547–1585.

[10] M. Kassmann, R.W. Schwab, Regularity results for nonlocal parabolic equations, Riv. Mat. Univ. Parma 5 (2014) 183–212.

[11] T. Kulczycki, M. Ryznar, Gradient estimates of Dirichlet heat kernels for unimodal Lévy processes, Math. Nachr. 291 (2–3) (2018) 374–397.

[12] F. Otto, J. Sauer, S. Smith, H. Weber, Parabolic equations with rough coefficients and singular forcing, preprint, arXiv:1803.07884v2.

[13] A. de Pablo, F. Quirós, A. Rodríguez, J.L. Vázquez, Classical solutions for a logarithmic fractional diffusion equation, J. Math. Pures Appl. (9) 101 (6) (2014) 901–924.

[14] W.E. Pruitt, S.J. Taylor, The potential kernel and hitting probabilities for the general stable process in R^N, Trans. Am. Math. Soc. 146 (1969) 299–321.

[15] X. Ros-Oton, Nonlocal elliptic equations in bounded domains: a survey, Publ. Mat. 60 (1) (2016) 3–26.

[16] G. Samorodnitsky, M.S. Taqqu, Stable Non-Gaussian Random Processes. Stochastic Models with Infinite Variance, Stochastic Modeling, Chapman & Hall, New York, 1994.

[17] J.L. Vázquez, A. de Pablo, F. Quirós, A. Rodríguez, Classical solutions and higher regularity for nonlinear fractional diffusion equations, J. Eur. Math. Soc. 19 (7) (2017) 1949–1975.

[18] T. Watanabe, Asymptotic estimates of multi-dimensional stable densities and their applications, Trans. Am. Math. Soc. 359 (6) (2007) 2851–2879.